# **Species distribution model of Warmouth (*Lepomis gulosus*) at Point Pelee National Park**

Megan McCusker

Central and Arctic Region
Fisheries and Oceans Canada
867 Lakeshore Road
Burlington, ON
L7S 1A1

2017

**Canadian Manuscript Report of**
**Fisheries and Aquatic Sciences 3143**

Fisheries and Oceans Canada  Pêches et Océans Canada

Canada

## Canadian Manuscript Report of Fisheries and Aquatic Sciences

Manuscript reports contain scientific and technical information that contributes to existing knowledge but which deals with national or regional problems. Distribution is restricted to institutions or individuals located in particular regions of Canada. However, no restriction is placed on subject matter, and the series reflects the broad interests and policies of Fisheries and Oceans Canada, namely, fisheries and aquatic sciences.

Manuscript reports may be cited as full publications. The correct citation appears above the abstract of each report. Each report is abstracted in the data base *Aquatic Sciences and Fisheries Abstracts*.

Manuscript reports are produced regionally but are numbered nationally. Requests for individual reports will be filled by the issuing establishment listed on the front cover and title page.

Numbers 1-900 in this series were issued as Manuscript Reports (Biological Series) of the Biological Board of Canada, and subsequent to 1937 when the name of the Board was changed by Act of Parliament, as Manuscript Reports (Biological Series) of the Fisheries Research Board of Canada. Numbers 1426 - 1550 were issued as Department of Fisheries and Environment, Fisheries and Marine Service Manuscript Reports. The current series name was changed with report number 1551.

## Rapport manuscrit canadien des sciences halieutiques et aquatiques

Les rapports manuscrits contiennent des renseignements scientifiques et techniques qui constituent une contribution aux connaissances actuelles, mais qui traitent de problèmes nationaux ou régionaux. La distribution en est limitée aux organismes et aux personnes de régions particulières du Canada. Il n'y a aucune restriction quant au sujet; de fait, la série reflète la vaste gamme des intérêts et des politiques de Pêches et Océans Canada, c'est-à-dire les sciences halieutiques et aquatiques.

Les rapports manuscrits peuvent être cités comme des publications à part entière. Le titre exact figure au-dessus du résumé de chaque rapport. Les rapports manuscrits sont résumés dans la base de données Résumés des sciences aquatiques et halieutiques.

Les rapports manuscrits sont produits à l'échelon régional, mais numérotés à l'échelon national. Les demandes de rapports seront satisfaites par l'établissement auteur dont le nom figure sur la couverture et la page du titre.

Les numéros 1 à 900 de cette série ont été publiés à titre de Manuscrits (série biologique) de l'Office de biologie du Canada, et après le changement de la désignation de cet organisme par décret du Parlement, en 1937, ont été classés comme Manuscrits (série biologique) de l'Office des recherches sur les pêcheries du Canada. Les numéros 901 à 1425 ont été publiés à titre de Rapports manuscrits de l'Office des recherches sur les pêcheries du Canada. Les numéros 1426 à 1550 sont parus à titre de Rapports manuscrits du Service des pêches et de la mer, ministère des Pêches et de l'Environnement. Le nom actuel de la série a été établi lors de la parution du numéro 1551.

Canadian Manuscript Report of
Fisheries and Aquatic Sciences 3143

2017

# SPECIES DISTRIBUTION MODEL OF WARMOUTH (*LEPOMIS GULOSUS*) AT POINT PELEE NATIONAL PARK.

by

M. McCusker

Central and Arctic Region
Fisheries and Oceans Canada
867 Lakeshore Road
Burlington, ON
L7S 1A1

PDF version: Cat. No. Fs97-4/3143E-PDF  ISBN 978-0-660-23984-2 ISSN 1488-5387

Correct citation for this publication:

McCusker, M. 2017. Species distribution model of Warmouth (*Lepomis gulosus*) at Point Pelee National Park. Can. Manuscr. Rep. Fish. Aquat. Sci. 3143: iv + 26 p.

## TABLE OF CONTENTS

## ABSTRACT

Warmouth (*Lepomis gulosus*) has been assessed as ‘Endangered’ in Canada (COSEWIC 2015), although it is listed as ‘Special Concern’ under the *Species at Risk Act*. Species distribution modeling was conducted for Warmouth at Point Pelee National Park to map suitable habitat within the park for conservation planning. Species distribution modeling was conducted with machine learning using environmental layers (substrate, vegetation, depth, temperature, water quality) within a geographic information system. Warmouth is widespread in Point Pelee National Park and present in all seven of the major ponds. This model provided insight into the location of the most suitable habitat for Warmouth in the park and the environmental characteristics of that habitat. Results of the model indicated that Warmouth habitat at Point Pelee National Park is characterized by shallow water, abundant aquatic vegetation, low turbidity, low pH, high conductivity, and lack of clay substrate.

## RÉSUMÉ

Le crapet sac-à-lait (*Lepomis gulosus*) a été désigné comme étant une espèce en péril au Canada (COSEPAC 2015), bien qu’il ait été inscrit sur la liste des espèces préoccupantes aux termes de la *Loi sur les espèces en péril*. Une modélisation de la répartition de l’espèce a été effectuée pour le crapet sac-à-lait dans le parc national de la Pointe-Pelée dans le but d’y cartographier un habitat convenable aux fins de planification de la conservation. Une modélisation de la répartition de l’espèce a été effectuée à l’aide d’une méthode d’apprentissage automatique basée sur des aspects environnementaux (le substrat, la végétation, la profondeur, la température, la qualité de l’eau) au sein d’un système d’information géographique. Le crapet sac-à-lait est très répandu dans le parc national de la Pointe-Pelée et est présent dans sept des étangs majeurs. Ce modèle a permis de déterminer l’habitat le plus convenable pour le crapet sac-à-lait dans le parc et d’établir les caractéristiques environnementales de cet habitat. Les résultats du modèle indiquent que l’habitat du crapet sac-à-lait dans le parc national de la Pointe-Pelée se caractérise par des eaux peu profondes, une végétation aquatique abondante, une turbidité faible, un pH faible, une conductivité élevée, et une absence de substrat argileux.

# INTRODUCTION

Freshwater fishes have declined worldwide primarily due to habitat loss related to direct and indirect effects of human activities (Abell 2002; Jelks et al. 2008). The identification of suitable habitat is an essential component of both conservation planning and the protection of habitat for species at risk. Warmouth, a member of the family Centrarchidae, is found in only a few locations in Canada, all on Lake Erie: Point Pelee National Park, Rondeau Bay, Long Point Bay, and Turkey Point Provincial Park (Edwards and Staton 2009). The species is found both inside parks and in unprotected areas. The species is assessed as ‘Endangered’ by COSEWIC (COSEWIC 2015), and listed as ‘Special Concern’ under the *Species at Risk Act*. In the United States, its national status is ‘secure’ (Edwards and Staton 2009). Less than 5% of the global distribution of Warmouth is found in Canada, and Canadian populations are considered vulnerable to habitat alteration, particularly loss of aquatic vegetation (Edwards and Staton 2009).

Warmouth is a warmwater species and its current distribution in Canada may be limited by temperature (COSEWIC 2005). Warming conditions expected with climate change may lead to an expansion of available habitat (Mandrak 1989; Edwards and Staton 2009). Warmouth is thought to prefer heavily vegetated, clear, shallow waters in streams and lakes (Edwards and Staton, 2009). Although it may prefer clear water, the species is considered tolerant of both high turbidity (Larimore 1957; McMahon et al. 1984) and low dissolved oxygen concentrations, with a lower tolerance limit of approximately 3.0 ppm (Larimore 1957). Adults are found in water depths of 0.1-5.0 m with sand or silt substrates (Lane et al. 1996a; COSEWIC 2005), but spawning and nursery habitat is found in depths up to 2 m (Lane et al. 1996b; Lane et al. 1996c). Warmouth habitat generally consists of submergent vegetation growing in sand, silt, or gravel, often with stumps and rocks (COSEWIC 2005). Spawning occurs in the spring on muddy bottoms of streams and lakes at water temperatures of 18-32 °C (Edwards and Staton 2009). The species is a sit-and-wait ambush predator and uses vegetation or debris as shelter (Edwards 1997). Its large mouth allows the species to consume crustaceans, crayfishes, molluscs, small fishes, and aquatic insect larvae (Larimore 1957).

Warmouth were first discovered in Canada in 1966 and Crossman et al. (1996) suggested that Warmouth may have colonized Canadian waters relatively recently (Edwards and Staton 2009). However, the lack of nearby populations from which the species may have expanded its range suggests it may have colonized Canada in the more distant past (COSEWIC 2005; COSEWIC 2015). Although a recent and deliberate introduction of the species into the park is possible, such an explanation seems unlikely. Perhaps a more probable explanation is that the species has been in Canada since the last Ice Age and has gone undetected until its discovery in 1966 (COSEWIC 2015). Sizes of Canadian populations are largely unknown, although a fish community survey conducted at Point Pelee National Park in 2002 and 2003 revealed a sizeable, established population, which may be the healthiest in Canada (Surette 2006; Edwards and Staton 2009).

Point Pelee National Park is situated on the northern shore of Lake Erie (Figure 1). The park extends across a small peninsular sand spit and consists of marsh and woodland habitat. Marshes comprise about 70% of the park and are separated from Lake Erie by sandy beaches on either side that function to close the marshland from lake water, except when lake water levels are very high (Surette 2006). As such, limited potential exists for Warmouth to immigrate into the park from Lake Erie (Edwards and Staton 2009), and Point Pelee essentially represents a closed population.

Species habitats and distributions are increasingly modeled within a Geographic Information System (GIS) framework. MAXENT is a presence-only program that uses a maximum entropy model to identify suitable habitat (Phillips et al. 2006) and has been shown to perform well compared to other species distribution modeling methods (Elith et al. 2006). Here, the distribution and preferred habitat of Warmouth are modeled at Point Pelee National Park using environmental variables within a geographic information system. The results can be used to guide habitat protection for this endangered species.

## METHODS

Data used for this project were provided by Parks Canada, including two ArcGIS shapefiles and several environmental datasets (PPNP Open Water Classification 2010; PPNP ParkBoundary 2013; H Surette Thesis 2002 Data; H Surette Thesis 2003 Data; PPNP Marsh Water Quality Data and Metadata 1971-2014.) All datasets were provided by Point Pelee National Park of Canada (407 Monarch Lane, Leamington, ON N8H 3V4). The available datasets were the Surette 2002 and 2003 data, Parks Canada's Ecological Integrity Monitoring Program (EIMP) survey data from 2008-2015, and an additional small survey conducted in 2005 (Razavi 2006). Species distribution modeling was ultimately conducted using 2003 data from Surette's thesis (Surette 2006), as this was the most comprehensive dataset for any given year (in 2002, water quality data were not collected by Surette, and vegetation data were lacking for some sites). However, data were compiled across all surveys conducted between May and August to provide a composite layer for comparison (see Table 1 for a summary of environmental data considered in this study).

## DATA COLLECTION

### Spatial Extent

ArcGIS shapefiles were provided by Parks Canada for the Point Pelee National Park boundary from 2013 and Point Pelee National Park open-water classification from 2010. The open-water classification layer was simplified for the analysis by removing small ponds that lacked environmental sampling data. Only the seven large ponds in Point Pelee National Park, which had been the focus of environmental surveys, were retained (Sanctuary, Bush, West Cranberry, East Cranberry, Lake, Girardin, and Redhead ponds). The boundary of Lake Erie was downloaded from Scholars GeoPortal (http://geo2.scholarsportal.info/) (Figure 1).

### Substrate

Substrate data were collected in 2002 and 2003 as part of a M.Sc. thesis on the fish communities of the seven large ponds on Point Pelee National Park (Surette 2006).

Percent coverage of substrate types (muck, debris, clay, sand, gravel, pebble, cobble, rubble, boulder) was condensed into three main subtypes (muck/ debris; clay; and, sand/ gravel/ pebble/ cobble/ rubble/ boulder) to simplify analysis (Figure 2). The percent cover of gravel, pebble, cobble, rubble, and boulder was very limited in the dataset, and the inclusion of more variables increased the probability of over-fitting within MAXENT (Radosavljevic and Anderson 2014).

### Vegetation

Vegetation data, i.e. percent cover of submergent, emergent, and floating vegetation, were also collected in 2002 and 2003 (Surette 2006). Each of these layers was interpolated separately (see below), in addition to a 'no vegetation' layer, which was simply 1- the other vegetation types (Figure 3). Vegetation sampling occurred in all seven ponds in 2003; however, for 2002, vegetation data were not available for Bush Pond, and very limited for Redhead Pond. Therefore, only 2003 data were used for species distribution modeling.

### Water Depth

Minimum and maximum water depths at each site were collected by Surette in 2002 and 2003 (Figure 4). Total depth was also sampled as part of Parks Canada EIMP surveys in spring and summer of 2009. However, EIMP sampling occurred predominantly in the centre of ponds with much smaller spatial coverage than Surette's surveys, therefore, species distribution modeling was conducted using the Surette data only, but composite data layers were compiled and used in correlation analysis.

### Temperature

Temperature data were collected by Surette in 2002 and 2003 in the spring (May-June) and summer (July-August). In 2002, temperature was recorded at all sites but, in 2003, summer data were absent from Sanctuary Pond. Therefore, only the spring temperature data were used in 2003 (Figure 4).

### Water quality

Water-quality data collected in 2003 as part of Surette's fish community survey included conductivity, pH, dissolved oxygen, turbidity, and chlorophyll-a (Surette 2006) (Figure 5). Similar water quality data were also collected by Parks Canada as part of the EIMP surveys 2008-2015; however, EIMP sampling predominantly occurred in the centre of ponds. Not all ponds were sampled every year in EIMP surveys, therefore, species distribution modeling was conducted using the Surette data only, but composite data layers were compiled and used in correlation analysis.

### Warmouth locality data

Fish sampling was conducted across all ponds in 2002 and 2003 using inshore hoop nets, offshore hoop nets, Windermere nets, minnow traps, trap nets, bag seines, and straight seines (Surette 2006). In 2002, 353 Warmouth specimens were collected ranging from 26-300 mm in length. In 2003, 304 Warmouth specimens were collected ranging from 27-197 mm in length. In total, 657 specimens were captured. The spatial extent of sampling conducted in 2003 from Surette (2006), including that of specific gear types, is shown in Figure 6. In addition to the Surette survey data, two Warmouth specimens were caught by Fisheries and Oceans Canada in 2002 as part of a Spotted Gar survey at Point Pelee National Park, both of which were located close to sites from Surette's survey in 2002.

### *Phragmites australis*

Riparian vegetation data were also collected as part of Surette's survey data in 2002 and 2003 (Surette, 2006). The common reed, *Phragmites australis*, was found along the eastern shore of Point Pelee, specifically, on Bush, Lake, and Redhead ponds in both years. These data were used to assess potential threats to the species, but were not included in any formal analysis.

## SAMPLE LOCALITY AND ENVIRONMENTAL DATA ANALYSIS IN ARCGIS

Environmental data were interpolated using the Spline with Barriers function in ArcMAP 10 with a cell size of 0.0001 (or approximately 10 m x 10 m). Spline with Barriers was used to maintain pond integrity and to avoid interpolating data from one pond into another. The geographic extent of layers was defined using the open-water classification shapefile provided by Parks Canada, edited to only include the seven main ponds. Warmouth sampling data were edited slightly before inclusion in the species distribution modeling. If two Warmouth localities occupied the same cell within ArcGIS, one was removed for a maximum of one locality per cell. This ensured that training and testing sites were distinct and did not have overlapping cells. To be conservative, one sample locality was randomly removed when two sites were less than 20 m from one another. In total, 82 distinct Warmouth localities in 2002, and 63 distinct localities in 2003 were included in the study. A small number of Warmouth localities were moved (<2 m) to fit on the grid in ArcGIS.

Interpolated environmental data were collected (for more than 900 data points) for correlation analysis using the 'Extract Multi Values to Points' function in ArcGIS as sample sites varied slightly among surveys and years. Correlation analysis was conducted to evaluate whether environmental variables showed consistent patterns across datasets (and were, therefore, reliable environmental indicators) and to assess multicollinearity among variables. Correlations were examined using the package 'corrplot' in RStudio ver. 0.99.473 (RStudio Team 2015). Species distribution modeling was conducted using those environmental variables that exhibited consistency across years, but were not highly correlated with other environmental variables.

## RUNNING MAXENT

The distribution of Warmouth was modeled using MAXENT (Maximum Entropy Species Distributional Modeling, Version 3.3.3k), a presence-only, machine-learning method based on maximum entropy (Phillips et al. 2006). Like all species distribution modeling methods, MAXENT assumes that all environmental variables have ecological importance to the species and that the species' distribution is limited by environmental variables rather than other factors (e.g. physical barriers). Another assumption is that environmental variables are not highly correlated with one another, although MAXENT is fairly robust to correlations among environmental variables (Elith et al. 2011). A regularization parameter of 3.5 was used in MaxEnt to minimize over-fitting (Radosavljevic and Anderson 2014). Analyses were performed with 5-fold cross-validation method. The data were partitioned into five 'folds', with one fold (20% of the data) used as the 'test' data, and the remaining 80% of the data (the training data) used to build the model. Five models were created, with each fold acting as the 'test' data once. The final model prediction was based on an average of all the models.

Model performance was evaluated using the area under the receiver-operating curve (AUC), a threshold-independent estimate of statistical support (Phillips *et al*., 2006). The AUC represents the fit of the model or the ability of the model to distinguish between locations where the species was found and all other locations. Models with an AUC of 0.9 are often considered outstanding, 0.8 is considered good, and 0.7 acceptable (Hosmer and Lemeshow 2000). Elith et al. (2006) suggested that an AUC value of 0.75 indicates a useful model for understanding species distributions with MaxEnt. The one-tailed binomial probability that the model predicted the test data no better than random was evaluated using two different threshold values (maximum test sensitivity plus specificity threshold, and the $10^{th}$ percentile training presence threshold) to distinguish 'suitable' from 'unsuitable' habitat. Each environmental variable was also examined to assess its importance within the model using the jackknife approach in MaxEnt. Importance of each variable is estimated by the decrease in AUC when the parameter in question is randomly permuted across the study site.

To address the possibility of spatial bias in sampling effort, a second species distribution model was created in MaxEnt using only Warmouth localities from the black hoop net and white hoop net as these gear types show the least amount of bias (Figure 6). Given the small sample size of Warmouth localities (n=16), a 10-fold cross validation was used to allow more samples in the training dataset (n=14) and n=2 for testing. This analysis was conducted for comparative purposes to evaluate how sampling bias may have influenced the model.

## RESULTS

### DATA AND CORRELATIONS

Strong correlations were observed between Surette 2003 and composite data layers for substrate, vegetation, and depth (Figure 7), illustrating that Surette 2003 data were representative of all available data. Muck/debris was highly inversely correlated with sand/cobble (r = -0.9). Therefore, one layer (muck/debris layer) was arbitrarily removed to comply with MaxEnt assumptions about correlated variables. The 'no vegetation' layer was inversely correlated with submergent vegetation (r = -0.7), but the correlation was considered moderate and both variables were retained in the MaxEnt model.

Correlations among temporal estimates of water-quality data revealed that turbidity (r~0.7), pH (r~0.7), and conductivity (r~0.6) were all relatively stable across years (Figure 8). However, dissolved oxygen and chlorophyll had notably lower consistency across surveys, with r~0.4 and r~0.3, respectively. Therefore, only turbidity, pH, and conductivity were used for species distribution modeling.

### SPECIES DISTRIBUTION MODELING

The test AUC value for the species distribution model using 2003 data was 0.78, with AUC values ranging from ~0.6-0.85 for individual variables (Figure 9). The one-tailed binomial probability was highly significant ($p<0.0005$) using the maximum test sensitivity plus specificity threshold. The final model specifications indicated that Warmouth habitat was characterized by limited clay substrate. Suitable habitat tended to be characterized by minimal depths, vegetation (submergent and floating), low turbidity, low pH, and high conductivity, and sand/cobble substrate (Figure 10). Probability of suitable habitat varies fairly linearly with changes in many of these environmental characteristics, although conductivity values > 350 µS/cm) and submergent vegetation > 25% may represent important thresholds (Figure 10). Minimum depth was considered the most important variable within the model (38% importance), followed by floating vegetation (28% importance), clay substrate (10%), and pH, conductivity, turbidity, and submergent vegetation (all ~5%).

Good quality habitat was predominantly found in the eastern portion of Point Pelee National Park. Virtually all of Bush and Redhead ponds were considered highly suitable, as well as large portions of East Cranberry and Girardin ponds (Figure 11). Within West Cranberry and Lake Pond, much of the shoreline was considered suitable, although the species was not particularly abundant in West Cranberry Pond. Sanctuary Pond provided the least amount of suitable habitat as the species was only found at two sites in 2003 and none in 2002. These patterns are also easily visualized with a 'maximum test sensitivity plus specificity' threshold (of 0.38), and the $10^{th}$ percentile training presence threshold (of 0.23) (Figure 12).

A comparison of the original model with a second model created with a more limited locality dataset (black and white hoop net data only) is shown in Figure 13. The test AUC value for this model was 0.689, illustrating lower confidence with fewer data points. This model indicated that environmental characteristics of suitable habitat were similar to those previously identified, including minimal depths, limited clay substrate, vegetation (submergent and floating), low turbidity, low pH, and high conductivity, although the presence of sand and cobble substrate is no longer a good predictor (Figure 14).

## DISCUSSION

Warmouth is widely distributed in Point Pelee National Park. Statistical support (AUC values) for the habitat model was strong, but not excellent, perhaps reflecting the difficulty of distinguishing suitable from unsuitable habitat when the species is so widespread within the study area. With the exception of most of Sanctuary Pond and, perhaps, the inner parts of West Cranberry Pond, most of the freshwater habitat at Point Pelee National Park appears to provide suitable habitat for Warmouth. Habitat suitability models were highly consistent with previous expectations regarding Warmouth habitat characteristics (Edwards and Staton, 2009). Shallow depths, abundant aquatic vegetation, low turbidity, low pH, and high conductivity were all predictive of suitable habitat. Clay substrate appears to be less favorable, perhaps because vegetation is less abundant on this substrate. However, clay was only found in a few isolated patches, so additional evidence would be required to conclude that clay is not suitable. Sand/cobble and muck/debris both appear to provide suitable habitat, although sand/cobble was slightly more favorable, perhaps because it was found in shallower parts of ponds.

The sampling surveys of Surette (2006) were extensive and invaluable to this study. However, some potential sources of bias were identified, including a sampling regime that was weighted towards nearshore samples and the use of a variety of fish-capture methods. Soft substrates and extensive macrophyte coverage in some areas did not allow for uniform sampling methods throughout the study area. To address this bias, a second analysis was performed using gear types that appeared to have low bias in their distribution (i.e. black and white hoop nets). The green hoop net sampling was biased towards the nearshore areas. The bag seine and trap nets were biased towards the eastern part of the park, although these gear types did not contribute many sites, and hardly any unique sites to the study. The second model largely supported the previous findings, albeit with lower overall support, possibly due to the smaller sample size of Warmouth localities. One notable exception was that sand and cobble substrate was not associated with higher predicted suitability in the second model (Figure 14) unlike in the first (Figure 10). Sand and cobble were dominant in the eastern part of Lake Pond, an area that was not well sampled by black and white hoop nets (Figure 6).

The degree to which bias has influenced the model depends on whether Warmouth localities are representative of the true distribution of the species. MAXENT was, in part, designed to take advantage of museum collections where presence localities were recorded, but absence data were not (Phillips et al. 2006). The method relies on the assumption that over a long period of time, an area has been well sampled. It follows that if the selection of sampling sites at Point Pelee National Park was informed by a history of sampling and knowledge about where the species is found, then the model will still perform well despite apparent bias. Sampling in this study may well have been biased towards true Warmouth habitat. For example, the high capture rate in the nearshore areas by the green hoop net (Figure 6), suggests that this gear type may have sampled predominantly in true Warmouth habitat. Nevertheless, given that the distribution of 'true' Warmouth habitat is not known, and given the bias in sampling site selection, a second model was created for comparison. In general, the similarity between the two models suggests that the models are performing relatively well despite some bias in the sampling. Characteristics of suitable habitat, according to both models, were largely consistent with findings from the literature (e.g., preference for shallow water, abundant vegetation, sand and silt substrate, and low turbidity) (COSEWIC 2005).

Conservation planning requires the identification of suitable habitat, however, the selection of an appropriate threshold to differentiate suitable from unsuitable habitat is not always clear. Threshold selection has implications for whether error is primarily that of omission (leaving out suitable habitat) or commission (including habitat that is not actually suitable). Although the 'maximum test sensitivity plus specificity' threshold has been found to perform well compared to other threshold methods (Liu et al. 2005), this threshold may be too high for a conservation application (Phillips et al. 2006). Lower thresholds are likely warranted for species at risk as the implications for omission are worse than for commission when identifying habitat for protection. In this case, a low threshold seems warranted as all ponds clearly provide at least some suitable habitat.

Point Pelee National Park contains possibly the best habitat for Warmouth in Canada and the population is thought to be healthy (Edwards and Staton 2009). As Point Pelee National Park is a protected area, the risk of habitat destruction is low.

However, climate change and invasive species still present a threat. Warmouth was identified as a species at risk with high sensitivity to climate change due to its reliance on nearshore habitat and its limited distribution in Canada (Doka et al. 2006). Furthermore, lower water levels from increased evaporation rates due to climate change will likely negatively affect the extent of Warmouth habitat in the park. Conversely, Warmouth is considered a warmwater species and may even benefit from thermal changes associated climate change (Edwards and Staton 2009). Therefore, the ultimate impact of climate change on Warmouth populations at Point Pelee is unclear and will likely depend on the magnitude of changes to water temperature and evapotranspiration. In addition to climate change mediated alterations, *Phragmites australis* presents a potential threat to habitat, which may be exacerbated as water levels decrease. *Phragmites australis* can have a significant impact on nearshore environments as it tends to dominate wetlands with its large dense stems, outcompeting other types of vegetation (Bourgeau-Chavez et al, 2015) and converting open water to dense stands of emergent growth (Crisman et al. 2014). Surveys in 2003 identified *Phragmites australis* along the eastern border of the park, where some of the most suitable habitat for Warmouth exists according to this study. Therefore, continued monitoring is required, and active management of *Phragmites australis* within Point Pelee National Park may be warranted.

## ACKNOWLEDGEMENTS

This study would not have been possible without help from Hiroo Sawada, Geomatics Technician, and Tammy Dobbie, Park Ecologist with Parks Canada at Point Pelee National Park who provided environmental data for this study. The geo-referenced fish and habitat data, collected as part of Heather Surette's M.Sc. thesis, were invaluable to this study.

## REFERENCES

Abell R (2002) Conservation biology for the biodiversity crisis: a freshwater follow-up. Cons. Bio. 16, 1435-1437.

Bourgeau-Chavez L, Endres S, Battaglia M, Miller ME, Banda E, Laubach Z, Higman P, Chow-Fraser P, Marcaccio J (2015) Development of a bi-national Great Lakes coastal wetland and land use map using three season PALSAR and landsat imagery. Remote Sens. 7: 1-27.

COSEWIC (2005) COSEWIC assessment and update status report on the warmouth *Lepomis gulosus* in Canada. Committee on the Status of Endangered Wildlife in Canada. Ottawa. vi + 16 pp. (www.sararegistry.gc.ca/status/status_e.cfm).

COSEWIC (2015) COSEWIC assessment and status report on theWarmouth *Lepomis gulosus* in Canada. Committee on the Status of Endangered Wildlife in Canada. Ottawa. x + 47 pp. (www.registrelepsararegistry.gc.ca/default_e.cfm).

Crisman TL, Alexandridis TK, Zalidis GC, Takavakoglou V (2014) *Phragmites* distribution relative to progressive water level decline in Lake Koronia, Greece. Ecohydrology, 7: 1403-1411.

Doka S, Bakelaar C, Bouvier L (2006) Coastal wetland fish community assessment of climate change in the lower Great Lakes, in: Mortsch L., Ingram J., Hebb A., Doka S. (Eds.), Great Lakes wetland communities: vulnerability to climate change and response to adaptation strategies. Environment Canada and the Department of Fisheries and Oceans. Toronto, Ontario, pp 101-128.

Edwards RJ (1997) Ecological profiles for selected stream-dwelling Texas freshwater fishes. Texas Water Development Board. 89 pp.

Edwards AL, Staton SK (2009) Management plan for the Blackstripe Topminnow, Pugnose Minnow, Spotted Sucker and Warmouth in Canada. *Species at Risk Act* Management Plan Series. Fisheries and Oceans Canada, Ottawa. viii + 43 pp.

Elith J, Graham CH, Anderson RP, Dudı´k M, Ferrier S, Guisan A, Hijmans RJ, Huettmann F, Leathwick JR, Lehmann A, Li J, Lohmann LG, Loiselle BA, Manion G, Moritz C, Nakamura M, Nakazawa Y, Overton J, Peterson AT, Phillips SJ, Richardson KS, Scachetti-Pereira R, Schapire RE, Sobero´n J, Williams S, Wisz MS, Zimmermann NE (2006) Novel methods improve prediction of species' distributions from occurrence data. Ecography 29: 129-151.

Elith J, Phillips SJ, Hastie T, Dudik M, Chee YE, Yates CJ (2011) A statistical explanation of MAXENT for ecologists. Diversity Distrib. 17: 43-57.

Hosmer DW, Lemeshow S (2000) Applied logistic regression, 2nd edition. New York; Toronto; Wiley.

Jelks HL, Walsh SJ, Burkhead NM, Contreras-Balderas S, Diaz-Pardo E, Hendrickson DA, Lyons J, Mandrak N, McCormick F, Platania SP, Porter BA, Renaud CB, Schmitter-Soto JJ, Taylor EB, Warren MLJ (2008) Conservation status of imperiled North American freshwater and diadromous fishes. Fisheries 33: 372-407.

Lane JA, Lane CB, Minns CK (1996a) Adult habitat characteristics of Great Lakes fishes. Canadian Manuscript Report of Fisheries and Aquatic Sciences 2358: v + 43 pp.

Lane JA, Lane CB, Minns CK (1996b) Nursery habitat characteristics of Great Lakes fishes. Canadian Manuscript Report of Fisheries and Aquatic Sciences 2338: v + 42 pp.

Lane JA, Lane CB, Minns CK (1996c) Spawning habitat characteristics of Great Lakes fishes. Canadian Manuscript Report of Fisheries and Aquatic Sciences 2368: v + 48 pp.

Larimore RW (1957) Ecological life history of the warmouth (Centrarchidae). Bull. Illinois Nat. Hist. Surv. 27(1):1-83.

Liu C, Berry PM, Dawson TP, Pearson RG (2005) Selecting thresholds of occurrence in the prediction of species distributions. Ecography 28: 385-393.

Mandrak NE (1989) Potential invasion of the Great Lakes by fish species associated with climatic warming. Journal of Great Lakes Research, 15: 306-316.

McMahon TE, Gebhart G, Maughan OE, and Nelson PC (1984) Habitat suitability index models and instream flow suitability curves: Warmouth. U.S. Fish Wild. Serv. FWS/OBS-82/10.67. 21 pp.

Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. Ecol. Model. 190: 231-259.

Radosavljevic A, Anderson RP (2014) Making better MAXENT models of species distributions: complexity, overfitting and evaluation. J. Biogeogr. 41: 629-643.

Razavi R (2006) Assessing historical and present-day fish habitat in the marshes of Point Pelee National Park. Senior Honours Thesis. Department of Biology. McMaster University, Hamilton. April 2006. 65 pp.
RStudio Team (2015). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL http://www.rstudio.com/.
Surette HJ (2006) Processes influencing temporal variation in fish species composition in Point Pelee National Park. M.Sc. Thesis. University of Guelph, Guelph, ON. 105 pp.

*Table 1. Geo-referenced environmental data collected at Point Pelee National Park. Data used in the final species distribution model are indicated.*

| Source | Year collected | Category | Specific Details | In final model |
|---|---|---|---|---|
| Surette (2006) | 2003 | Substrate | Clay | * |
| Surette (2006) | 2003 | Substrate | Sand, Cobble | * |
| Surette (2006) | 2003 | Substrate | Muck, Debris | |
| Surette (2006) | 2003 | Vegetation | Emergent | * |
| Surette (2006) | 2003 | Vegetation | Submergent | * |
| Surette (2006) | 2003 | Vegetation | Floating | * |
| Surette (2006) | 2003 | Vegetation | No Vegetation** | * |
| Surette (2006) | 2003 | Depth | Minimum Depth | * |
| Surette (2006) | 2003 | Depth | Maximum Depth | * |
| Surette (2006) | 2003 | Temperature | Temp- Spring (May, June) | * |
| Surette (2006) | 2003 | Temperature | Temp- Summer (July, August) | |
| Surette (2006) | 2003 | Water Quality | Dissolved oxygen | |
| Surette (2006) | 2003 | Water Quality | Conductivity | * |
| Surette (2006) | 2003 | Water Quality | pH | * |
| Surette (2006) | 2003 | Water Quality | Turbidity | * |
| Surette (2006) | 2003 | Water Quality | Chlorophyll | |
| Surette (2006) | 2002 | Substrate | Clay | |
| Surette (2006) | 2002 | Substrate | Sand, Cobble | |
| Surette (2006) | 2002 | Substrate | Muck, Debris | |
| Surette (2006) | 2002 | Vegetation | Emergent | |
| Surette (2006) | 2002 | Vegetation | Submergent | |
| Surette (2006) | 2002 | Vegetation | Floating | |
| Surette (2006) | 2002 | Vegetation | No Vegetation** | |
| Surette (2006) | 2002 | Depth | Minimum Depth | |
| Surette (2006) | 2002 | Depth | Maximum Depth | |
| Surette (2006) | 2002 | Temperature | Temp- Spring (May, June) | |
| Surette (2006) | 2002 | Temperature | Temp- Summer (July, August) | |
| Surette (2006) | 2002 | Water Quality | Dissolved oxygen | |
| Surette (2006) | 2002 | Water Quality | Conductivity | |
| Surette (2006) | 2002 | Water Quality | pH | |
| Surette (2006) | 2002 | Water Quality | Turbidity | |
| Surette (2006) | 2002 | Water Quality | Chlorophyll | |
| Razavi (2006) | 2005 | Water Quality | Conductivity | |
| Razavi (2006) | 2005 | Water Quality | pH | |
| Razavi (2006) | 2005 | Water Quality | Turbidity | |
| Razavi (2006) | 2005 | Water Quality | Chlorophyll | |
| EIMP | 2008-2015 | Water Quality | Dissolved oxygen | |
| EIMP | 2008-2015 | Water Quality | Chlorophyll | |
| EIMP | 2008-2015 | Water Quality | Turbidity | |

| Source | Year collected | Category | Specific Details | In final model |
|---|---|---|---|---|
| EIMP | 2008-2015 | Water Quality | Conductivity | |
| EIMP | 2008-2015 | Water Quality | pH | |

** Note: 'No Vegetation' is plotted as 'All Vegetation' in Figure 3 for colour consistency within the figure.

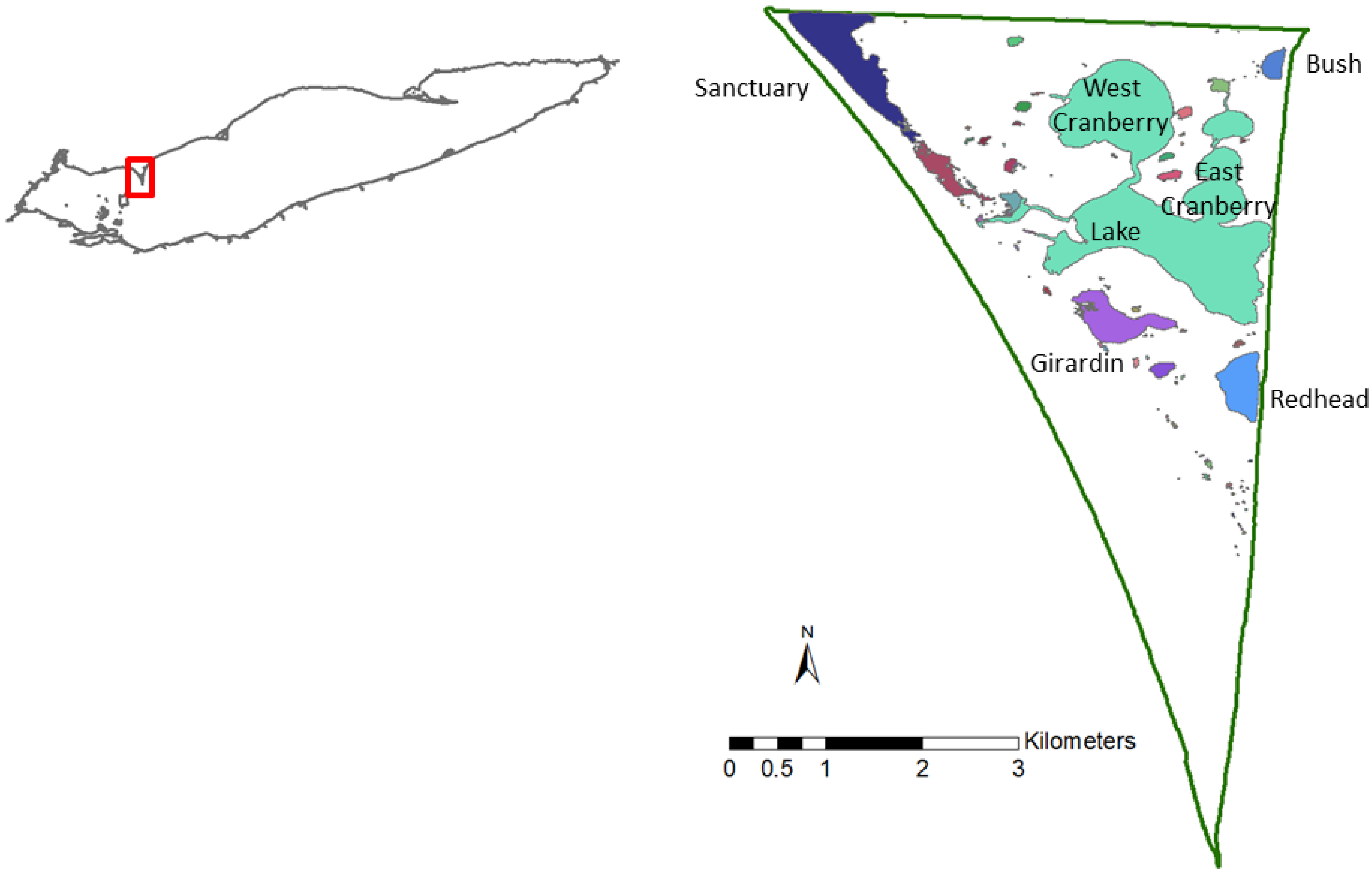


*Figure 1. Lake Erie with Point Pelee National Park highlighted in red (top). The freshwater ponds in Point Pelee National Park, with the park boundary highlighted in green, and pond names indicated (right).*

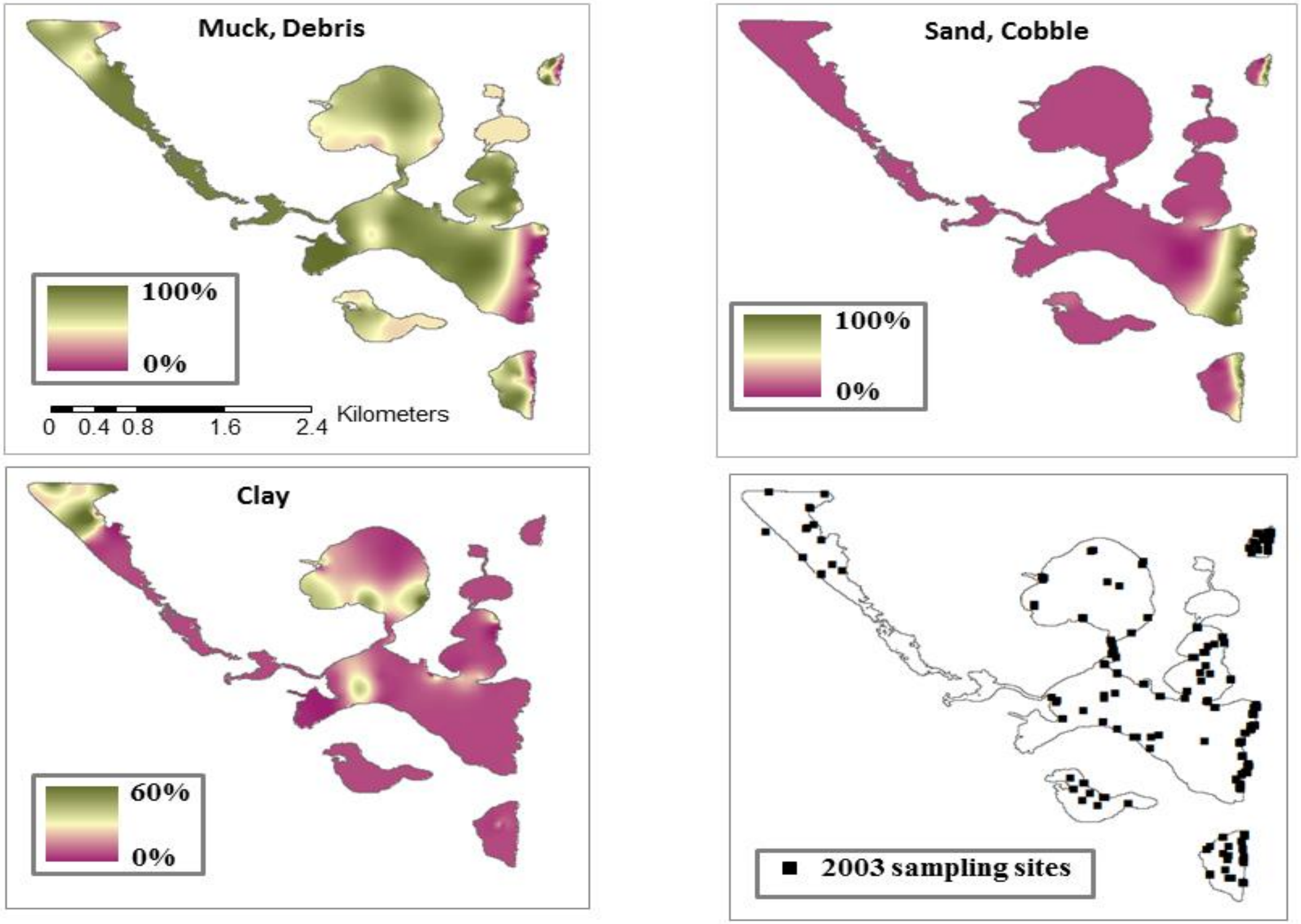


*Figure 2. Interpolated substrate data from ArcGIS. Condensed substrate layers are based on 2003 data from Surette (2006). Sampling sites used by Surette in 2003 (bottom right) apply to substrate, vegetation, and depth (see Figures 3, 4).*

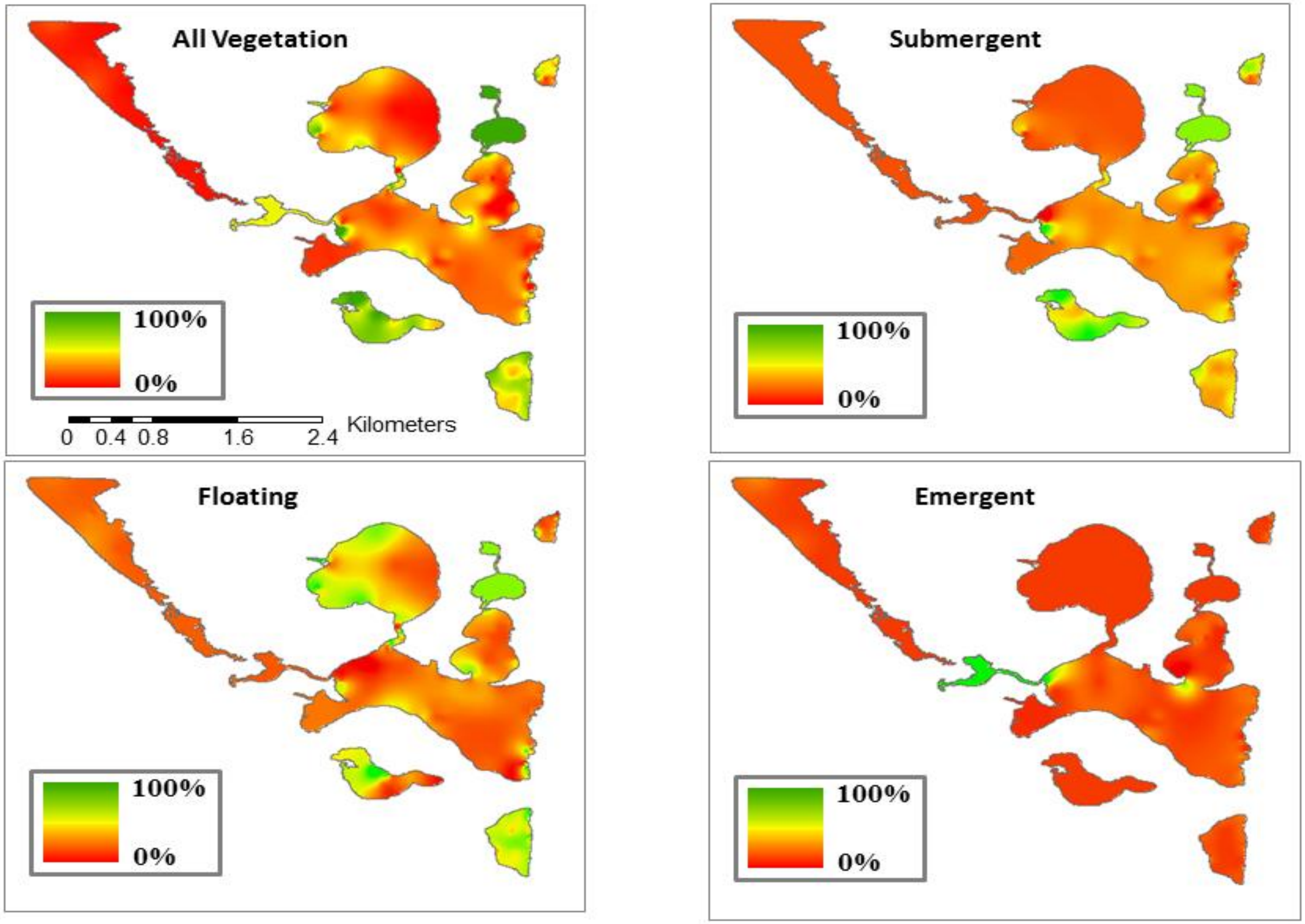


*Figure 3. Interpolated vegetation layers based on 2003 data from Surette (2006). See Figure 2 for sampling sites. Note: 'All Vegetation' is used for colour consistency within the figure ('All Vegetation' = 1 – 'No Vegetation').*

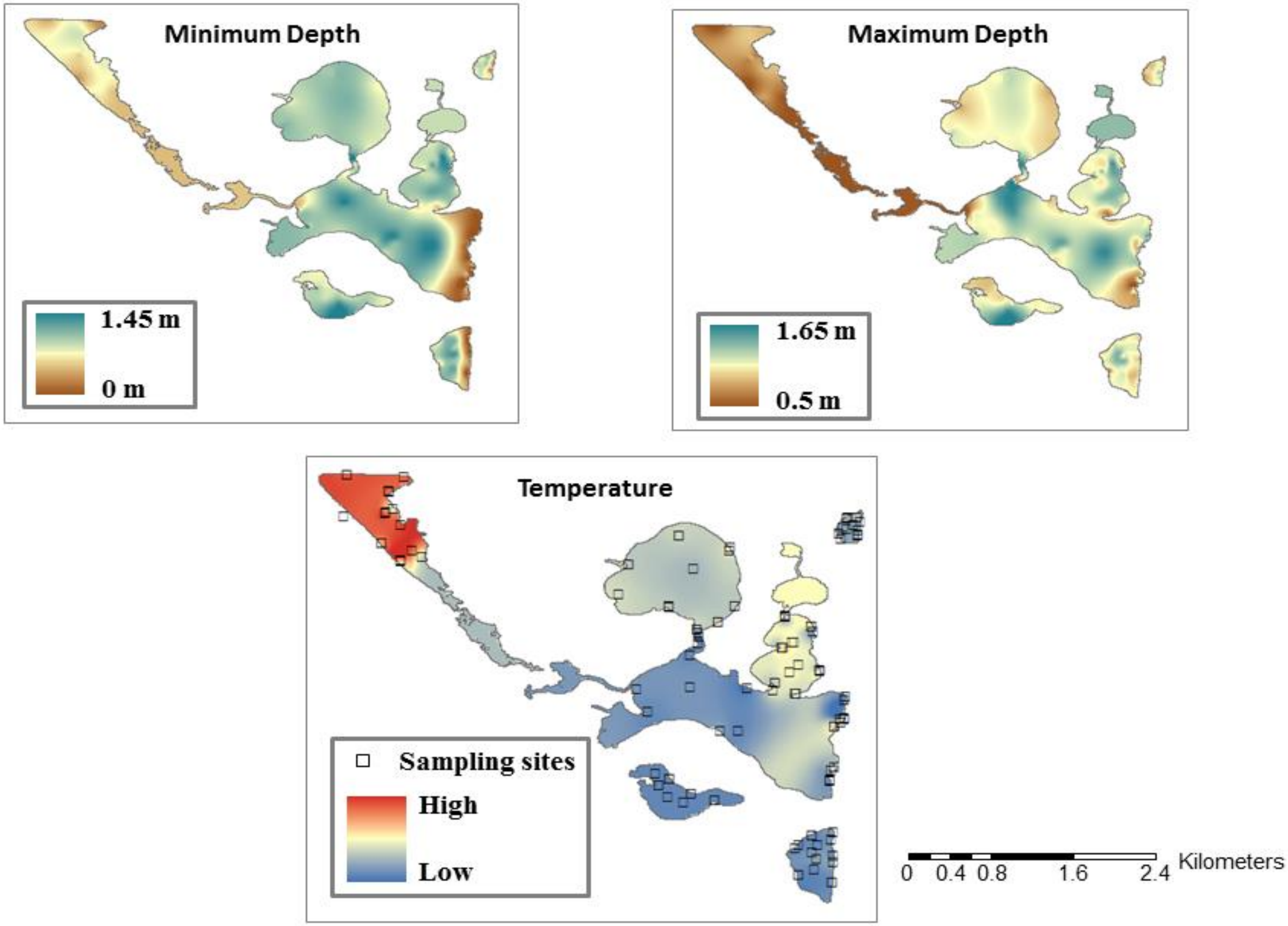


*Figure 4. Interpolated depth and temperature data from 2003 survey in Surette (2006). See Figure 2 for sampling sites related to depth. Temperature data were from the spring of 2003 (May-June survey) (Surette 2006).*

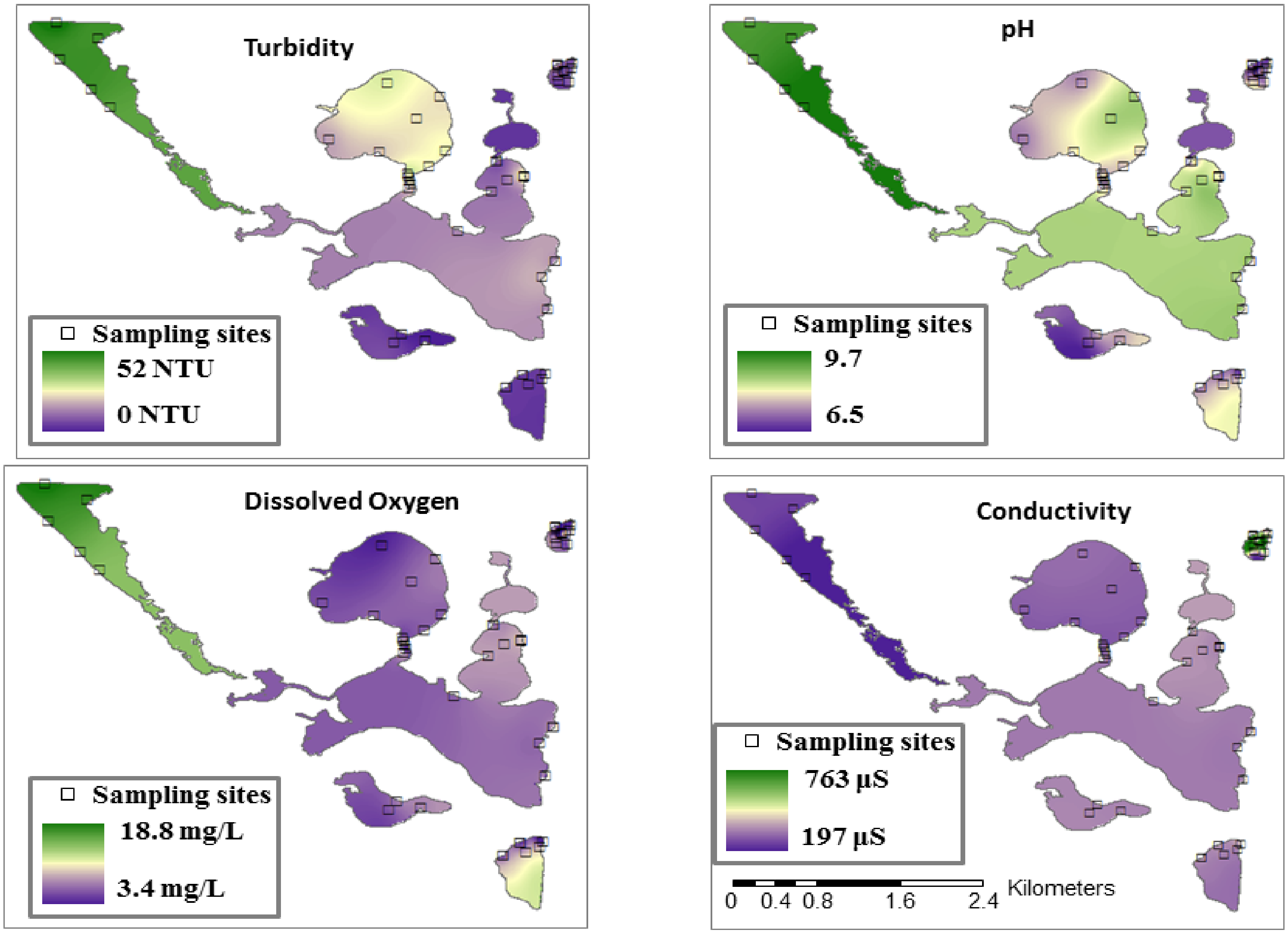


*Figure 5. Interpolated water quality data for 2003 from Surette (2006), with respective sampling sites.*

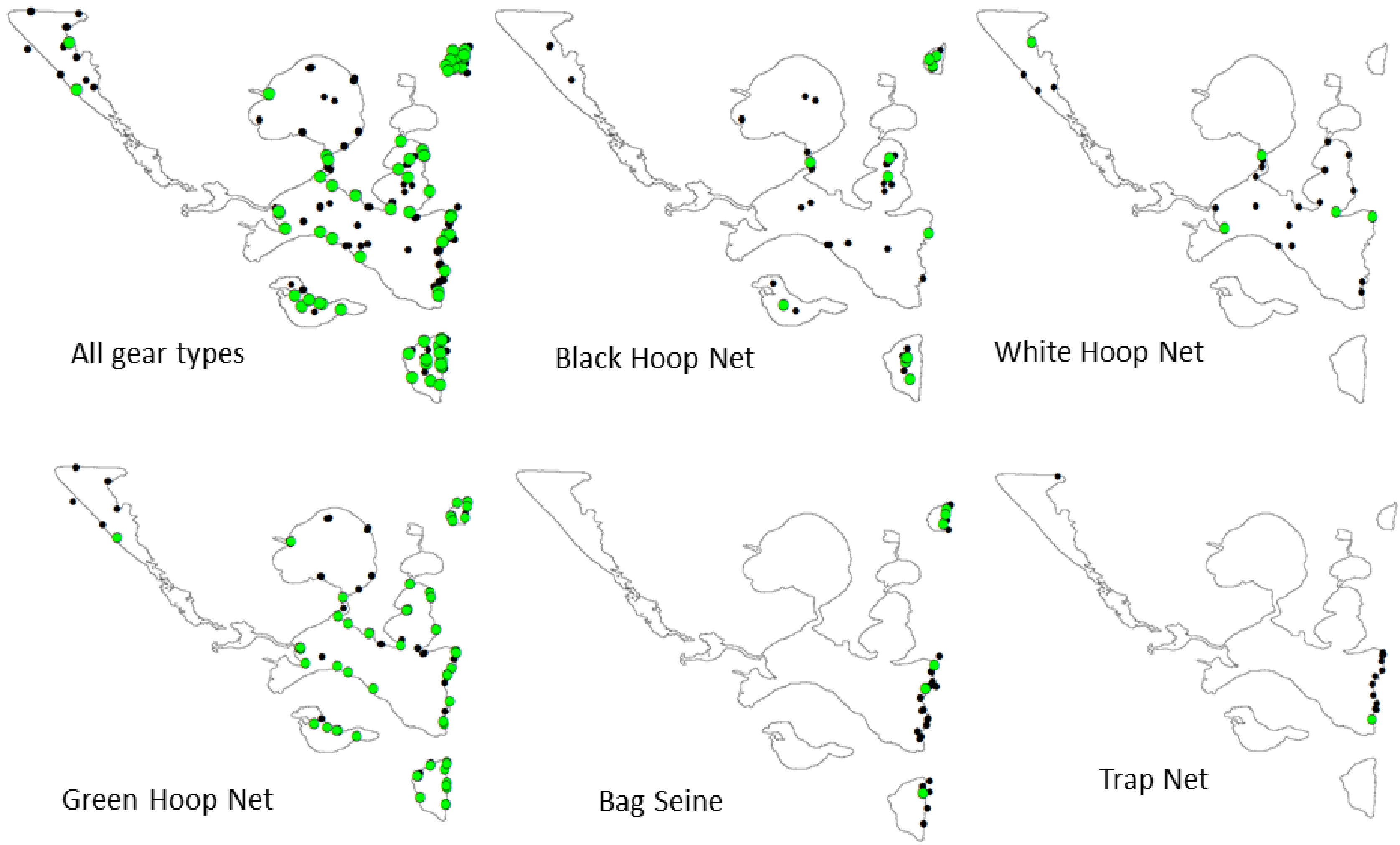


*Figure 6. Spatial extend of sampling conducted in 2003 (Surette 2006), shown together as well as separated by gear type. Black circles represent sampling sites where no Warmouth were captured. Green circles represent sampling sites where Warmouth were captured.*

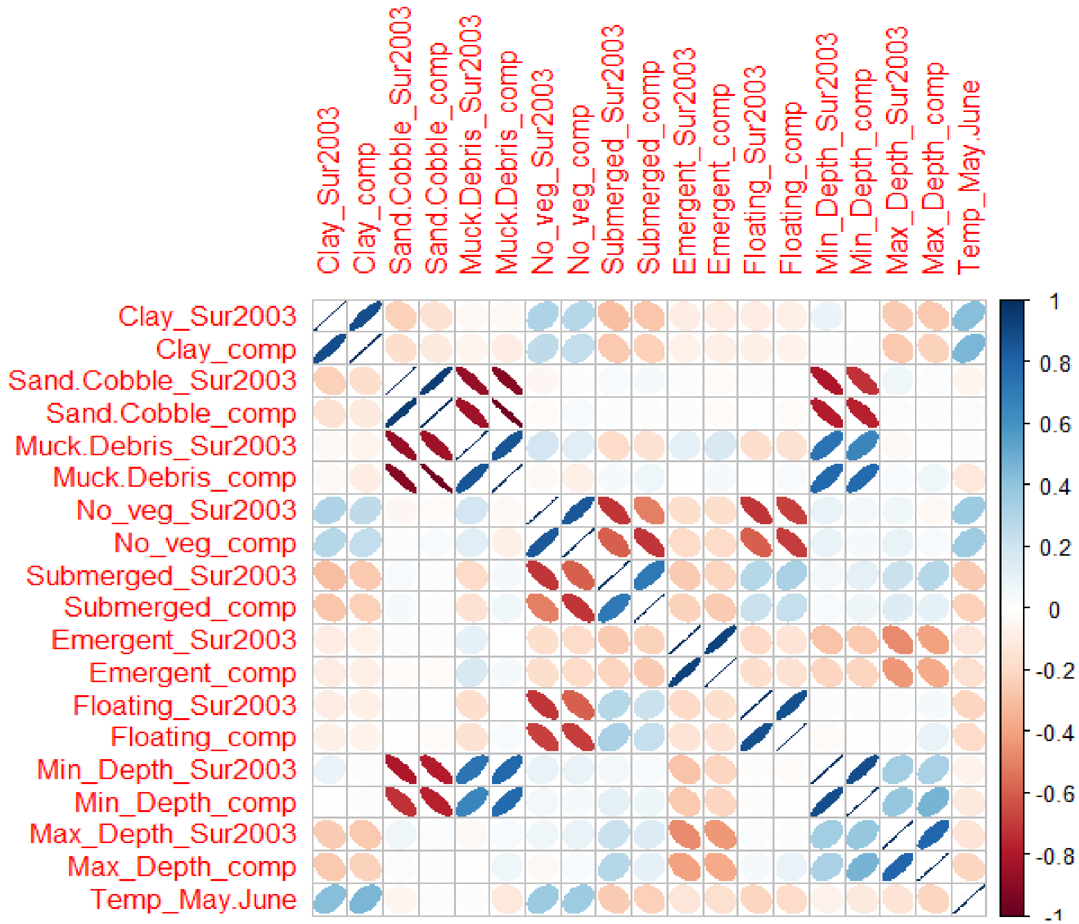


*Figure 7. Correlations among interpolated substrate, vegetation, depth, and temperature data from Surette's 2003 data and a composite layer of all data available (Note: 'no vegetation' in this figure refers to 1- 'all vegetation' from Figure 3). The strength of the correlation is indicated by color in legends shown on the right.*

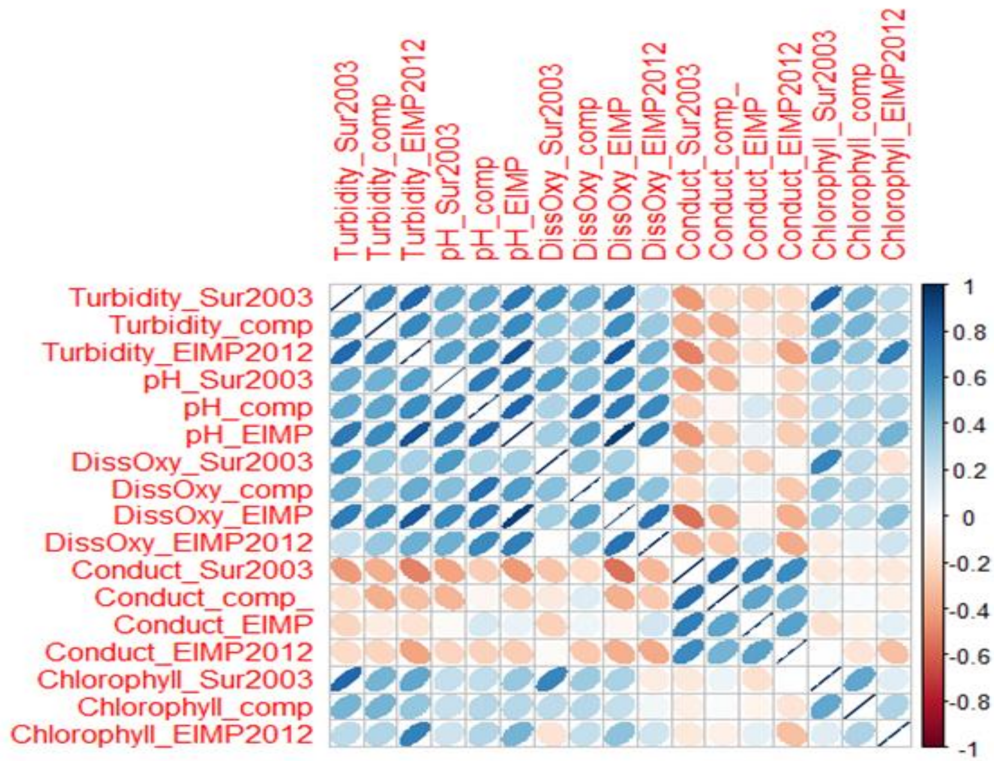


*Figure 8. Correlations among interpolated water quality data layers. 'Sur2003' represents Surette's 2003 data; 'comp' represents a composite layer of all data available, including Surette's 2002 and 2003 data, EIMP data from 2008-2015, and Razavi (2006) 2005 data; 'EIMP' represents all EIMP survey data from 2008-2015; and, 'EIMP2012' represents EIMP survey data from 2012 only, when the spatial coverage was greater than in other EIMP survey years. The strength of the correlation is indicated by color in the legend shown to the right.*

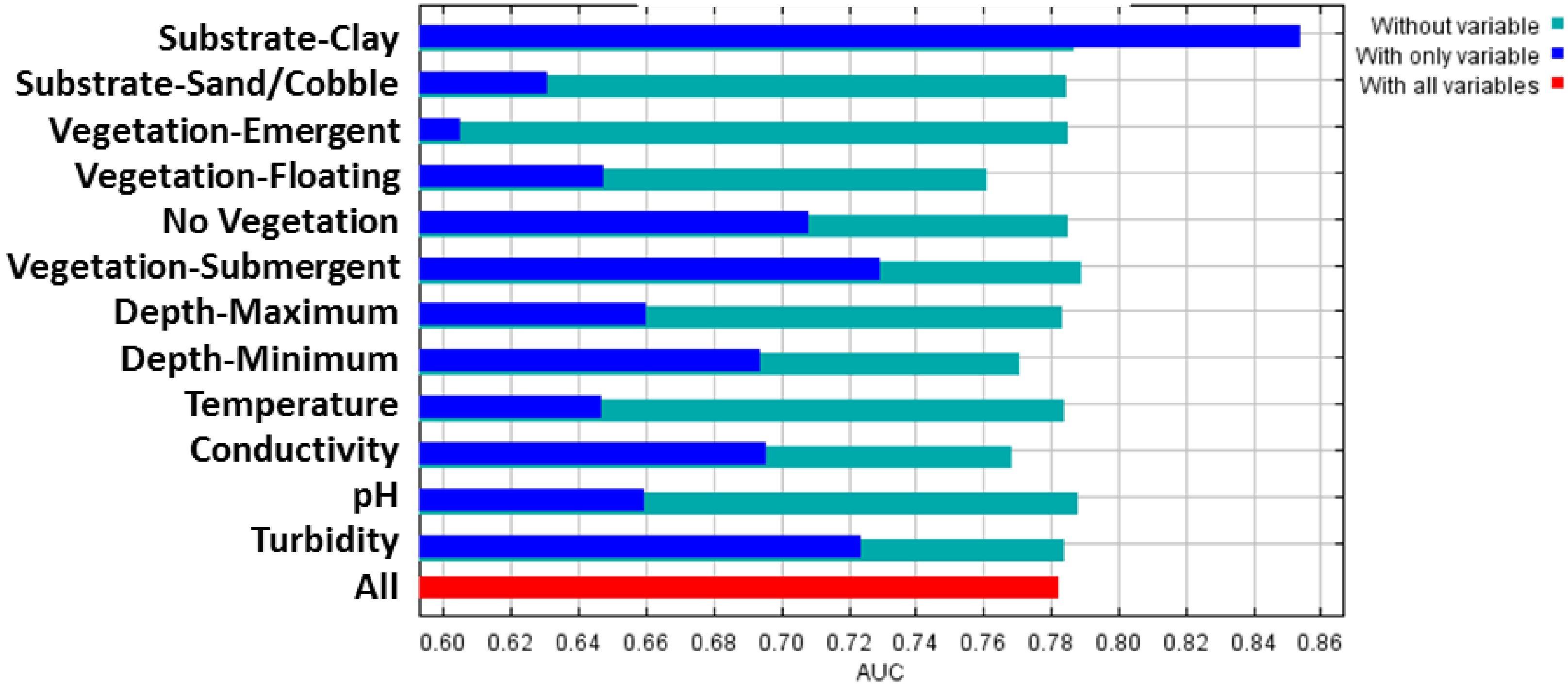


*Figure 9. Test AUC results from species distribution modeling using MaxEnt. Support for the model based on all variables (red bar) is approximately 0.78. Statistic support for models based on single variables are represented with dark blue bars. Statistical support for models based on all variables other than the single variable is represented with light blue bars.*

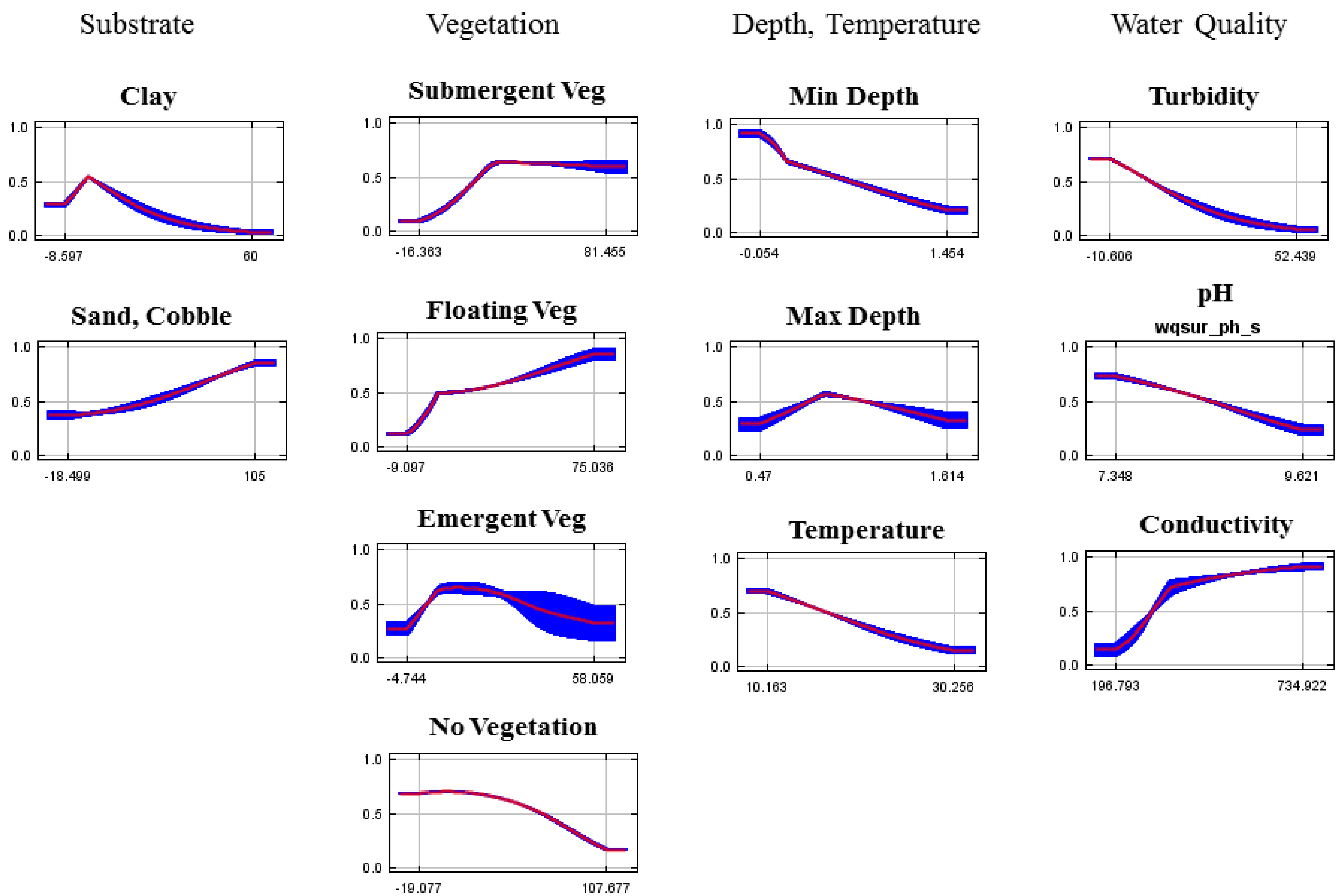


*Figure 10. Results of MaxEnt models created using only a single environmental variable. These graphs show how the logistic prediction of habitat suitability changes as the environmental variable changes. Higher values on the y-axis indicate higher habitat suitability.*

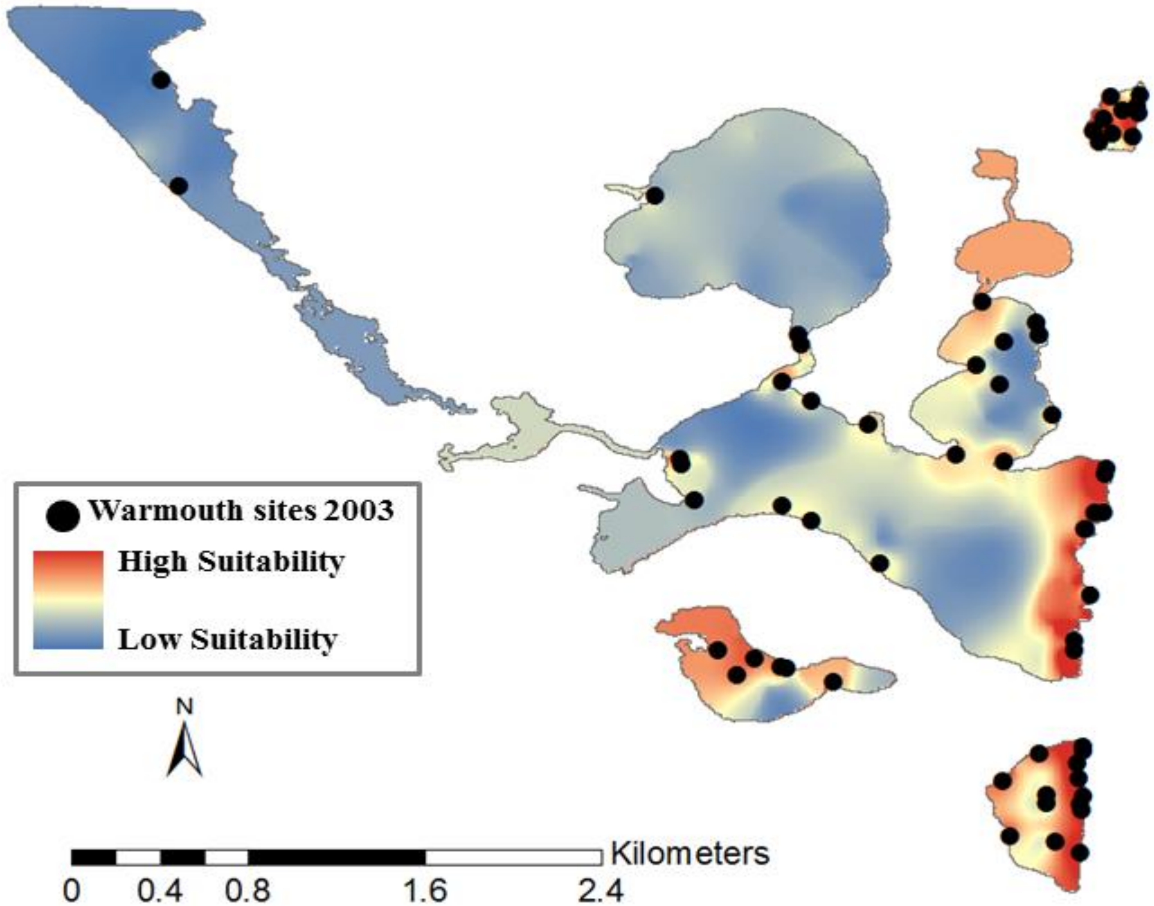


*Figure 11. Species distribution model predications of suitable habitat as indicated by continuous probabilities. Warmouth presence localities from 2003 are also shown.*

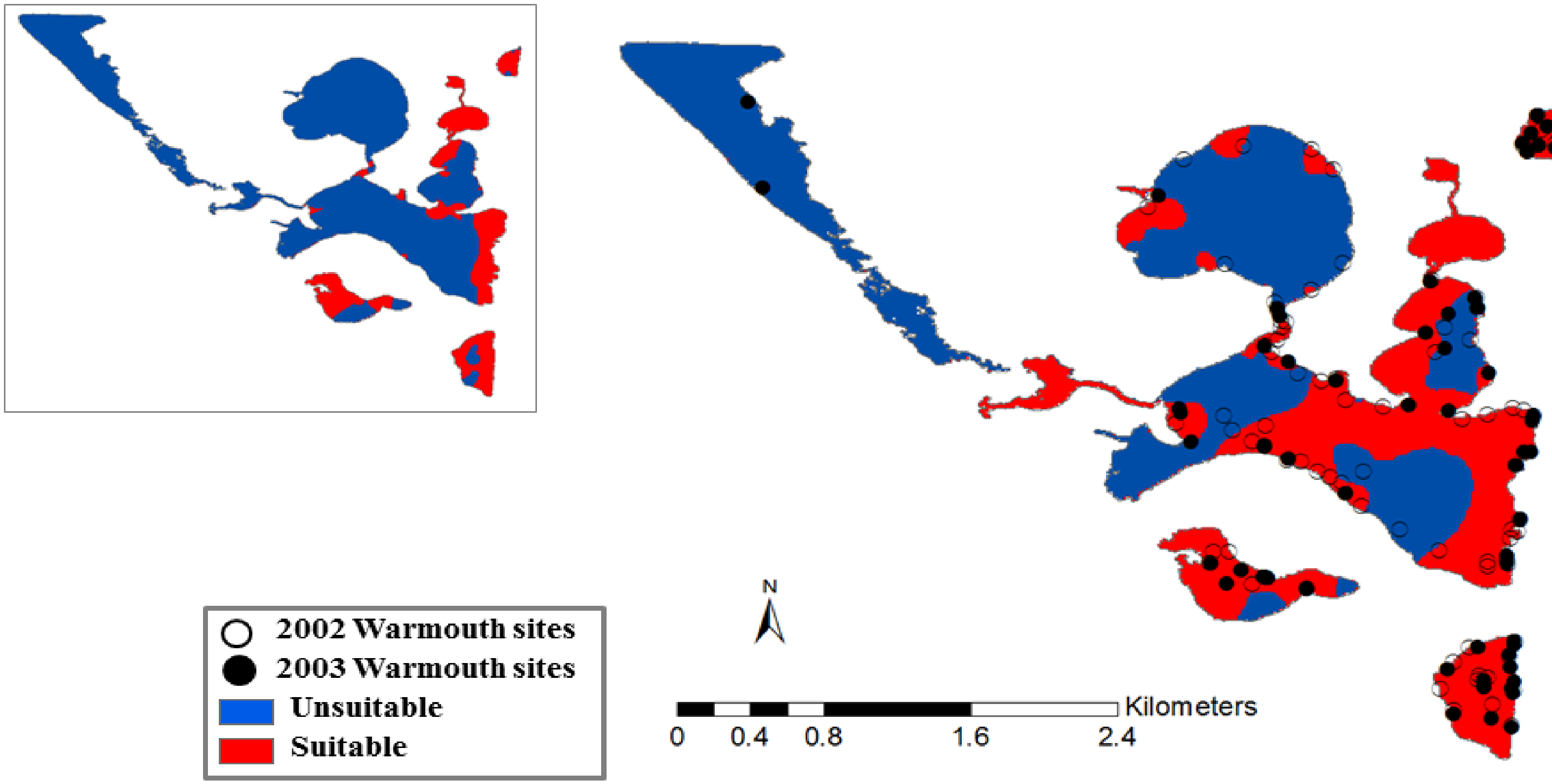


*Figure 12. Suitable and unsuitable habitat based on two thresholds of suitability. Maximum test sensitivity plus specificity threshold (0.38) is shown on the left; and, the 10$^{th}$ percentile training presence threshold (0.23) is shown in the centre. Warmouth presence sites from both 2002 and 2003 were plotted to illustrate how well the predictive region, which is based on 2003 data, captures Warmouth sites from both years.*

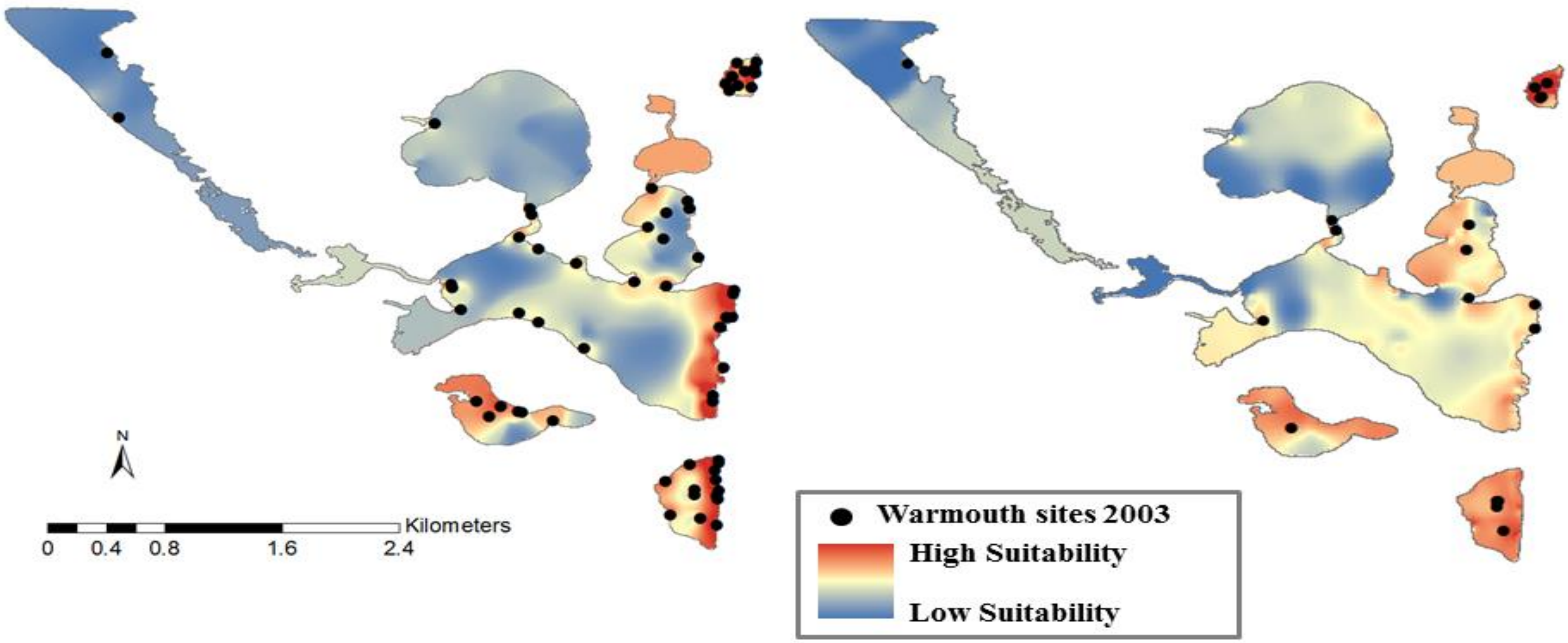


*Figure 13. Comparison of species distribution model predictions of suitable habitat from the original model using all locality data (left) and using only locality data from the black hoop net and white hoop net (Figure 7). All Warmouth presence localities from 2003 are shown on the original prediction (left), but only those localities identified with the black hoop net and white hoop net are shown on the right.*

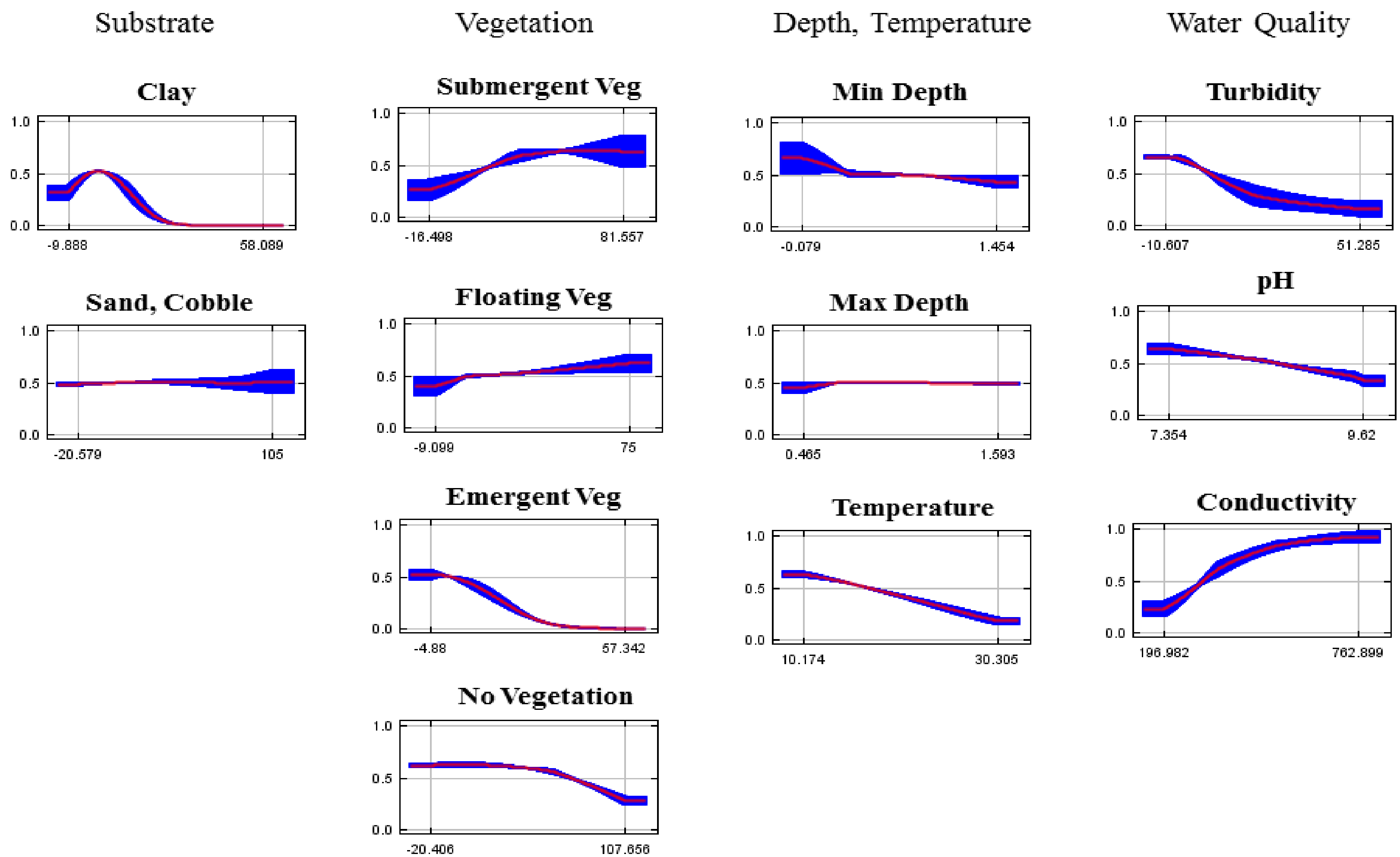


*Figure 14. Results of the second MaxEnt model (using only black and white hoop net data) created using only a single environmental variable. These graphs show how the logistic prediction of habitat suitability changes as the environmental variable changes. Higher values on the y-axis indicate higher habitat suitability.*