Development of a GIS Model to Predict Muskellunge Spawning Habitat in Northern Wisconsin Lakes

Joel Nohner

A thesis submitted
in partial fulfillment of the requirements
for the degree of
Master of Science
in the School of Natural Resources and Environment
of the University of Michigan
April 2009

Thesis Committee:
Dr. James S. Diana, Chair
Dr. James E. Breck

Acknowledgments

I would like to thank the Wisconsin Department of Natural Resources, Musky Clubs Alliance of Wisconsin, and the University of Michigan for funding this project. Thank you also to the Alvan Macauley Fellowship, which partially supported my graduate education.

I would especially like to thank Dr. Jim Diana for his guidance, enthusiasm, kindness, and patience. I am grateful to committee members Dr. Jim Diana and Dr. Jim Breck for providing a thoughtful commentary and critique for this thesis. Thanks also to Michelle Lelli, Sean Ledwin, and Barb Diana for their comments on drafts for this paper.

This research project would not have been completed without the effort of over fifty individuals logging over 2,000 volunteer hours. Thank you to Bob Haase and the Musky Clubs Alliance of Wisconsin volunteers. A special thanks to Tom Aepelbacher, Kyle Battige, Jason Bies, Leah Ettema, Corrine Higley, Andy Layman, Sean Lewandoski, Dan Oele, Dana Rudy, Jon Schultz, and Sam Upton, all of whom contributed up to 18 hour days for almost a month to this project. Thanks also to Paul Cunningham, Tim Simonson, and Steve Bolssen, of the Wisconsin Department of Natural Resources. I would also like to thank the Kemp Natural Resources Station staff for their hospitality and logistical support.

I have had the fortune of being inspired in the classroom and the field by group of premier scientists. For their dedication to teaching and ecology, I thank Dr. Jim Diana, Dr. Jennifer Tank, Dr. Doug Wilcox, Dr. Mike Wiley, Dr. Gary Lamberti, Dr. David Lodge, Dr. Gary Belovsky, Dr. George Kling, and Dr. Sally Entrekin.

Finally, I would like to thank my family, Kevin, Mary, Emily, Nick, and Mitch Nohner, for their love, support, and inspiration. I was introduced to the outdoors at a young age, and owe my passion to the outdoor experiences they have shared with me.

Table of Contents

Page

Acknowledgments …….…….…… ii

List of Tables……..……..iv

List of Figures……v

Abstract…….….1

Introduction……....2

Methods…….…….3

Results……...….8

Discussion……11

References……..…..19

Tables……...……25

Figures……..……29

List of Tables

Page

Table 1. A summary of lakes used in this study ……..... 25

Table 2. Characteristics used to categorize submersed aquatic vegetation in the study lakes …….....26

Table 3. Characteristics used to categorize emergent aquatic vegetation in the study lakes ……....26

Table 4. Results from the χ^2 analyses for differences between spawning site and available habitat frequencies for each variable.……...27

Table 5. Summary of muskellunge spawning habitat preferences calculated from spawning site and available habitat data ...……..……28

List of Figures

Page

Figure 1. Locations of study lakes in northern Wisconsin, USA….....29

Figure 2. Conceptual diagram of GIS metrics…….……29

Figure 3. Training gain for MSHM1 and MSHM2……...……..30

Figure 4. Probability of spawning from MSHM1 for each variable…….….…….31

Figure 5. Probability of spawning from MSHM2 for each variable…….….…….32

Figure 6. Training and test area under the curve analyses for MSHM1 and MSHM2...33

Figure 7. MSHM1 applied to Birch Lake, Vilas Co., WI.……...…...……34

Figure 8. MSHM2 applied to Birch Lake, Vilas Co., WI……...……35

Abstract

This study determined the habitat preferences of spawning muskellunge in northern Wisconsin lakes and used these preferences to create two GIS-based models that predict the location of muskellunge spawning habitat. This information will enable efficient conservation of muskellunge spawning habitat, which has been implicated in declining natural reproduction. Muskellunge spawning sites were identified using spotlighting surveys and verified by the presence of muskellunge eggs. Aquatic vegetation and substrate maps were created using visual surveys to determine habitat preference and train the models. Vegetation was categorized structurally, and muskellunge preferred to spawn over emergent sedges and rushes as well as submersed short grasses and mat-forming vegetation. Muskellunge preferred sand, cobble, and coarse benthic organic matter substrates, areas with high potential groundwater flow, and areas adjacent to wetlands. Moderate to steep slopes were preferred for spawning, as were locations near bays and points. While shorelines facing east to north-east and south to south-west were slightly preferred, the biological connection to this pattern is likely tenuous. Muskellunge Spawning Habitat Models (MSHM) 1 and 2 were created using the Maxent modeling program. The models utilized the difference between characteristics of spawning sites and available habitat to assign probabilities of spawning across each variable. These probabilities were in general agreement with the spawning habitat preferences documented in this and other studies. While MSHM1 uses only data which can be obtained remotely in Wisconsin, MSHM2 utilizes low-cost habitat surveys to slightly improve model performance. MSHM1 and MSHM2 were tested by withholding 25% of the spawning sites from model training for testing. Both models performed significantly better than random at predicting spawning locations using a binomial test, and the area under the curve analyses are evidence that each model possesses reasonable efficiency. The models assign a probability of muskellunge spawning to cells in a raster grid, and these values can be used to rank the best spawning habitat in each lake. For example, using either MSHM1 or MSHM2, a manager could identify the best 10% of available habitat and protect approximately half of the muskellunge spawning sites. MSHM2, which includes variables from habitat surveys, appears to outperform MSHM1 in identifying the top 10% of available habitat. The muskellunge spawning habitat preferences identified by this study can inform habitat conservation and restoration. The spawning habitat models identify the locations of likely spawning habitat, allowing managers to efficiently protect these critical areas from the removal of vegetation and woody debris which muskellunge preferred for spawning.

Introduction

The muskellunge *Esox masquinongy* is an ecologically and economically important species throughout its range, acting as a top predator in aquatic ecosystems and driving multi-million dollar state fishing industries (Menz and Wilton 1983; Younk and Cook 1992; Bozek et al. 1999). However, the species faces an uncertain future due to declines in natural reproduction throughout its native range (Dombeck et al. 1986, Inskip 1986). Stocking has countered these losses by supplementing or replacing natural reproduction. However, stocking is expensive (Margenau 1992), can spread or be affected by disease, and influences the genetic composition of individual fish stocks (Miller et al., in press). In the face of increasing human populations and their effects on muskellunge spawning habitat, the most cost effective long term approach will be to preserve natural reproduction.

Research has found that muskellunge home to large spawning grounds (Crossman 1990), but little is known about homing to particular sites, what cues determine site fidelity, at what stage muskellunge potentially imprint upon a site, and whether stocking impacts site fidelity and habitat preferences. Muskellunge typically spawn in shallow (<1.5 m) water near aquatic vegetation and woody habitat, but have also been found at off-shore locations as deep as 3.7 m (Scott and Crossman 1973; Haas 1978; Dombeck et al. 1984; Zorn et al. 1998; Pierce et al. 2007). Muskellunge egg mortality, typically due to hypoxic conditions or *Saprolegnia* spp. fungus, was lower over wood, sand, and gravel substrates than leaves, plants, and silt (Dombeck et al. 1984; Zorn et al. 1998). While individual plant species vary with location, commonly documented vegetation in muskellunge spawning areas includes sedges, *Chara* spp., *Elodea* spp., *Myriophyllum* spp., *Potamogeton* spp., *Najas flexilis,* and *Vallisneria americana* (Craig and Black 1986; Werner et al. 1996).

A number of studies have provided insight to muskellunge spawning habitat requirements, although the methodologies used often limit the interpretation of their results. Typical muskellunge densities and the substantial effort required to identify spawning sites limit the sample size in many studies (e.g., Dombeck et al. 1984; Craig and Black 1986; Miller and Menzel 1986; Strand 1986; Werner et al. 1996; Murry and Farrell 2007; Pierce et al. 2007). Additionally, selecting study sites based on professional judgment has the potential to bias habitat characterizations (e.g., Craig and Black 1986). Methods that fail to statistically compare spawning habitat to available habitat can only make conclusions about habitat usage, not preference (e.g., Craig and Black 1986; Strand 1986; Pierce et al. 2007). Finally, studies at the whole-lake scale on habitat characteristics which contribute to natural reproduction lack the resolution to determine habitat preferences (e.g., Dombeck et al. 1986; Zorn et al. 1998; Rust et al. 2002). This study investigated the habitat preferences of spawning muskellunge by comparing the characteristics of a large number of spawning sites to available habitat. Spawning sites were identified by visual spotlighting surveys conducted from boats at night, covering the entire shallow (<1.5 m) area of each lake. This approach minimized the potential bias toward certain habitat types by surveying the entire shallow area of each lake.

Models for muskellunge spawning habitat have been created (Dombeck et al. 1986; Rust et al. 2002), but are limited to whole-lake classifications of successful natural reproduction. Farmer and Chow-Fraser (2004) created a conceptual model with parameters for three primary requirements for spawning habitat: temperature, dissolved oxygen (DO), and spatial separation of the eggs. While the scientific literature provides a good idea of *what* type of habitats muskellunge need to reproduce successfully, and even *how* to restore these

habitats (e.g., Dombeck 1986), there is no efficient way to determine *where* muskellunge actually spawn in a given lake. This information is critical to both managers and landowners throughout muskellunge's range, where the vital ecological services such as natural reproduction, provided by aquatic vegetation and woody habitat, are being threatened by rapidly expanding human populations (Christensen et al. 1996; Radomski and Goeman 2001; Rust et al. 2002; Peterson et al. 2003).

Advances in Geographic Information Systems (GIS) and machine-learning techniques have enabled the production of advanced spatial models. GIS layers with land cover, bathymetry, and sometimes even aquatic vegetation data are now available to management agencies. These data enable models which move beyond simple binary maps of presence and absence (e.g., Scott and Crossman 1973) to continuous probability distributions mapping the likelihood of species presence. Common techniques include generalized linear models, generalized additive models, genetic algorithms such as GARP (Stockwell and Peters 1999), and maximum-entropy techniques such as Maxent (Phillips et al. 2006). This study used Maxent, because it generally outperformed the commonly used GARP, Domain, and Bioclim models in a series of tests modeling the distributions of 18 species (Hernandez et al. 2006). While designed to map species distributions, these maps have also been applied to mapping the habitat requirements of species during life-history phases such as reproduction (Yost et al. 2008).

The purpose of this study is to create two GIS-based models that can be used by managers to identify the location of muskellunge spawning sites in any northern Wisconsin lake. This was accomplished using datasets of spawning sites from 31 northern Wisconsin lakes, spatial habitat data, and optional habitat surveys. The Maxent modeling program used this information to train and test the Muskellunge Spawning Habitat Models.

Methods

This study identified muskellunge spawning habitat in 25 lakes located in the Northern Highlands lake district of Wisconsin (Table 1, Figure 1). These inland lakes were created by glaciers approximately 10,000 ybp and are surrounded by a mix of glacial geological features including outwash plains and moraines (Martin 1965). The lakes studied represent a wide diversity of lake types, from oligotrophic to eutrophic. Lake size ranged from small lakes of 41 ha to larger 1450-ha lakes. To avoid potential effects of stocking on site selection, lakes were chosen that had not been stocked for at least ten years prior to being surveyed. All lakes were classified as either category one or two muskellunge lakes by the Wisconsin Department of Natural Resources (WDNR), which means they receive little or no stocking and support naturally reproducing populations (WDNR 1996).

Spawning Site Determination

Spawning site locations were determined using nighttime spotlighting surveys during the spawning period (after Zorn et al. 1998). Surveys were completed using 1,000,000 candlepower or greater spotlights to visually locate potentially spawning muskellunge. The surveys were performed from boats operated at speeds of approximately 3 km/h and powered by electric trolling motors. This method allowed surveyors to identify muskellunge up to 15 m from the boat. Areas wider than 30 m were covered by multiple passes, so that the entire shallow water area (<1.5 m) of each lake was surveyed. The size of spawning muskellunge, use of spotlights, and depth of the water surveyed contributed to a very low likelihood (<5%)

of failing to see or correctly identify muskellunge within 15 m. Stained lakes and those surveyed under poor weather conditions had lower visibility, but surveys were only conducted if researchers were able to confidently identify fish at depths of 1.5 m. Surveys were completed during the spawning period, 23 April to 9 May 2007 and 5 to 23 May 2008. Surveys were targeted for each lake after the daytime surface temperature reached 11.5°C, but actual surface temperatures of verified sites during the surveys ranged from 6.6 to 15°C. The location of each adult (visually estimated > 0.6 m) muskellunge sighted was recorded using a Trimble GeoXM GPS unit and identified as a potential spawning site. A spawning site was defined as the location of a muskellunge or pair of muskellunge identified during spotlighting surveys and verified through egg surveys. A grouping of multiple spawning sites is sometimes referred to as a spawning ground (e.g., Crossman 1990), but the unit of study for this research was a spawning site. While the spotlighting method excluded fish spawning in deeper water (e.g., Pierce et al. 2007), these deepwater sites are likely to be an exception. The average depth of muskellunge spawning sites in the Pierce et al. (2007) study was 1.1 m, and Farrell (2001) found that muskellunge avoided spawning at depths greater than 1.5 m.

Each potential spawning site was searched for eggs within eleven days of the spotlighting survey (after Zorn et al. 1998). Sediments were searched with a D-frame net until a muskellunge egg was found or an effort of 1.5 person-hours of search had been achieved. Due to a high number of sites, sites at North Twin Lake were limited to 1 person-hour searches. Oehmcke et al. (1958) described muskellunge behavior as broadcast spawning over "several hundred yards," and I observed muskellunge spawning over distances of more than 140 m in this study. Therefore, muskellunge eggs found within 200 m of a site were considered a verification of spawning for that site. A subsample of eggs were hatched and grown to 30 d to verify identification. All spawning sites included in the habitat selection analysis and model creation were verified by positively identified muskellunge eggs.

GIS Data for Habitat Analysis

A series of habitat maps were used to describe available habitat and spawning sites. Each verified spawning site was represented by a point, and the habitat conditions underlying that point were attributed to it using Hawth's Analysis Tools (Beyer 2004). All maps were created using ArcGIS (Environmental Systems Research Institute, Redlands, CA).

Habitat surveys for substrate, emergent, and submersed aquatic vegetation (SAV) were conducted at each lake between 28 May and 16 June 2008. Surveys required two to six hours per lake, depending upon shoreline length and water clarity. For each survey, a boat moving at approximately 6 km/h followed a visually estimated 1.5-m depth contour along the entire shoreline. Visual observations classified the major substrate and presence or absence of vegetation along a cross-sectional line between the 1.5-m contour and the shore. A polyline feature representing the habitat classification of the cross-section was created in ArcPad on Trimble GeoXM GPS units (Sunnyvale, CA). When the habitat classification along the cross section changed, the polyline was ended and a new polyline was begun. Inorganic substrates were visually categorized using a modified Wentworth Scale as silt, sand, pebble/gravel, cobble, and boulder (Allan 1995, after Cummins, 1962). Muck (e.g., gyttja) was included in the silt category. The final substrate category consisted of coarse benthic organic matter (CBOM) such as leaves and woody debris. Since previous research has emphasized the structural aspect of aquatic vegetation for egg survival (Dombeck et al. 1984), vegetation was categorized by structural groups (Tables 2 and 3). The minimum

criterion for recording the presence of SAV was 50% coverage along the cross sectional line. Emergent vegetation was recorded similarly, with the difference that the minimum criterion was simply presence. Due to equipment failure, approximately half of the vegetation survey data were lost. Lakes with habitat survey information used in this study are noted in Table 1. To explicitly allow for interactions between habitat variables within the model, the "combined habitat" variable was created. This variable represented each possible combination of substrate, SAV, and emergent vegetation and was used only for the models.

Bathymetric maps were obtained from the WDNR as images of scanned hard-copy bathymetric maps (WDNR 1939-1978). These images were georeferenced, and the shoreline and 5-ft (~1.5-m) contours were digitized manually. A triangulated irregular network bathymetric model was created, from which 5-m raster data sets were created for slope (percent slope), aspect (degrees), and plan curvature. Plan curvature (referred to as curvature from this point) is calculated as the second derivative of the near-shore bathymetric surface, perpendicular to the direction of the slope (Figure 2). A high curvature value represents a concave feature such as a bay, while a low value represents a convex feature such as a point. Aspect represents the compass direction which the bathymetric slope faces, so a value 180° represents a slope on the southern shore facing north (Figure 2). Nearby shallow area, defined as the total area between the 0- and 1.5-m contours within 100 m of a given cell, was calculated as a metric of available potential spawning habitat within 100 m (Figure 2). Nearby shallow area was correlated to slope, but included in the analysis because it provides additional information that slope does not contain and has been referenced as a characteristic of muskellunge spawning habitat (Farrell et al. 2007). To account for differences between dated bathymetric source data and current geospatial data, the raster cells between the 0- and 1.5-m contours were converted to points, then to Thiessen polygons, and finally back to a raster format. While some cells represented the nearest measured value for each variable, most (86%) represented the actual value at that location.

Land cover data derived from Landsat images and interpreted in the WISCLAND dataset were obtained from the WDNR (1998). The category "open water" was removed from this 30-m raster dataset, and the WISCLAND land cover classes were grouped as agriculture, coniferous forest, deciduous forest, developed /barren, grassland, forested wetland, shrubland, or wetland. Thiessen polygons were used to extrapolate values over the water, so each cell represented the land cover at the nearest point on the shoreline.

The potential for groundwater flow was determined using a GIS-based application of Darcy's Law in a 30m raster dataset (Baker et al. 2003). The area of open water defined by the most recent GIS maps (WDNR 2007) was removed from the analysis. A Thiessen polygon approach similar to the land cover data preparation was used, so that each cell represented the groundwater potential at the nearest point on the shoreline. The habitat preference analysis used groundwater flow categories of positive (groundwater discharge), negative (groundwater recharge) and zero for χ^2 tests.

There are anecdotal references to muskellunge preferring spawning habitats located near streams (Dombeck et al. 1984; Zorn et al. 1998). In order to test muskellunge's preference for sites near streams, the distance to the mouth of the nearest inflowing or outflowing stream was calculated. A cost-distance map was created, which ensured that the distance represented the shortest possible distance traveled through the water in meters. All distances larger than 3 km were coded as 3 km, since any effects from stream flow would be

negligible above this distance. The habitat preference analysis used stream distances categorized by break points of 10, 50, and 100 m.

Spawning Habitat Preference

The habitat preference of spawning muskellunge was analyzed using Ivlev's index of electivity (1961), where electivity, E = (r-p)/(r+p). The proportion of verified spawning sites (r) was compared to background proportions for each habitat category (p). Habitat preference is a relative judgment by individuals of the species in question. Habitat electivity is inferred from the difference between available habitat and the habitat utilized, which requires that the investigator defines "available habitat." When using Ivlev's index of electivity, the magnitude and direction of the preference for a habitat type is dependent upon the definition of available habitat (Johnson 1980). Most descriptions identify muskellunge spawning habitat as water less than 1.5 m deep (e.g., Scott and Crossman 1973) and the spotlighting methodology limits observations of fish at greater depths, so this study did not address habitat deeper than 1.5 m. Since muskellunge have been shown to travel 5-21 km to spawn (Strand 1986), the entire lake in which each fish lived was assumed to be available. Background frequencies representing the available habitat to each fish were determined by randomly assigning one thousand points per spawning site to each lake (Winter and Ross 1982). These points were distributed between the 0-m and 1.5-m depth contours, representing the habitat available for selection by any fish in that lake. Data from all background points were then combined across lakes to form a composite background representing the combined habitat available to spawning muskellunge in this study.

A χ^2 test was used to determine significant differences between the proportion of spawning sites and available habitat for each variable (SPSS Inc., Chicago, IL). The null hypothesis was that the proportion of spawning sites in each category was equal to the background proportion ($\alpha = 0.05$). An assumption of this test is that the expected value for each category is greater than or equal to five. Therefore, categories with less than five expected spawning sites were removed from the χ^2 analysis, or in the case of continuous variables, combined with numerically adjacent categories. A post-hoc Z-test for proportions was used to test each individual category of habitat against the background frequency (Zar 1999), with the null hypothesis that the two frequencies were equivalent ($\alpha = 0.05$). Those habitat categories with significantly different frequencies from background rates were concluded to represent selection for or against that habitat characteristic.

Muskellunge Spawning Habitat Models

Muskellunge spawning sites were predicted using the machine learning program Maxent (Phillips et al. 2006, Phillips et al. 2008). Maxent has been used to predict species distributions of plants, amphibians, birds, and reptiles based on geospatial habitat data (Phillips and Dudik 2008). The program creates an output map with each cell representing the probability of a species' presence between 0 and 1. In this case, muskellunge spawning habitat was treated conceptually as a species, and the distribution of this "species" was modeled. The Maxent program iteratively fits a species' predicted distribution to each set of environmental variables. It does this using a maximum likelihood algorithm that maximizes entropy to optimize fit. Most techniques require both presence and absence data. However, Maxent models are designed to predict species distributions in cases such as this study, where only presence data exist. Maxent utilizes data from background points that represent

pseudo-absence locations to train the model. This study used the available habitat points described above as background points.

All model parameters were set to the Maxent program's default settings, except the regularization multiplier β was set to 1.25. This smoothed the additive contribution of each continuous variable across its range, reducing the potential for overfitting in the model. Trial models with β set to 1 (default) showed large changes in the assigned probability of spawning across very small ranges of certain variables. These were more likely due to the modeling technique than biological relationships, and were reduced by increasing β. Trial models also showed decreased performance with test data when β was too high, so 1.25 was selected to minimize overfitting and maximize test performance. Categorical data included habitat combination and land cover variables. Continuous data included groundwater potential, slope, aspect, curvature, and amount of nearby shallow area. To be included in the model, a variable needed to have a significant χ^2 value from the habitat preference analysis. The output data is in the form of a 5-m raster grid file, with the value of each cell converted from the raw Maxent output to the "predicted probability of presence" using a logistic function (Phillips and Dudík 2008).

Two Maxent models were created using different sets of habitat variables. Muskellunge Spawning Habitat Model 1 (MSHM1) was created using aspect, curvature, groundwater flow potential, land cover, nearby shallow area, and slope. These variables can be obtained without conducting the habitat surveys described above. Therefore, data collected by Johnson (2000) using the same methods and meeting the standards of this study were included in MSHM1. Johnson's sites were scanned from hardcopy maps, georeferenced, and attributed habitat data using the methods described above. MSHM1 was created from 347 spawning sites in 31 lakes. Muskellunge Spawning Habitat Model 2 (MSHM2) included all of the MSHM1 variables plus the combined habitat variable obtained through habitat surveys. This model was created using a smaller (144) spawning site dataset, which consisted of sites in the 13 lakes with both aquatic vegetation and substrate surveys.

To understand how a specific variable impacted the predicted likelihood of muskellunge spawning, the logistic probability of spawning was plotted across the range of that variable's values for a model which held all other variables at their average value. Outputs near one indicate a high probability of spawning, while outputs near zero indicate a low probability. This analysis is heuristic; the actual models incorporate interactions between variables for increased accuracy. However, it is the best illustration of how each variable contributed to the multivariate analysis, because the variables were held at the average value for spawning sites instead of simply the average values for the available habitat.

The extent to which each variable contributed to the model was estimated using a jackknife test of variable importance on training data (Yost 2008). This technique measures the training gain from a model based solely on one variable and a model based on all but that variable. Training gain, which represents the increase in likelihood between the average spawning site and a uniform probability distribution, is calculated for each model. Gain is the average log probability of presence samples minus a constant that sets the uniform distribution to zero. For example, a gain of 2 would indicate that the average likelihood of spawning at the training spawning sites is 7.4 (e^2) times higher than a random background pixel.

To test the models' performance, 25% of spawning sites (96 for MSHM1 and 36 for MSHM2) were randomly withheld from model training. These were used for two tests to

assess and compare the models. The first tested the null hypothesis that the model performed no better than a random model at predicting the locations of sites selected by muskellunge for spawning, using a binomial test ($\alpha = 0.05$). A significant binomial test would provide evidence that the area deemed "presence" by the model encompassed significantly more spawning sites than a randomly selected area of equal size. This test requires presence and absence categories, while the model output is a continuous probability of presence from 0 to 1. Therefore, data were reclassified as presence or absence based on the output value. The cutoff, or threshold, for this reclassification was the value which equalized the model's specificity and sensitivity for test data. This threshold represents the output value at which rates of false positive (the model classifies a location as spawning habitat that is not actually used) and rates of false negative (muskellunge spawn at a site not indicated as spawning habitat by the model) instances are equal.

A second, threshold-independent analysis of the models was also completed by an area under the curve (AUC) analysis for both training and test data. The AUC represents "the probability that a random positive instance and a random negative instance are correctly ordered by the classifier" (Phillips et al. 2006). It is calculated by estimating the area below the curve of training and test data in a plot of sensitivity (1 - omission rate) against specificity (fractional area predicted). Therefore, a point (x, y) on this plot signifies a model that classified x percent of the area as presence, which included y percent of the spawning sites. AUC values were analyzed for both training data and test data. An AUC value of 0.5 represents the predictive power of a random model, and a value of 1 is ideal. However in practice, models created from solely presence data have a maximum possible AUC of less than 1 (Wiley et al. 2003).

Results

Spotlighting surveys identified 424 potential spawning sites. Due to the misidentification of a catostomid egg as a muskellunge egg at North Twin Lake, all sites at this lake and the adjoining South Twin Lake searched prior to the discovery of this error (58) were removed from the analyses. The remaining sites were verified according to protocol with the positive identification of a muskellunge egg. Of the 366 remaining potential sites, 247 (67%) were verified by egg surveys. These spawning sites were used for spawning habitat preference analysis and model creation.

Spawning Habitat Preference

Habitat survey data were compiled for all ten variables. Nine of those variables showed a significant difference between spawning site and available habitat frequencies (Table 4). The only variable not showing a significant difference was stream distance. Only 15 sites (6%) were located within 100 m of a stream, which was the same as the available habitat (6%). This pattern was consistent for sites not adjacent to wetlands, of which 6% of both spawning sites and available habitat were also located within 100 m of a stream. This suggested that muskellunge did not select for or against spawning locations near streams regardless of the presence of adjacent wetlands. The variable was dropped from further analysis and was not included in the models.

Muskellunge showed a strong preference for certain types of substrates (Table 5). The percentage of spawning sites in areas with sand (55%) was lower than the proportion of available sand habitat (66%). Therefore, although the majority of sites were over sandy

substrates, the index of electivity provides evidence that muskellunge selected against sand when choosing a spawning site. There was also negative selection for gravel substrates. Muskellunge selected positively for CBOM and cobble substrates.

Sites where SAV was absent represented 69% of spawning sites, and 79% of the available habitat (Table 5). While the majority of spawning sites were absent of SAV, the index of electivity suggests that muskellunge do not prefer to spawn over such sites. Sites with mat-forming vegetation and short grasses were preferred spawning areas. There were very low frequencies for both dissected- and simple-leafed vegetation found during the habitat surveys, which was likely due to the short time period between ice-out and the surveys. Because these two categories did not meet the requirements of the χ^2 test, no Z-test was applied.

Emergent aquatic vegetation showed a similar pattern to SAV, with muskellunge selecting against sites absent of emergent vegetation (Table 5). While 52% of sites had no emergent vegetation, 70% of available habitat was absent of emergent vegetation. There was a strong selection for sites with rushes, and similar selection for sedges. Wild rice and cattails were removed from the statistical analysis, because there were fewer than 5 sites expected for these categories in the χ^2 analysis.

Muskellunge selected against spawning sites with very shallow slopes (0–3%, Table 5). However, there was positive selection for sites in areas with moderate slopes (9.1–12.2%). Due to low expected values and the requirements of the χ^2 test, multiple categories were combined to represent steep slopes (12.2–32.9%). There was positive selection for this category as well.

The index of electivity indicated that muskellunge preferred to spawn at sites with little nearby shallow area (Table 5). Spawning sites located in areas with less than 0.9 ha of shallow area within 100 m were positively selected and represented a high percentage (44%) of spawning sites. This metric showed a strong inverse covariance with slope, since the denominator used to calculate slope is also a factor in calculating area.

Bathymetric slopes facing east to north-east (46–90°) and south to south-west (181–225°), were slightly preferred for spawning (Table 5). Sites on slopes facing north to north-west (316–360°) were selected against.

Sites with convex bathymetries represent point features, and muskellunge showed a preference for spawning in these areas (Table 5). Additionally, spawning sites with concave curvature, which represent more protected bays, were also preferred. In contrast, there was a negative preference for sites along straight shorelines with no curvature at all.

There was a positive selection for sites adjacent to wetlands (Table 5). Sites near forested wetlands were selected against. Areas adjacent to deciduous forests were also selected against, despite the fact that they represented 53% of spawning sites. It is important to note the relatively pristine land cover conditions surrounding these lakes, as areas adjacent to barren and developed land comprised less than 1% of the available habitat.

There was a clear relationship between spawning site selection and potential groundwater flow. Muskellunge selected against sites with low potential flow velocities (0.01-2.5 m/d, Table 5). However, those sites with higher negative (groundwater recharge) or positive (groundwater discharge) potential velocities showed positive selection (<-2.5 and >5 m/d). This bimodal selection for high rates both into and out of the lake suggests that the rate may be more important than the direction.

Muskellunge Spawning Habitat Models

The relative contribution of each variable to the overall probability of spawning was consistent between the two models with a few exceptions (Figure 3). This was shown by the training gain for each variable, which is an indicator of the predictive power that the variable contributes to the model. The combined substrate, emergent, and submersed vegetation (habitat combination) category in MSHM2 showed the highest overall training gain. This variable more than doubled the gain of the next most important variable, groundwater potential. The related slope and nearby shallow area variables both contributed moderately to the model. Aspect and curvature also contributed but were not strong factors. Training gain loss was greater in MSHM1 than MSHM2 when the full models were compared to their counterpart without land cover. Because the combined habitat variable was the major difference between MSHM1 and MSHM2, this confirms a partial redundancy between the land cover and combined habitat variables.

There was a general agreement between the probabilities of spawning assigned by MSHM1 and MSHM2 and the results of the habitat preference analysis (Figures 4 and 5). Sites with a high potential for groundwater discharge (positive) or recharge (negative) were more likely to provide spawning habitat. Both models predicted spawning probability minima at areas of very low magnitude potential groundwater flow. Areas adjacent to wetlands were more likely to be spawning habitat than all categories except grasslands, which had a very low sample size (Table 5). MSHM2 predicted higher probabilities over intermediate slopes, while MSHM1 predicted intermediate to steep slopes as more likely spawning sites. In contrast to the univariate approach of the habitat analysis, the model assigned higher likelihoods for spawning to locations with large amounts of nearby shallow area (holding slope at its average spawning site value). Aspect did not show any strong difference in likelihood across its range, but was slightly higher between values of 3 and 5 radians (shorelines facing between north and east southeast). Sites with sand, cobble, and CBOM showed high probabilities, while silt was assigned a low probability. Rushes and sedges resulted in greater spawning habitat probabilities. SAV categories with mat-forming or absent of vegetation generally were also likely to be spawning habitat.

The binomial tests strongly support both models. The MSHM1 breakpoints for classifying locations as presence or absence were 41.6% (training data) and 39.7% (test data). These values represent a cutoff designating 26.5% and 29.2% of the available habitat as spawning area, respectively. At these thresholds, MSHM1 performs significantly better than a model generating random predictions for both training and test data ($p < 0.0001$). The breakpoints for MSHM2 were 29.5% for training and 36.7% for test data, designating 36.1% and 27.8% of the available habitat as spawning habitat, respectively. The probability values give a strong indication that this model outperforms a random model ($p < 0.0001$ and $p = 0.0003$, respectively).

An analysis of the AUC values indicates good performance for both models (Figure 6). A training AUC value of 0.802 for MSHM1 indicates that approximately 80% of the time this model will correctly identify a random presence and absence point. The test AUC value (0.786) is very similar to the training AUC, indicating a robust model. Following the *x*-axis in Figure 6 to 0.1, one can investigate the performance of a model that designates 10% of the available habitat as presence. At this level, the sensitivity (along the y-axis) indicates that MSHM1 would include 50% and 45% of the training and test spawning sites in this area,

respectively. Increasing the fractional predicted area to 25%, MSHM1 included 71% and 64% of training and test spawning sites.

By comparison, MSHM2 showed training and test AUC scores of 0.842 and 0.702. This training data AUC is higher than the MSHM1 value, while the test AUC is lower. MSHM2 classified 60% and 50% of the observed training and test sites correctly with only 10% of the area. These values are higher than the MSHM1 model's, indicating a superior ability of MSHM2 to identify the top 10% of shallow areas for muskellunge spawning. If MSHM2 designates 25% of the shallow area as spawning habitat, it encompasses 71% and 58% of the training and test spawning sites. This performance is worse than MSHM1 in terms of test site omission rates. The MSHM2 test curve shows poorer performance than MSHM1 between fractional predicted areas of 12% and 84%, indicating that the latter may classify marginal habitats more accurately. However the training curve shows the opposite conclusion over this range, suggesting that MSHM2 outperforms MSHM1 over most of this range. Overall, the AUC comparison shows that MSHM2 is more efficient at predicting the best 10% of shallow area, and is inconclusive in determining which model is more efficient across the middle ranges which provide marginal spawning habitat.

Model output maps were created for all lakes in the study. Figures 7 and 8 show examples of these outputs for MSHM1 and MSHM2, respectively, at Birch Lake. Since both models rate habitat pooled across the sets of lakes from which they were created, the distribution of probabilities across each lake is variable. This is evidence that some lakes have better spawning habitat than others. For comparison between models, Figures 7 and 8 identify spawning habitat as those areas rated in the top 50% for Birch Lake. A visual test of the identified spawning habitats corroborates the model performance metrics. The areas that MSHM1 predicted muskellunge use for spawning included 14 of 18 (78%) observed spawning sites in this lake. Similarly, areas designated as spawning habitat by MSHM2 included 13 out of 18 (72%) spawning sites.

By inspection, both models appear to predict the same general areas as high probability spawning locations in each study lake. The predictions of MSHM1 and MSHM2 at Birch Lake identify spawning areas in the southeastern bay, the southwestern bay, and along the northeastern shoreline (Figures 7 and 8). These areas share characteristics that both models identify as spawning habitat, such as moderate slopes. However, MHSM2 accurately predicts muskellunge spawning in the large and shallow sloping area along the eastern shoreline while MSHM1 predicts an absence of spawning. This area is a large bed of rushes and mat-forming vegetation with a sand substrate, characteristics to which MSHM2 assigned very high probabilities of spawning (Figure 5). The inclusion of this variable enabled MSHM2 to correctly identify these spawning sites where MSHM1 did not. While both models score habitats differently, inspection shows strong performance and a general agreement between the two.

DISCUSSION

Spawning Habitat Preference

The habitat surveys in this study supported some of the paradigms in muskellunge spawning habitat research. Muskellunge preferred spawning in areas with vegetation present. Sedges and rushes were preferred emergent vegetation. Mat-forming vegetation and short grasses were preferred SAV categories. Muskellunge selected for CBOM and cobble while selecting against sand, and preferred locations with the potential for groundwater flow. These

findings support previous literature using a large sample of spawning sites identified from surveys of the entire shoreline to determine spawning habitat preference. The data also showed unexpected results with regard to the low proportions of sites with simple- and dissected-leaf SAV, the lack of a preference for spawning sites near streams, and a negative selection for gravel substrates.

Muskellunge preferred to spawn in areas with CBOM substrates, as expected. Previous work has reported that muskellunge egg mortality was lowest over organic substrates (Dombeck et al. 1984), that successful natural reproduction was correlated to the density of shoreline deadfall trees (Rust et al. 2002), and that stocked fall fingerlings selected for woody habitat (Hanson and Margenau 1992). The low proportion of CBOM substrates for both spawning sites and background levels in this study is likely a result of the classification criteria for the "majority constituent," since sites with low levels of CBOM coverage were present but not classified as CBOM. The data from this study suggest that while areas with a majority coverage of CBOM are rare, they are likely locations for spawning.

The finding that muskellunge selected for cobble but not gravel was unexpected. CBOM and cobble likely reduce the most significant sources of mortality found by Dombeck et al. (1984): hypoxic conditions and *Saprolegnia* spp. fungus. Dombeck et al. (1984) did not test cobble, but found that muskellunge eggs survived longer over wood and gravel than most natural substrates. Wood, sand, and gravel had the highest DO concentrations in the same study. This supports the interpretation of our data that the preference for spawning over CBOM and cobble may exist in part to provide sufficient DO for the eggs. Dombeck et al. (1984) and Zorn et al. (1998) found lower DO levels at spawning sites in lakes with poor natural reproduction. Zorn et al. (1998) found that self sustaining lakes had higher percentages of wood substrate and that survival to hatching *in situ* appeared slightly higher in wood treatments. The physical isolation in the interstitial spaces of cobble and elevated on particles of organic material should decrease mortality by limiting the spread of aquatic fungi (Dombeck et al. 1984). Therefore, the finding that muskellunge selected against gravel was unexpected. The data show that sand was widely available, but was also selected against for spawning habitat. This is consistent with the previous interpretation, since sandy substrates lack interstices large enough to retain and separate muskellunge eggs. In contrast to this study, Hanson and Margenau (1992) found that young-of-year (YOY) muskellunge stocked in the fall selected sandy substrates. This apparent inconsistency can be explained by the major sources of mortality facing each life stage. While eggs are subject to fungal infections, YOY mortality is based more heavily on predation and the ability to forage (Dombeck et al. 1984; Margenau 1992). Thus, avoiding sandy substrates for egg deposition may still be an important survival strategy despite the apparent YOY preference for sand.

The emergent vegetation structural categories of sedges and rushes were preferred, which was also expected. This confirms the preference for sedges and rushes found in other studies for northern Wisconsin lakes. Margenau and Hanson (1992) found that *Scirpus* spp. were utilized by age 5-month fry. Additionally, Craig and Black's (1986) study in Georgian Bay of Lake Huron found that sedges composed approximately 52% of the emergent plant community in nursery habitats, and that rushes were present at 72% of nursery areas. However, Johnson (2000) was unable to find evidence for a relationship between percent cover of *Scirpus* spp. and recruitment at the whole-lake scale. While the structural distinction between categories of sedges and rushes in this study differs from the taxonomic distinction

between the families of sedges Cyperaceae and rushes Juncaceae, this difference is mitigated by that fact that both structural categories showed positive selection.

Muskellunge also selected against sites where emergent or submersed vegetation were absent, as expected. However 52% and 69% of spawning sites lacked emergent and submersed vegetation, respectively. Because an even greater proportion of background sites were absent of vegetation (70% and 79%, respectively), this can be interpreted as evidence that sites without vegetation are the most common habitat, and sometimes represent the best available spawning habitat due to other characteristics. The low frequencies of vegetation in both spawning and available habitat may be due to a combination of factors. Vegetation surveys were conducted between 1 and 1.5 months after ice-out, so the lack of aquatic vegetation is partially explained by the short growing period. This lack of vegetation was likely even more pronounced when muskellunge spawned approximately one month earlier. Farrell (2001) found that muskellunge spawned over new growth of SAV in shallow water areas, which were often scoured by wind and ice. Additionally, aggressive intraspecific competitive interactions were observed during spotlighting and egg surveys. One example occurred at a location with prime spawning habitat in Horsehead Lake. An apparent alpha male repeatedly chased a subordinate male (sex inferred by size and subsequent behavior) from the ~0.5-ha spawning ground. After a number of chases, a larger female moved into the area, and both males swam alongside attempting to spawn. These observations, and the fact that the majority of fish exhibited scars from apparent bouts, support the hypothesis that strong competition exists for spawning habitat. This competition would further explain the usage of marginal habitats (i.e., absent vegetation) by some individuals in lakes with high adult densities and minimal vegetation present.

The finding that muskellunge were only selecting for short grasses and mat-forming SAV was unexpected. While previous research has shown a preference for *Chara* spp. and *Vallisneria americana*, it has also shown preferences for spawning in areas with *Myriophyllum* spp., *Potamogeton* spp., *Najas flexilis,* and *Elodea canadensis* among others (Dombeck et al. 1984; Craig and Black 1986; Werner et al. 1996; Murry and Farrell 2007; Pierce et al. 2007). The mat-forming and short grass vegetation category represented 27% of spawning sites. It included both *Chara* spp. and *Vallisneria americana,* which was included due to its short stature at this time of year, and verified that muskellunge in Wisconsin's inland lakes shared a preference for these species of SAV. The species included in the dissected- and simple-leaf categories represented less than 2% of the weighted available habitat. Their lack of a majority presence in both spawning and available habitats suggests muskellunge may be simply utilizing the best available habitat. It should also be noted that plants from both categories were often found during eggs surveys of spawning sites even if they did not represent the major form of vegetation during habitat surveys.

Muskellunge spawning sites showed strong selection for slope, aspect, nearby shallow area, and curvature. There is little evidence for a direct relationship between egg mortality and these bathymetric indices. More likely, these factors create conditions which are favorable for muskellunge fry and eggs. For example, Farmer and Chow-Fraser's (2004) conceptual model of muskellunge spawning habitat relates primary drivers of successful spawning (i.e., water temperature, DO, and egg separation) to secondary variables (e.g., depth, current, and particle size). The bathymetric variables in the present study are analogous to Farmer and Chow-Fraser's secondary variables, in that they often cause or are correlated with the primary variables.

Moderate to high slopes (>9.1%) and low levels of nearby shallow area (< 0.45 ha) were selected for spawning habitat. While neither of these characteristics directly impacts egg survival, the resulting physical conditions that they create may be important. Steep slopes result in a closer proximity to deep water, which facilitates mixing with the limnetic zone and could stabilize temperatures. Conversely a large, flat area may experience greater temperature fluctuations. To evaluate such conditions, StowAway WT56–8K temperature recorders were placed in eight known spawning locations at Horsehead Lake in 2008. A site facing southwest and located in a shallow bay experienced a 4.99°C temperature drop on 23 May 2008; in contrast, a site with moderate slope and facing northeast showed only a 2.7°C decline. Mitigation of such temperature fluctuations is likely to be important to muskellunge egg and fry survival (Bean 1908; Oehmcke et al. 1958; Zorn et al. 1998), and may be correlated to slope and nearby shallow area. It is interesting to note that while the univariate habitat preference analysis indicated muskellunge selected sites with little nearby shallow area, the multivariate habitat models, which included the correlated slope variable, indicated larger amounts of nearby shallow area were preferred. This suggests that the primary preference is for sites with moderate to steep slopes, but also that a secondary preference may exist for increased shallow water habitat that provides food and shelter for fry. This conclusion is supported by the observation of Farrell et al. (2007) that spawning sites in the St. Lawrence River were "clustered in areas that have a relatively high abundance of shallow littoral habitats."

Muskellunge preferred to spawn near both concave (bays) and convex (points) shorelines. While most authors describe muskellunge as utilizing bays for spawning (Oehmcke et al. 1958; Becker 1983; Crossman 1990), points are not mentioned. One possible explanation for this finding is that points and bays often occur in close proximity, since the geological processes that form them are similar. Crossman's (1990) definition of a spawning ground as very large area (189 ha) and observations of muskellunge broadcast spawning over distances greater than 140 m from this study make it likely that areas with complex shorelines will include spawning sites both on points and in bays. One possible biological explanation for this would be protection from severe wave action in areas with a higher shoreline development factor.

Shoreline aspect showed a weak pattern of preference by spawning muskellunge, and was relatively unimportant in contributing to MSHM1 and MSHM2. The variable was included because it was hypothesized that the northeastern shores would warm earlier in the spring due to the prevailing wind and angle of the sun. Dombeck (1979) showed that muskellunge preferred sites that were approximately 1.5°C warmer than adjacent waters, and Dombeck et al. (1984) noted that many of their sites were in areas of the lake that thawed earlier. The data from both the univariate habitat selection preferences and the multivariate models showed a weak preference for shores facing west-south-west through east-north-east. Data compiled over at least a four year period between 1997 and 2001 indicates that the prevailing winds in this region are from the southwest (Global Energy Concepts 2002). Since the direction of the prevailing wind shows little meaningful pattern, and the preferred categories are nearly opposite on another, a biological interpretation of the preference for sites facing certain directions based on the effects of wind would be speculative.

Spawning muskellunge selected for sites with extreme potential groundwater velocities. This finding is corroborated by previous studies which have suggested that muskellunge spawning sites were associated with groundwater discharge (Zorn et al. 1998)

and experienced earlier ice-out due to groundwater or streams (Dombeck et al. 1984). It also presents evidence that areas of groundwater recharge are preferred spawning habitat. There are three main arguments for the function of groundwater. First, groundwater flowing into nearby Sparkling Lake was shown to increase DO levels during the summer (Hagerthey and Kerfoot 1998), and likely has the same effect in the spring. Areas of recharge are likely to have higher DO levels due to the flow of water down through the sediments. Second, groundwater potentially stabilizes temperatures. The temperature drop of 4.3°C over two days on 10 May 2008 at Horsehead Lake and the increase of 5.4°C on 15 May 2008 are evidence of the thermal variation that eggs and fry experience. The mitigation of such drastic thermal variation may be important for successful recruitment (Bean 1908; Oehmcke et al. 1958; Zorn et al. 1998). Third, groundwater influences the emergent and submersed vegetative communities. Areas of groundwater discharge are often marked by patches lacking vegetation or composed of *Chara* spp. amidst other vegetation (Rosenberry et al. 2000). These areas with springs could contribute to an intermediate vegetation density which could confer foraging and protective benefits to fry without the significant reduction in DO associated with dense vegetation (Werner et al. 1996).

While the muskellunge showed a preference for spawning near sites with water flow generated by groundwater discharge, there was no preference detected for sites near streams. This was unexpected, since a number of studies have described spawning sites in close proximity to streams (Dombeck 1979; Dombeck et al. 1984; Zorn et al. 1998). While some spawning sites were located in or near streams (6% within 100 m), these occurred at nearly the same ratio as the available habitat (6%). Despite previous observations, these data support the conclusion that muskellunge in northern Wisconsin did not select for spawning sites near streams. Dombeck et al. (1984) and Zorn et al. (1998) described a preference for spawning areas with *either* groundwater discharge *or* a close proximity to streams. This hypothesis was investigated by removing sites adjacent to wetlands from the analysis and testing the remaining sites. The pattern did not change with this modification, indicating that there was no strong preference for sites near streams, regardless of the presence of adjacent wetlands.

The preference of muskellunge for wetland land cover was expected. Multiple descriptions of spawning habitat include an association with aquatic vegetation that characterizes wetlands (Scott and Crossman 1973, Craig and Black 1986). The 30-m resolution of Landsat images and the resulting WISCLAND data set mean that only large patches of wetlands could be identified. While these wetlands represented approximately 4% of available habitat, 11% of spawning sites occurred near them. It is interesting to note the discrepancy between MSHM1 and MSHM2 in the probabilities of spawning assigned to wetlands. Univariate habitat selection analysis and multivariate MSHM1 both designated wetlands as likely for spawning. However, the multivariate MSHM2 predicted that wetlands are unlikely spawning habitat, while assigning high values to individual vegetation categories. The increased resolution of the habitat surveys may account for their high relative importance in MSHM2. Areas with wetland land cover were assigned a high likelihood through the combined habitat variable, which superseded the coarser WISCLAND dataset in MSHM2.

This study found that shoreline land cover is an important factor for muskellunge spawning, a conclusion supported by Rust et al. (2002). Rust et al. investigated the relationship between shoreline land use (developed, undeveloped, and forested) and natural

reproduction at the whole-lake scale, finding that developed shorelines had a negative impact due to their correlation with decreased woody debris. Since the lakes in this study showed little to no development (<1% of the available habitat), these results and categories are not directly comparable.

Muskellunge Spawning Habitat Models

Muskellunge Spawning Habitat Models 1 and 2 showed relatively strong performance. Both models performed significantly better than random, as determined by the binomial tests.

The strength of each model is evidenced by the respective AUC values. The MSHM1 training (0.802) and test (0.786) AUC values indicate the model has the ability to correctly distinguish between two randomly selected sites where spawning does and does not occur approximately 80% of the time. AUC values between 0.7 and 0.9 indicate a "reasonable discrimination ability" (Pearce and Ferrier 2000). MSHM2 also showed good AUC values. The large difference between training (0.842) and test (0.702) scores potentially indicates overfitting, resulting from the high number of potential combinations of habitat survey variables. Trial models created without combined habitat variables showed a similar pattern, so the decision to combine habitat variables had little impact on this effect. Another possible explanation for this difference is the low number of samples from which the MSHM2 test AUC was calculated (36), resulting in higher variation between model runs.

Investigation of the AUC plots shows the relative utility of each model. By increasing the probability of spawning which the manager designates as an acceptable cutoff between presence and absence of spawning habitat, a decreasing proportion of the potential area will be classified as spawning habitat. MSHM2 relies heavily on habitat survey variables, and is therefore grounded in more direct relationships with influences upon egg and fry survival (Dombeck et al. 1984; Farmer and Chow Fraser 2004). The training data and test data show that MSHM2 was more efficient at classifying the top 10% of available habitat. Efficiency in classifying relatively suboptimal spawning habitats (those below the top 10% of available habitat) was unclear, with MSHM2 outperforming MSHM1 for training data but vice versa for test data. While MSHM2 utilized more direct variables and slightly outperformed MSHM1, the costs of the habitat surveys required for MSHM2 may be prohibitive.

Maxent models discriminate presence from absence based upon the difference between points where muskellunge spawned and the random points that represent available habitat. The low density of muskellunge populations, the large areas used for broadcast spawning, and fact that each lake was only surveyed once increases the likelihood of a failure to detect actual spawning habitat. Therefore, a portion of the available habitat used for modeling was actually utilized for spawning by muskellunge. Available habitat represents pseudo-absence data, for which the maximum AUC value is actually $1-\alpha/2$, where α is the fraction of the available area covered by the species' distribution (Phillips et al. 2006). Maximum potential AUC scores for species with wide distributions across the available habitat are therefore lower. This may be the case with muskellunge, which were found spawning in a variety of habitats.

The models and data from which they were created suggest that muskellunge in northern Wisconsin utilized a wider range of habitat types for spawning than previously thought. For example, while many studies suggest that spawning occurs almost entirely over vegetated areas (e.g., Craig and Black 1986; Werner et al. 1996), the SAV and emergent

variables showed a majority of spawning sites were absent of vegetation. While sand appeared to be the most common substrate upon which muskellunge spawned, it was also the most widely available. Despite evidence of variable recruitment from eggs deposited over certain substrates (e.g., Dombeck et al. 1984; Zorn et al. 1998), the data from this study show that eggs are laid over a diverse set of habitats. This could potentially increase the chance of success for eggs deposited in a suitable microhabitat given interannual variability in water levels, stream flow, groundwater discharge, wind patterns, and temperature. The distances over which eggs are broadcast spawned also may contribute to the variation in spawning habitat found by this and other studies. Evidence for the wide variety of habitats utilized can be found in the AUC values for MSHM1 and MSHM2. Common AUC values for comparably sized data sets are between 0.70 and 0.99 (e.g., Hernandez et al. 2006; Phillips et al. 2006; Peterson et al. 2007). Species with relatively low AUC scores are ones that utilize a wider diversity of habitats. The moderate AUC values for muskellunge spawning habitat indicate the usage of multiple habitat types in northern Wisconsin.

The variation in spawning habitat characteristics may be in part due to the variation in habitats available in individual lakes. For example, while some lakes have abundant vegetation present, others are more sparsely vegetated. Therefore, muskellunge are forced to choose suboptimal spawning habitat based on the available choices in that lake or lake chain. Some lakes had median spawning habitat probabilities near 75%, which indicates a high proportion of the lake is suitable spawning habitat. To deal with such variation, managers can evaluate the lake's best spawning habitat using the probabilities of spawning. By classifying the top 50% of available habitat rated by the models, such as in Figures 7 and 8, managers can identify likely spawning habitat in lakes with high proportions of good or poor habitat.

One advantage to the Maxent modeling process is that the multivariate approach allows interpretation of the relative strength of and interactions between variables. Generally, the model converged upon probability assignments that agreed with the habitat preference analysis. Because habitat survey data were combined into one variable, the comparisons between the model and habitat selection analysis are less direct. However, variables with high electivity scores generally translated to habitat combinations with high likelihoods of spawning in the model. There were two major instances of disagreement. First, MSHM2 assigned relatively low values to areas near wetlands. This disagreement is explained above, and was largely a result of the co-occurrence of wetlands and emergent vegetation categories with a high likelihood for muskellunge spawning. The second disagreement was the low likelihood of spawning assigned to steep slopes. While both the index of electivity and MSHM1 classified steep slopes as likely spawning habitat, MSHM2 assigned a low likelihood. This is partially explained by the fact that the majority of the sites with the steepest slopes (7/11) also belonged to the combined habitat category with gravel substrate and absence of emergent and submersed vegetation. In this case the combined habitat variable superseded the slope variable in describing the likelihood of spawning. The greater relative importance of the combined habitat variable is consistent with the fact that habitat surveys were a more direct measure of spawning site suitability than slope measurements, which may serve partially as a proxy for vegetative or substrate characteristics.

Model gain is a good estimate of the amount of information that each variable contains. As expected, the combined habitat variable had the highest gain in MSHM2. In fact, the combined habitat variable had a higher gain (0.71) than MSHM1 did with all other variables (0.61). A comparison of MSHM1 to MSHM2 with the combined habitat variable

removed showed similar training gains. This is evidence that, aside from the combined habitat variable, both models function nearly equivalently. Besides combined habitat, the groundwater variable explained the most variation in the models. This supported the observations by Dombeck et al. (1984) and Zorn et al. (1998) that spawning sites often were located near groundwater flow. Slope and nearby shallow area were the next most important, likely because of the correlations of these variables with vegetation and dissolved oxygen. Land cover showed intermediate importance, although it provided more information to MSHM1, which lacked in-lake habitat survey data. Finally, curvature and aspect were relatively unimportant in the models, contributing less than 0.1 to the training gain. This means that the average spawning site, using only model outputs from aspect or curvature, is rated by the model as only 1.1 times more likely to be used by spawning muskellunge than available habitat ($e^{0.1} \approx 1.1$). The indirect connection between these two variables and factors that influence spawning success are likely the reason for their low importance.

These models are difficult to compare to the conceptual model of Farmer and Chow-Fraser (2004). While the conceptual model used direct variables (i.e., temperature, DO, spatial separation), MSHM1 and MSHM2 necessarily used indirect measures from surveys or remotely sensed data. However, the conceptual model, MSHM1, and MSHM2 all assigned high importance to substrates larger than the diameter of a muskellunge egg or those that contributed to spatial separation (i.e., CBOM). All three models also considered the presence of SAV an important factor in site selection, specifically mat-forming vegetation. Furthermore, groundwater and the bathymetric variables all likely contributed to stable DO and temperatures that allowed eggs to hatch and fry to survive, which the conceptual model also emphasizes.

Both MSHM1 and MSHM2 model the potential niche for spawning habitat in northern Wisconsin lakes. Fielding and Bell (1997) describe the difficulty with modeling habitat usage of species in competition, noting that the presence of competitors in a given region influences the habitats utilized. There is evidence of competition for spawning habitat between muskellunge and northern pike *Esox lucius*, resulting in muskellunge utilizing different habitats for spawning to avoid predation from northern pike fry (Strand 1986). In these cases, muskellunge are suspected to use off-shore areas with *Chara* spp. to separate the fry from northern pike fry. However, there is also evidence which suggests that muskellunge and pike utilize the same habitats in a system in which the two species have existed sympatrically for at least 600 years (Farrell et al. 1996). In order to compensate for the potential influences of northern pike upon muskellunge spawning habitat preferences, this study used fourteen lakes that contained known populations of northern pike and seventeen that had no record of their presence. This ensured that if muskellunge used different habitats to spawn in the presence of northern pike, these habitats were also included in the model. While muskellunge may also use different habitats in the presence of northern pike, there was no clear difference in spawning habitat usage observed in this study.

Management Implications

While muskellunge spawning habitat has been implicated in declining natural reproduction (Dombeck et al. 1986), attempts to locate spawning habitat in individual lakes have been limited by the time required for surveys. For example, it would be impractical for the WDNR to survey the state's 700+ muskellunge lakes for spawning locations during the limited spawning season. Both MSHM1 and MSHM2 provide an efficient method for

management agencies and conservation groups to designate spawning habitat for protection. The model output provides users with a likelihood of spawning at a given location, from which conservation decisions can be made depending upon cost and feasibility. Furthermore, the visual nature of the model output can provide a useful tool to communicate the location and significance of spawning habitat to the public. While all models represent a simplification of reality, and therefore can not perfectly represent the real world, MSHM1 and MSHM2 are an efficient method for identification of potential spawning sites.

These models were developed for muskellunge in northern Wisconsin's inland lakes, so the transferability outside these systems should be tested. Fish in lotic or Great Lakes ecosystems may select different habitats. Furthermore, the models were created from populations of northern strain muskellunge, and spawning habitat selection may differ between strains (Cook and Solomon 1987). Therefore, the models require further testing for Great Lakes and Ohio River strain muskellunge.

The spawning habitat selection data collected in this study using a relatively bias-free method support many of the current spawning habitat descriptions. The importance of CBOM and aquatic vegetation such as rushes, sedges, and stonewort for muskellunge spawning habitat is clear. Using habitat preferences described by this study, managers can conserve or restore vegetation, CBOM, and riparian habitats, focusing especially upon areas identified as likely spawning habitat by MSHM1 and MSHM2.

References

Allan, J. D. 1995. Stream ecology: structure and function of running waters. Springer Press, Dordrecht, The Netherlands.

Baker, M. E., M. J. Wiley, M. L. Carlson, and P. W. Seelbach. 2003. A GIS model of subsurface water potential for aquatic resource inventory, assessment, and environmental management. Environmental Management 32:706-719.

Bean, T. H. 1908. The muskalonge [sic] of the Ohio basin. Transactions of the American Fisheries Society 37:145-151.

Becker, G. C. 1983. The fishes of Wisconsin. The University of Wisconsin Press, Madison.

Beyer, H. L. 2004. Hawth's analysis tools for ArcGIS. Available: http://www.spatialecology.com/htools. (December 2008).

Bozek, M. A., T. M. Burri, and R. V. Frie. 1999. Diets of muskellunge in northern Wisconsin lakes. North American Journal of Fisheries Management 19:258-270.

Christensen, D. L., B. R. Herwig, D. E. Schindler, and S. R. Carpenter. 1996. Impacts of lakeshore residential development on coarse woody debris in North Temperate lakes. Ecological Applications 6:1143-1149.

Cook, M. F., and R. C. Solomon. 1987. Habitat suitability index models: muskellunge. U.S. Fish and Wildlife Service, Biological Report 82(10.148). Washington, D.C.

Craig, R. E., and R. M. Black. 1986. Nursery habitat of muskellunge in southern Georgian Bay, Lake Huron, Canada. Pages 79-86 *in* Hall, G. E., editor. Managing muskies. American Fisheries Society, Special Publication 15, Bethesda, Maryland.

Crossman, E. J. 1990. Reproductive homing in muskellunge, *Esox masquinongy*. Canadian Journal of Fisheries and Aquatic Sciences 47:1803-1812.

Cummins, K. W. 1962. An evaluation of some techniques for the collection and analysis of benthic samples with special emphasis on lotic waters. American Midland Naturalist 67:477-504.

Dombeck, M. P. 1979. Movement and behavior of the muskellunge determined by radio-telemetry. Wisconsin Department of Natural Resources, Technical Bulletin 113, Madison.

Dombeck, M. P. 1986. Muskellunge habitat with guidelines for habitat management. Pages 208-215 *in* Hall, G. E., editor. Managing muskies. American Fisheries Society, Special Publication 15, Bethesda, Maryland.

Dombeck, M. P., B. W. Menzel, and P. N. Hinz. 1984. Muskellunge spawning habitat and reproductive success. Transactions of the American Fisheries Society 113:205-216.

Dombeck, M. P., B. W. Menzel, and P. N. Hinz. 1986. Natural muskellunge reproduction in midwestern lakes. Pages 122-134 *in* Hall, G. E., editor. Managing muskies. American Fisheries Society, Special Publication 15, Bethesda, Maryland.

Farmer, B., and P. Chow-Fraser. 2004. A conceptual model of muskellunge spawning habitat. Undergraduate thesis. McMaster University, Hamilton, Ontario, Canada.

Farrell, J. M. 2001. Reproductive success of sympatric northern pike and muskellunge in an upper St. Lawrence River bay. Transactions of the American Fisheries Society 130:796-808.

Farrell, J. M., R. G. Werner, S. R. LaPan, and K. A. Claypoole. 1996. Egg distribution and spawning habitat of northern pike and muskellunge in a St. Lawrence River marsh, New York. Transactions of the American Fisheries Society 125:127-131.

Farrell, J. M., R. M. Klindt, J. M. Casselman, S. R. LaPan, R. G. Werner, and A. Sciavone. 2007. Development, implementation, and evaluation of an international muskellunge management strategy for the upper St. Lawrence River. Environmental Biology of Fishes 79:111-123.

Fielding, A. H., and J. F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental Conservation 24:38-49.

Global Energy Concepts, LLC. 2002. Wisconsin wind resource assessment program draft final report. Report of Global Energy Concepts, LLC to the Wisconsin Wind Resource Assessment Program. Available: http://www.focusonenergy.com/files/Document_Management_System/Renewables/windresourceassessmentprogram_finalreport.pdf. (April 2009).

Haas, R. C. 1978. The muskellunge in Lake St. Clair. Pages 334-359. *in* Kendall, R. L., editor. A symposium on selected coolwater fishes of North America. American Fisheries Society, Special Publication 11, Bethesda, Maryland.

Hagerthey, S. E., and W. C. Kerfoot. 1998. Groundwater flow influences the biomass and nutrient ratios of epibenthic algae in a north temperate seepage lake. Limnology and Oceanography 43:1227-1242.

Hanson, D. A., and T. L. Margenau. 1992. Movement, habitat selection, behavior, and survival of stocked muskellunge. North American Journal of Fisheries Management 12:474-483.

Hernandez, P. A., C. H. Graham, L. L. Master, and D. L. Albert. 2006. The effect of sample size and species characteristics on performance of different species distribution modeling methods. Ecography 29:773-785.

Inskip, P. D. 1986. Negative associations between abundances of muskellunge and northern pike: evidence and possible explanations. Pages 135-150 *in* Hall, G. E., editor. Managing muskies. American Fisheries Society, Special Publication 15, Bethesda, Maryland.

Ivlev, V. S. 1961. Experimental ecology of the feeding of fishes. Yale University Press, New Haven, Connecticut.

Johnson, A. J. 2000. Lake characteristics influencing spawning success of muskellunge *Esox masquinongy*. Master's thesis. University of Michigan, Ann Arbor.

Johnson, D. H. 1980. The comparison of usage and availability measurements for evaluating resource preference. Ecology 61:65-71.

Margenau, T. L. 1992. Survival and cost-effectiveness of stocked fall fingerling and spring yearling muskellunge in Wisconsin. North American Journal of Fisheries Management 12:484-493.

Martin, L. 1965. Physical geography of Wisconsin. University of Wisconsin Press. Madison.

Menz, F. C., and D. P. Wilton. 1983. An economic study of the muskellunge fishery in New York. New York Fish and Game Journal 30:12-29.

Miller, L. M., S. W. Mero, and J. A. Younk. In press. The genetic legacy of stocking muskellunge in a northern Minnesota lake. Transactions of the American Fisheries Society.

Miller, M. L., and B. W. Menzel. 1986. Movement, activity, and habitat use patterns of muskellunge in West Okoboji Lake, Iowa. Pages 51-61 *in* Hall, G. E., editor. Managing muskies. American Fisheries Society, Special Publication 15, Bethesda, Maryland.

Murry, B. A., and J. M. Farrell. 2007. Quantification of native muskellunge nursery habitat: influence of body size, fish community composition, and vegetation structure. Environmental Biology of Fishes 79:37-47.

Oehmcke, A. A., L. Johnson, J. Klingbiel, and W. Wistrom. 1958. The Wisconsin muskellunge. Its life history, ecology, and management. Wisconsin Conservation Department, Publication 225, Madison.

Pearce, J., and S. Ferrier. 2000. Evaluating the predictive performance of habitat models developed using logistic regression. Ecological Modelling 133:225-245.

Peterson, A. T., M. Papeş, and M. Eaton. 2007. Transferability and model evaluation in ecological niche modeling: a comparison of GARP and Maxent. Ecography 30:550-560.

Peterson, G. D., T. D. Beard, Jr., B. E. Beisner, E. M. Bennett, S. R. Carpenter, G. S. Cumming, C. L. Dent, and T. D. Havlicek. 2003. Assessing future ecosystem services: a case study of the Northern Highlands Lake District, Wisconsin. Conservation Ecology 7(3). Available: http://www.consecol.org/vol7/iss3/art1. (April 2009).

Phillips, S. J., and M. Dudík. 2008. Modeling species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography 31:161-175.

Phillips, S. J., R. P. Anderson, and R. E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. Ecological Modelling 190:231-259.

Phillips, S. J., M. Dudik, and R. E. Schapire. 2008. Maxent software for species habitat modeling, ver. 3.2.19. Available: http://www.cs.princeton.edu/~schapire/maxent/. (October 2008).

Pierce, R. B., J. A. Younk, and C. M. Tomcko. 2007. Expulsion of miniature radio transmitters along with eggs of muskellunge and northern pike–a new method for locating critical spawning habitat. Environmental Biology of Fishes 79:99-109.

Radomski, P., and T. Goeman. 2001. Consequences of human lakeshore development on emergent and floating-leaf vegetation abundance. North American Journal of Fisheries Management 21:46–61.

Rosenberry, D. O., R. G. Striegl, and D. C. Hudson. 2000. Plants as indicators of focused ground water discharge to a northern Minnesota lake. Groundwater 38:296-303.

Rust, A. J., J. S. Diana, T. L. Margenau, and C. J. Edwards. 2002. Lake characteristics influencing spawning success of muskellunge in northern Wisconsin lakes. North American Journal of Fisheries Management 22:834-841.

Scott, W. B., and E. J. Crossman. 1973. Freshwater fishes of Canada. Fisheries Research Board of Canada, Bulletin 184.

Strand, R. F. 1986. Identification of principle spawning areas and seasonal distribution and movements of muskellunge in Leech Lake Minnesota. Pages 62-73 *in* Hall, G. E., editor. Managing muskies. American Fisheries Society, Special Publication 15, Bethesda, Maryland.

Stockwell, D., and D. Peters. 1999. The GARP modeling system: problems and solutions to automated spatial prediction. International Journal of Geographical Information Science 13:143-158.

Werner, R. G., R. Klindt, and B. Jonckheere. 1996. Vegetative characteristics of muskellunge (*Esox masquinongy*) spawning and nursery habitat in the 1000 Islands section of the St. Lawrence River. Great Lakes Research Review 2(2):29-35.

WDNR (Wisconsin Department of Natural Resources). 1939-1978. Lake survey map series. Wisconsin Department of Natural Resources. Madison.

WDNR. 1996. Wisconsin muskellunge waters. Wisconsin Department of Natural Resources, Publication RS-919-96. Madison.

WDNR. 1998. Wisconsin land cover grid. Wisconsin Department of Natural Resources. Madison.

WDNR. 2007. Wisconsin open water from 1:24,000-scale sources. Wisconsin Department of Natural Resources. Madison.

WDNR. 2009. Wisconsin Department of Natural Resources and Wisconsin aquatic gap mapping application. Available: http://infotrek.er.usgs.gov/wdnrfish/. (April 2009).

Wiley, E. O., K. M. McNyset, A. T. Peterson, C. R. Robins, and A. M. Stewart. 2003. Niche modeling and geographic range predictions in the marine environment using a machine-learning algorithm. Oceanography 16(3):120-127.

Winter, J. D., and M. J. Ross. 1982. Methods in analyzing fish habitat utilization from telemetry data. Pages 273-279 *in* N. Armantrout, editor. Acquisition and utilization of aquatic habitat inventory information. American Fisheries Society, Bethesda, Maryland.

Yost, A. C., S. L. Peterson, M. Gregg, and R. Miller. 2008. Predictive modeling and mapping sage grouse (*Centrocercus urophasianus*) nesting habitat using Maximum Entropy and a long-term dataset from Southern Oregon. Ecological Informatics 3:375-386.

Younk, J. A., and M. F. Cook. 1992. Applications of an angler diary for muskellunge *Esox masquinongy*. Minnesota Department of Natural Resources, Investigational Report 420, St. Paul.

Zar, J. H. 1999. Biostatistical analysis. Prentice-Hall, Inc., New York.

Zorn, S.L., T. L. Margenau, J. S. Diana, and C. J. Edwards. 1998. The influence of spawning habitat on natural reproduction of muskellunge in Wisconsin. Transactions of the American Fisheries Society 127:995-1005.

Tables

Table 1. A summary of lakes used in this study, with their counties, surface area, the year spotlighting surveys were completed, number of spawning sites identified, whether data from that lake were generated from surveys of part of the shoreline (incomplete), and if the lake was partially (partial) or fully (X) surveyed for aquatic vegetation and substrate variables.

Lake	County	Area (ha)	Survey year	Number of spawning sites	Vegetation surveys	Substrate surveys
1. Amik*	Vilas	75	1999	22		
2. Annabelle	Vilas	96	2008	6		X
3. Big Carr†	Oneida	86	2007	3		X
4. Black*	Sawyer	52	1998	12		
5. Birch†	Vilas	214	2007	18	X	X
6. Chippewa*†	Bayfield	128	1998	12		
7. Circle Lily	Vilas	90	2008	11		X
8. Clear†	Oneida	342	2008	9	X	X
9. Fisher	Iron	183	2008	4		X
10. Hancock†	Oneida	105	2007	4		X
11. Harris	Vilas	205	2007	25		X
12. Hasbrook†	Oneida	122	2008	13		X
13. Horsehead	Vilas	95	2007	14		X
14. Jute	Vilas	79	2008	25	X	X
15. Kentuck	Vilas	387	2007	24	partial	X
16. Little Sissabagama*	Sawyer	120	1999	9		
17. Long	Iron	151	2008	2	X	X
18. Long (incomplete)†	Vilas	353	2008	3	X	X
19. Mineral*	Ashland	90	1998	11		
20. Mud/Callahan*	Sawyer	218	1998	45		
21. North Twin (incomplete) †	Vilas	1128	2007	11	X	X
22. Oxbow	Vilas	207	2007	23	X	X
23. Pelican†	Oneida	1451	2008	5	X	X
24. Pine	Iron	126	2008	2	X	X
25. Razorback†	Vilas	146	2008	9		X
26. Roberts†	Forest	183	2007	9	X	X
27. Sevenmile†	Oneida	204	2007	20	X	X
28. Snipe†	Vilas	97	2008	2		X
29. Spider	Oneida	48	2007	3		X
30. Third †	Oneida	42	2007	2	X	X
31. Tiger Cat* (incomplete)	Sawyer	89	1998	28		

* Data from Johnson (2000)

† Contains a verified population of northern pike (WDNR 2009)

Table 2. Characteristics used to categorize submersed aquatic vegetation in the study lakes.

Category	**Description**	**Common Examples**
Complex Leaf	Complex or brush-like leaves, leaves or leaflets typically short	*Ceratophyllum demersum,* *Elodea canadensis,* *Myriophylum sibiricum*
Simple Leaf	Simple, flattened leaves	*Potamogeton* spp.
Mat-Forming & Short Grasses	Simple stem, shorter than approximately 10 cm	*Eleocharis acicularis,* *Vallisneria americana, Chara* spp.
Absent	More than 50% absent vegetation	

Table 3. Characteristics used to categorize emergent aquatic vegetation in the study lakes.

Category	**Description**	**Common Examples**
Sedges	Triangular stem in cross section	*Carex* spp., *Schoenoplectus fluviatilis,* *Glyceria canadensis*
Rushes	Circular stem in cross section, few leaves	*Eleocharis palustris,* *Schoenoplectus acutus,* *Schoenoplectus tabernaemontani,* *Juncus canadensis,* *Juncus effusus*
Cattails	Circular in cross section, tall sword-shaped leaves	*Typha latifolia,* *Typha angustifolia*
Floating Leaf	Floating leaves or broad leaves which are supported at or near the surface	*Nymphaea odorata,* *Nuphar variegata,* *Pontederia cordata,* *Potamogeton natans,* *Sagittaria latifolia*
Absent	100% absent	

Table 4. Results from the χ^2 analyses for differences between spawning site and available habitat frequencies for each variable.

Category	**N**	χ^2	**df**	***P***
Aspect	247	49.069	7	<0.0005
Curvature	247	31.029	2	<0.0005
Emergent vegetation	144	72.109	5	<0.0005
Groundwater flow potential	247	185.258	6	<0.0005
Land cover	241	42.452	3	<0.0005
Nearby shallow area	247	89.345	4	<0.0005
Slope	247	96.673	4	<0.0005
Stream distance	247	0.119	2	0.942
Submersed aquatic vegetation	138	5.791	1	0.016
Substrate	228	34.126	4	<0.0005

Table 5. Summary of muskellunge spawning habitat preferences calculated from spawning site and available habitat data. Categories with fewer than five expected spawning sites were removed from the analysis (†).

Variable	Category	Proportion of spawning sites ± 1 SE	Proportion of available habitat ± 1 SE	Index of electivity	Z	*P*
Aspect (degrees from N)	1 to 45	8.5 ± 1.8	11.4 ± 0.1	-0.14	1.308	0.1936
	46 to 90	14.2 ± 2.2	8.1 ± 0.1	0.27	3.383	0.001
	91 to 135	15.4 ± 2.3	12.6 ± 0.1	0.09	1.143	0.2714
	136 to 180	12.2 ± 2.1	10.1 ± 0.1	0.09	0.987	0.3682
	181 to 225	12.6 ± 2.1	6.8 ± 0.1	0.30	3.459	0.0006
	226 to 270	8.5 ± 1.8	6.2 ± 0.1	0.16	1.399	0.1936
	271 to 315	15.8 ± 2.3	17.4 ± 0.2	-0.05	0.583	0.617
	316 to 360	13.0 ± 2.1	27.4 ± 0.2	-0.36	5.020	0.0002
Curvature	Concave	34.8 ± 3.0	24.6 ± 0.2	0.17	3.665	0.004
	Straight	40.1 ± 3.1	57.6 ± 0.2	-0.18	5.491	0.0002
	Convex	25.1 ± 2.8	17.7 ± 0.2	0.17	2.881	0.0052
Emergent aquatic vegetation	Absent	52.8 ± 4.2	69.9 ± 0.2	-0.14	4.376	0.0002
	Cattails	3.5 ± 1.5	2.3 ± 0.1	0.20	†	†
	Floating	6.3 ± 2.0	11.5 ± 0.2	-0.30	1.854	0.0718
	Rushes	27.1 ± 3.7	8.7 ± 0.2	0.51	7.612	0.0002
	Sedge	9.0 ± 2.4	4.8 ± 0.1	0.30	2.147	0.0358
	Wild Rice	1.4 ± 1.0	2.7 ± 0.1	-0.32	†	†
Groundwater potential (m/d)	<-2.5	13.4 ± 2.2	2.9 ± 0.1	0.64	9.548	0.0002
	-2.49 to -1	7.3 ± 1.7	3.8 ± 0.1	0.31	2.689	0.0094
	- .99 to 0	11.7 ± 2.1	5.6 ± 0.1	0.35	4.024	0.0002
	.01 to 1	16.2 ± 2.3	21.3 ± 0.2	-0.14	1.890	0.0718
	1.01 to 2.5	24.3 ± 2.7	43.5 ± 0.2	-0.28	6.006	0.0002
	2.51 to 5	17.4 ± 2.4	20.1 ± 0.2	-0.07	0.967	0.3682
	> 5	9.7 ± 1.9	2.8 ± 0.1	0.55	6.331	0.0002
Land cover	Agriculture	1.2 ± 0.7	0.9 ± 0.1	0.17	†	†
	Barren	0.8 ± 0.6	0.5 ± 0.1	0.23	†	†
	Coniferous	26.7 ± 2.8	21.5 ± 0.2	0.11	1.900	0.0718
	Deciduous	53.4 ± 3.12	61.7 ± 0.2	-0.07	2.596	0.0124
	Forested wetland	6.1 ± 1.5	11.2 ± 0.1	-0.30	2.463	0.0164
	Grassland	0.4 ± 0.4	0	NA	†	†
	Wetland	11.3 ± 2.0	4.2 ± 0.1	0.46	5.436	0.0002
Nearby shallow area (ha within 100 m)	0 to 0.45	44.1 ± 3.2	23.0 ± 0.2	0.31	7.769	0.0002
	0.45 to 0.9	34.4 ± 3.0	28.7 ± 0.2	0.09	1.925	0.0574
	0.9 to 1.35	15.0 ± 2.3	27.0 ± 0.2	-0.29	4.181	0.0002
	1.35 to 1.8	4.5 ± 1.3	15.7 ± 0.2	-0.56	4.761	0.0002
	1.8 to 2.7	2.0 ± 0.9	5.6 ± 0.1	-0.47	2.307	0.0214
Slope (percent slope)	0 to 3.0	31.2 ± 3.0	56.1 ± 0.2	-0.29	7.824	0.0002
	3.0 to 6.1	28.7 ± 2.9	24.1 ± 0.2	0.09	1.641	0.1096
	6.1 to 9.1	12.6 ± 2.1	8.9 ± 0.1	0.17	1.924	0.0574
	9.1 to 12.2	12.2 ± 2.1	4.4 ± 0.1	0.47	5.715	0.0002
	12.2 to 32.9	15.4 ± 2.3	6.5 ± 0.1	0.41	5.494	0.0002
Submersed aquatic vegetation	Absent	68.6 ± 3.9	78.6 ± 0.2	-0.07	2.762	0.007
	Dissected leaf	0	0.03 ± 0.1	NA	†	†
	Mat forming	27.1 ± 3.7	19.7 ± 0.2	0.16	2.112	0.0358
	Simple leaf	4.2 ± 1.7	1.7 ± 0.1	0.42	†	†
Substrate	Silt	1.6 ± 0.8	2.5 ± 0.1	-0.21	0.674	0.5486
	Sand	55.5 ± 3.2	65.8 ± 0.2	-0.09	3.349	0.001
	Gravel	12.2 ± 2.1	18.3 ± 0.2	-0.20	2.417	0.0164
	Cobble	9.3 ± 1.9	5.1 ± 0.1	0.29	2.871	0.0052
	Boulder	0	0.1 ± 0.1	NA	†	†
	CBOM	13.8 ± 2.2	7.3 ± 0.1	0.31	3.794	0.0002
	Marl	7.7 ± 1.7	0.9 ± 0.1	0.78	†	†

Figures

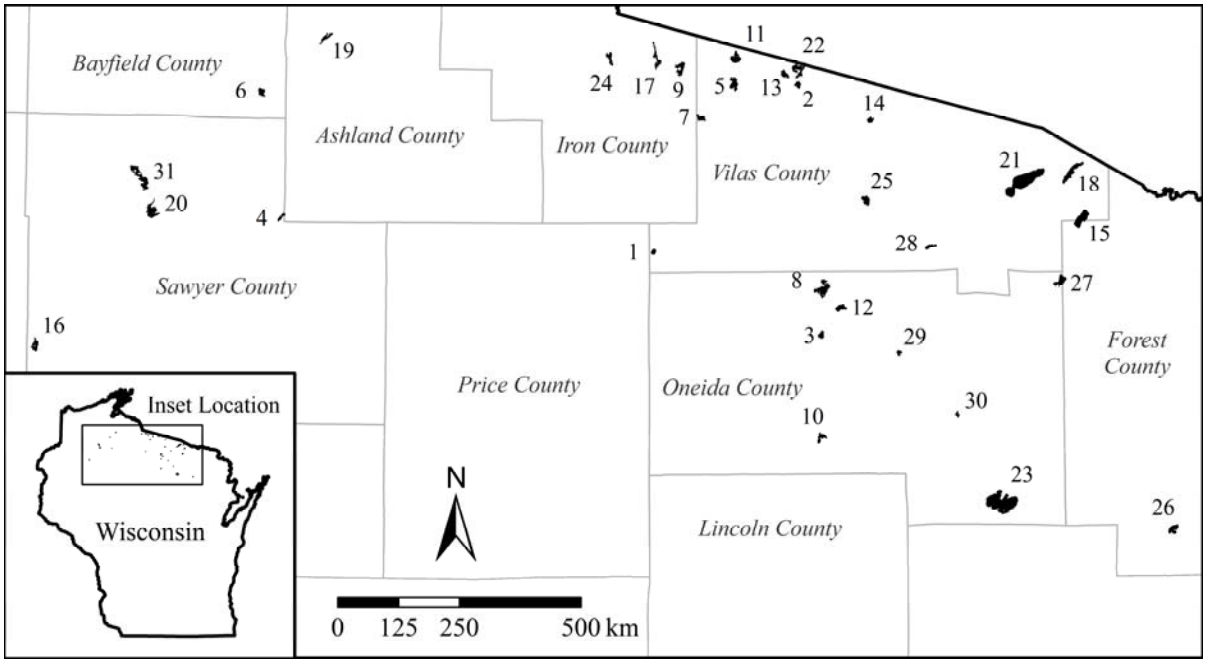

Figure 1. Locations of study lakes (black shapes) in northern Wisconsin, USA. Lakes are numbered for reference to Table 1.

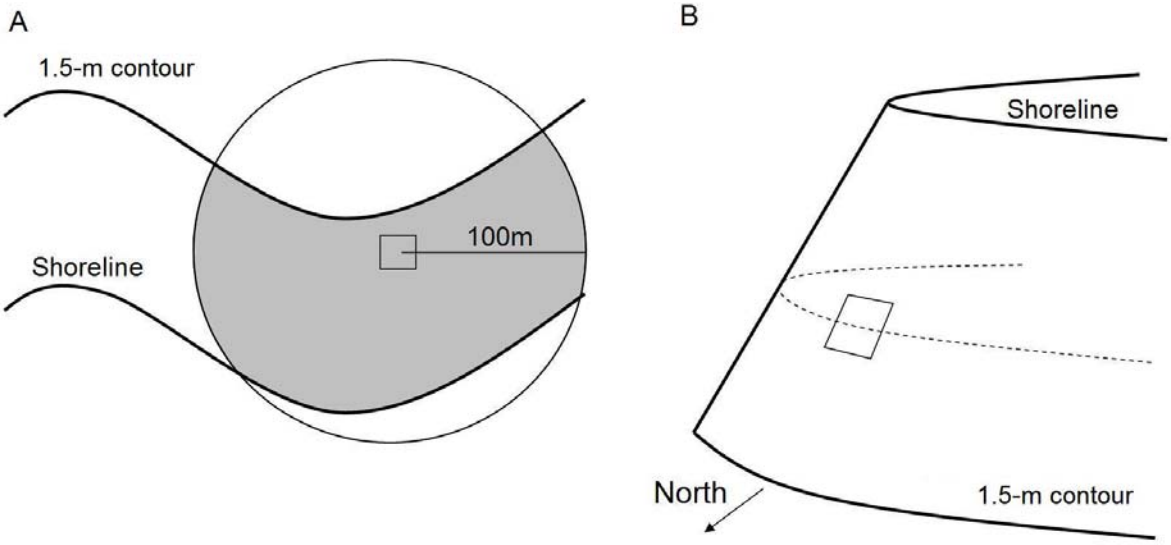

Figure 2. Conceptual diagram of GIS metrics. (A) Nearby shallow area (shaded) was calculated as the area between the shoreline and 1.5-m contour within 100 m of the cell center. (B) Curvature (dashed line) was calculated as the second derivative of the bathymetric contour that transected each grid cell. The aspect value for the cell shown, which faces north, would be 180°.

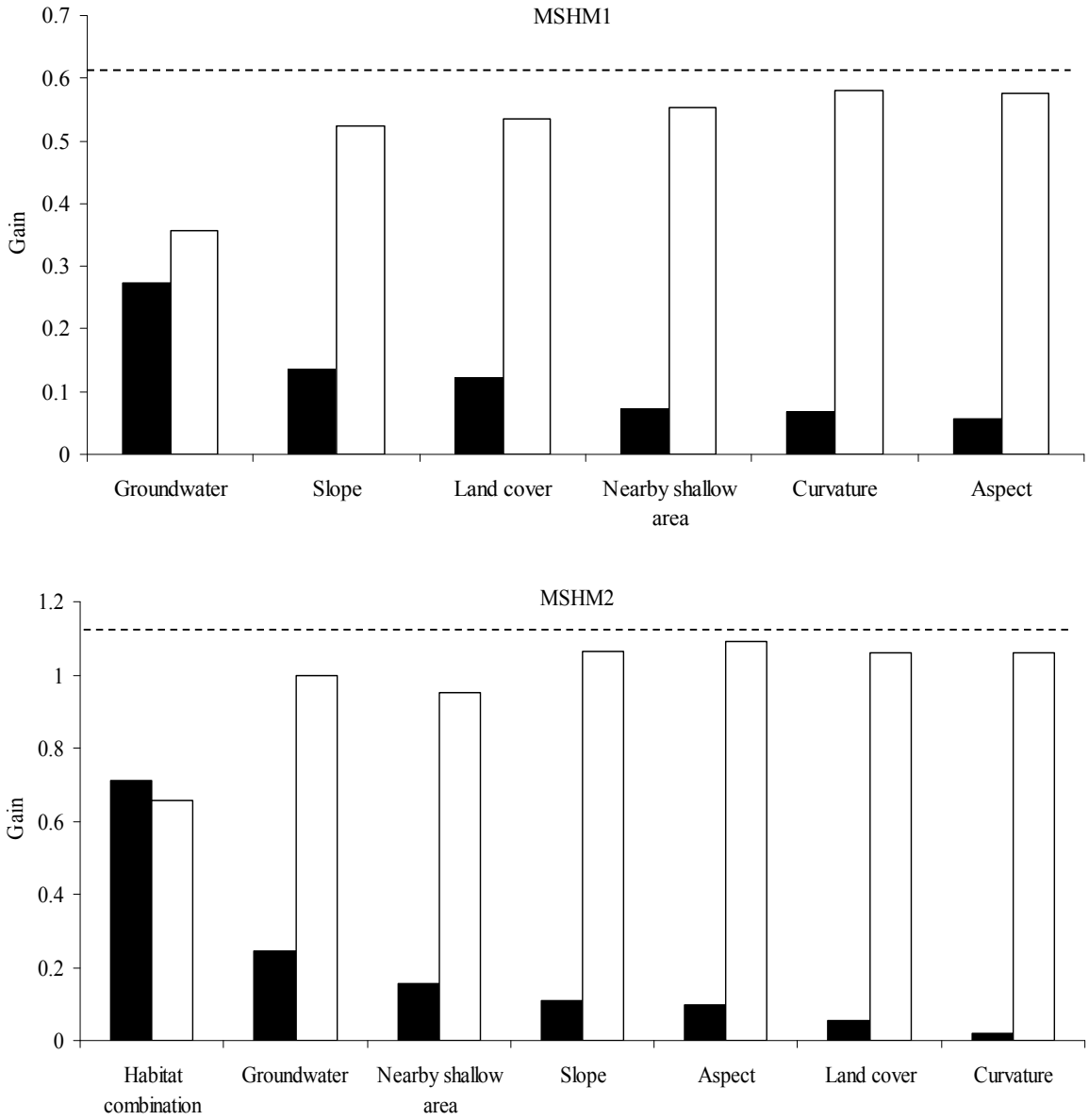

Figure 3. Training gain for MSHM1 (top) and MSHM2 (bottom). The training gain for a model created using only one variable (black) indicates the strength of a model using only that variable. The training gain for a model created without that variable (white) indicates the strength of a model using all variables except that one. Dashed lines represent the training gain for MSHM1 and MSHM2 with all variables included for reference.

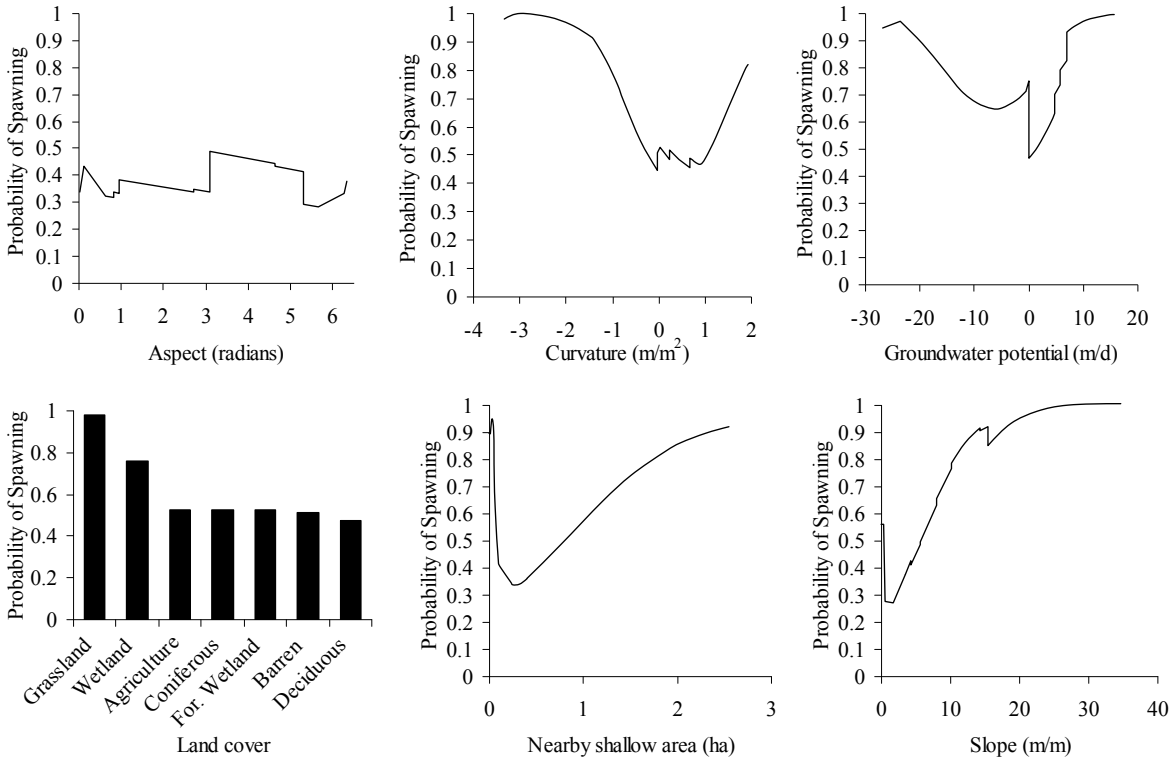

Figure 4. Probability of spawning from MSHM1 for each variable, plotted across the range of each significant variable with all other variables held at their average values. Actual MSHM1 outputs may differ based on interactions between variables.

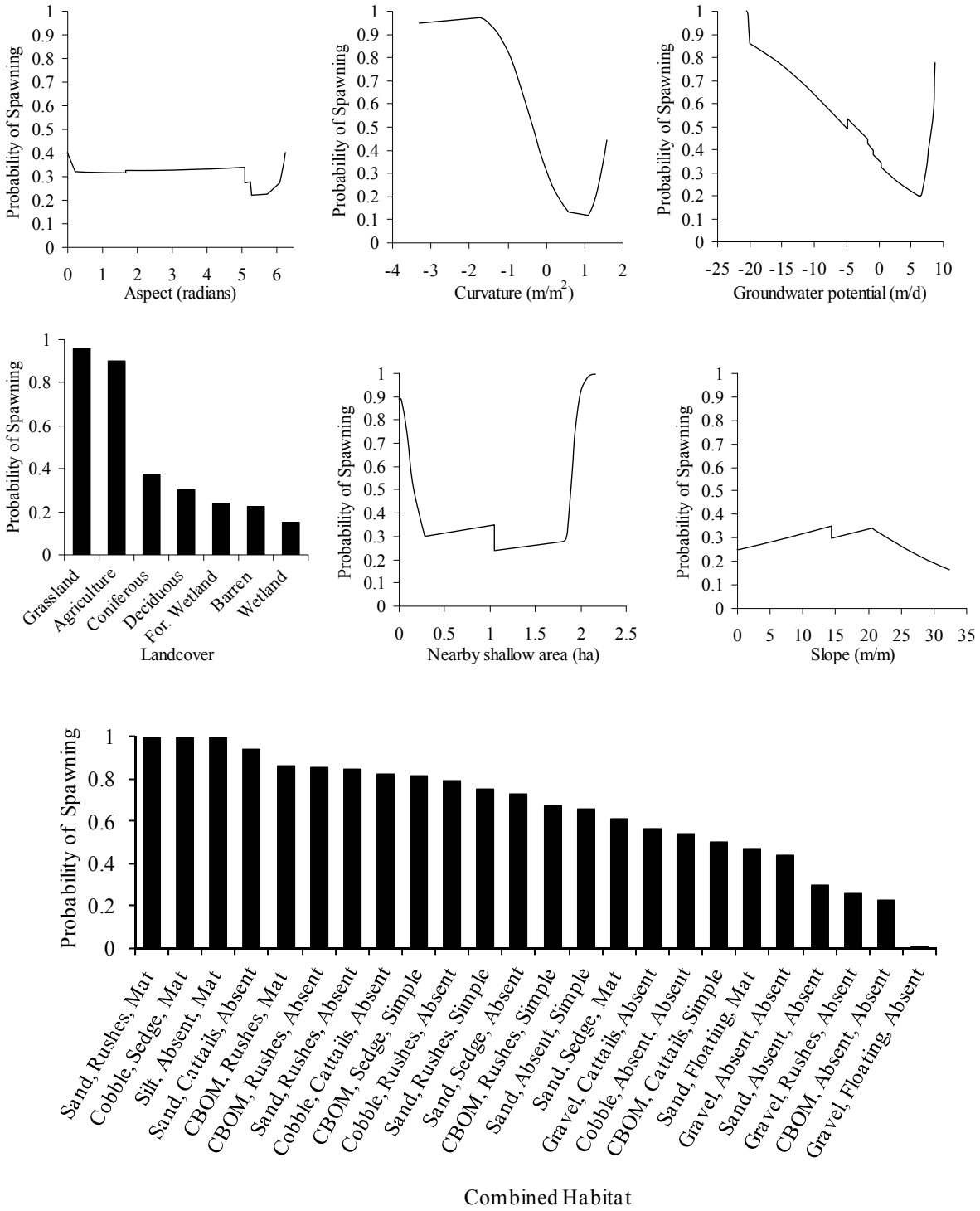

Figure 5. Probability of spawning from MSHM2 for each variable, plotted across the range of each variable with all other variables held at their average values. Combined habitat categories are represented by the abbreviations for substrate, emergent, and submersed aquatic vegetation respectively. All combined habitat categories not shown here had a probability of 0.46 ± 0.01, which is greater than the lowest four categories. Actual MSHM1 outputs may differ based on interactions between variables.

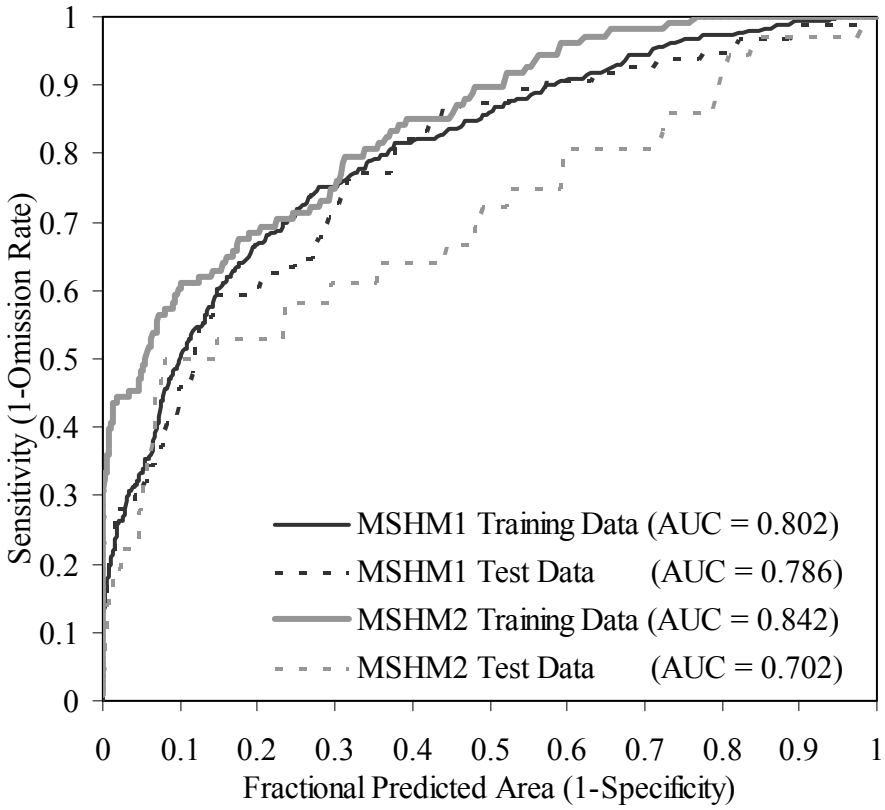

Figure 6. Training and test area under the curve analyses for MSHM1 and MSHM2 indicate a reasonable ability to discriminate between locations with muskellunge spawning present and absent.

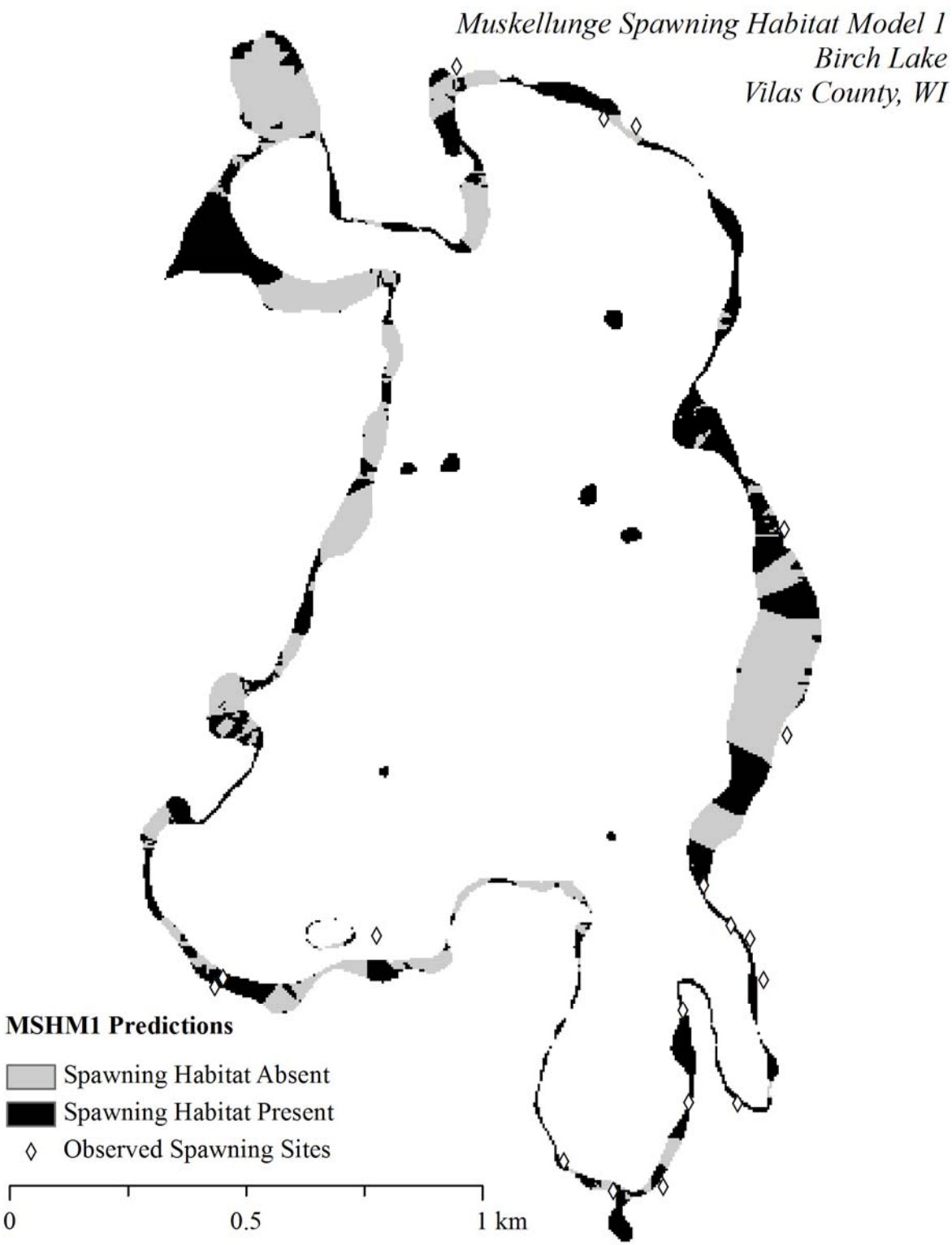

Figure 7. MSHM1 applied to Birch Lake, Vilas Co., WI. Black shading indicates areas rated in the top 50% of spawning habitat in this lake by MSHM1. Spawning sites (diamonds) corroborate the model's performance, as the majority (14/18) were located in areas the model classified as spawning habitat.

Figure 8. MSHM2 applied to Birch Lake, Vilas Co., WI. Black shading indicates areas rated in the top 50% of spawning habitat in this lake by MSHM2. Spawning sites (diamonds) corroborate the model's performance, as the majority (13/18) were located in areas the model classified as spawning habitat.