

Refining species distribution model outputs using landscape-scale habitat data: Forecasting grass carp and *Hydrilla* establishment in the Great Lakes region

Marion E. Wittmann<sup>a,b,\*</sup>, Gust Annis<sup>c</sup>, Andrew M. Kramer<sup>d</sup>, Lacey Mason<sup>e</sup>, Catherine Riseng<sup>e</sup>, Edward S. Rutherford<sup>f</sup>, William L. Chadderton<sup>c</sup>, Dmitry Beletsky<sup>g</sup>, John M. Drake<sup>d</sup>, David M. Lodge<sup>a,i</sup>

<sup>a</sup>Department of Biological Sciences, University of Notre Dame, Notre Dame IN 46556.

<sup>b</sup>University of California Natural Reserve System, University of California, Santa Barbara, CA, 93106-6150

<sup>c</sup>The Nature Conservancy, 101 East Grand River Lansing, MI 48906

<sup>d</sup>University of Georgia, Athens, GA 30602. Email: [kramera3@uga.edu](mailto:kramera3@uga.edu)

<sup>e</sup>University of Michigan, School of Natural Resources and Environment, 440 Church St, Ann Arbor, MI 48109-1041.

<sup>f</sup>NOAA Great Lakes Environmental Research Laboratory, 4840 S. State Rd. Ann Arbor, MI 48108-9719.

<sup>g</sup>Cooperative Institute for Limnology and Ecosystems Research, School of Natural Resources and Environment, University of Michigan, 4840 S. State Rd, Ann Arbor, MI 48108

<sup>h</sup>Environmental Change Initiative, University of Notre Dame, Notre Dame IN 46556

\*Corresponding author: [marion.wittman@ucsb.edu](mailto:marion.wittman@ucsb.edu); phone 805-893-6179

## **Abstract**

Forecasts of the locations of species invasions can improve by integrating species-specific climate and habitat variables and the effects of other invaders into predictive models of species distribution. We developed two species distribution models (SDMs) using a new algorithm to predict the global distributions of two nonindigenous species, grass carp (*Ctenopharyngodon idella*) and *Hydrilla* (*Hydrilla verticillata*), with special attention to the North American Great Lakes. We restricted the projected suitable habitat for these species using relevant habitat data layers including accumulated Growing Degree Days (GDD), submersed aquatic vegetation (SAV), wetlands, and photic zone. In addition, we restricted the grass carp niche by the projected *Hydrilla* niche to explore the potential spatial extent for grass carp given a joint invasion scenario. SDMs showed that climate conditions in the Great Lakes basin were often suitable for both species, with a high overlap between the areas predicted to be climatologically suitable to both species. Restricting *Hydrilla* regions by GDD and photic zone depth showed that the nearshore zones are primary regions for its establishment. The area of predicted habitat for grass carp increased greatly when including *Hydrilla* niche as a potential habitat for this species. Integrated risk maps can provide a means for the scientifically informed prioritization of management resources toward particular species and geographic regions.

## **Keywords**

nonindigenous species; management of invasions; habitat suitability; environmental niche

## **Introduction**

Biological invasions of freshwater ecosystems cause significant impacts on community structure and ecosystem function (Havel et al., 2015). The effects of freshwater invasive species occur globally, in part because of their ability for widespread dispersal through both natural pathways (e.g., active or passive movement through connected waterways) and human-mediated mechanisms (e.g., intentional stocking, accidental releases, hitchhiking on vessels or equipment). Often these human-mediated vectors are associated with commercial and recreational activities. It is expected that the effects of aquatic invasive species will increase as human populations and associated commerce and recreation also increase (Lockwood et al., 2013).

As a major center of commercial and recreational activity, the Great Lakes Basin represents a unique confluence of nonindigenous aquatic species (NAS) from across the globe (Rothlisberger and Lodge, 2013). Multiple pathways of introduction of NAS to the Great Lakes include transport on ship hulls, contamination of ballast tanks, regional overland movement from inland waterbodies on recreational boats, aquarium and horticulture trade, accidental release, or passage through waterway connections (MacIsaac et al., 2001). As a result, the Great Lakes have been subject to over 180 nonindigenous species establishments, some of which have caused irreversible ecological shifts and significant economic damages (Mills et al., 1993; Ricciardi and MacIsaac, 2000; Rothlisberger et al., 2012). As resources to manage biological invasions are typically scarce, there is value in understanding where NAS may establish prior to their establishment. Further, the ability to understand how different NAS may facilitate future invasions may also provide useful information for managers prioritizing prevention or control strategies.

A number of tools have been developed to forecast invasive species distributions based on the concept of the environmental niche (Guisan and Zimmermann, 2000). Termed "environmental niche model" or

"species distribution model" (hereafter referred to as SDM), these correlative predictive models combine known geographic locations or occurrences of a species with environmental data (often climate data such as temperature or precipitation) to predict species potential geographic distributions in novel environments (Elith and Leathwick, 2009; Hutchinson, 1958; Jiménez-Valverde et al., 2011; Pulliam, 2000). SDMs have been used in a wide range of applications such as locating rare and threatened species and habitats, predicting the spread of invasive species, and estimating the response of species to global climate change (Guisan and Thuiller, 2005). In the Great Lakes, SDMs have been used to anticipate invasions from the Caspian Sea (Fitzpatrick et al., 2013), assess suitable climate conditions for an organism in trade, grass carp (*Ctenopharyngodon idella*) (Wittmann et al., 2014), and to estimate the potential range expansion of common reed (*Phragmites australis*) (Carlson Mazur et al., 2014).

While SDMs are useful tools to estimate species distributions based on climate variables, they have been criticized for the lack of the integration of species-specific habitat information in model specification (Araújo and Peterson, 2012; Elith and Leathwick, 2009). The ability to incorporate specific habitat information is typically not possible because of the absence of relevant data at the appropriate spatial resolutions or extents (Gies et al., 2015). Further, predicting which combinations of species and habitats may facilitate or prevent biological invasions is difficult (Romanuk et al., 2009). In part, this is due to uncertainties in forecasting how nonindigenous species may interact with one another, and with their environments in invaded ecosystems (Grosholz et al., 2000; Johnson et al., 2009).

Here, we forecast biological invasions by combining an SDM algorithm and a recently developed spatially explicit habitat classification database to assess sole and joint invasion scenarios for two nonindigenous species that threaten the Great Lakes ecosystem, grass carp and *Hydrilla* (*Hydrilla verticillata*). First we apply "range bagging", a new technique of species distribution modeling that uses only species presence data (Drake, 2015) to predict suitable climate conditions for these species. Range

bagging draws on the concept of a species' environmental range and is inspired by the empirical performance of ensemble learning algorithms (e.g., boosted regression trees) in other areas of ecological research (Elith et al., 2008). Second, we evaluate localized habitat suitability for each of these species using Great Lakes habitat data layers (Wang et al., 2015) specific to the physiological limitations of these species as found in the published literature. Finally, we investigated the intersection of the restricted niches of both species to understand how the predicted suitable habitat for *Hydrilla* might enhance grass carp distribution.

The goal of this study was to identify potential habitat for two invasive species that currently threaten the Great Lakes region. For our specific study species, *Hydrilla* and grass carp, we tested the following three hypotheses for the Great Lakes region: (1) that there is suitable habitat for the potential invasion of *Hydrilla* and (2) grass carp, and that (3) the presence of *Hydrilla* can increase the amount of habitat in which grass carp could persist. We propose that the incorporation of habitat specific information in SDMs can focus management efforts on the locations where prevention, management, and monitoring programs will be most effective.

## **Methods**

### **Study Species**

A number of nonindigenous aquatic species currently threaten to invade the Great Lakes region (USACE, 2011). Due to their proximity to the Great Lakes and the documented negative impacts in other systems, grass carp (*Ctenopharyngodon idella*) and *Hydrilla* are of particular concern (Langeland, 1996; Michelan et al., 2014; Wittmann et al., 2014). Grass carp is a large cyprinid fish native to eastern Asia, with a native range extending from northern Vietnam to the Amur River along the Russia-China border (Fuller et al., 1999). An herbivore, it has been globally introduced for nuisance aquatic plant control and is also cultivated in China and other countries as a food source. Despite its widespread introduction and use as

a biocontrol agent for nuisance aquatic macrophytes, there has been a great deal of uncertainty about its ecological risk—particularly in the Great Lakes Region (Wittmann et al., 2014). Currently, both diploid and triploid grass carp remain widely available for stocking in the U.S., and feral, reproducing populations have been reported in the Illinois and Mississippi Rivers (Raibley et al., 1995), Lake Erie, and some tributaries of Lake Erie and Lake Michigan (Chapman et al., 2013; Wittmann et al., 2014).

*Hydrilla* is a submersed aquatic macrophyte native to central Asia and Australia (Cook and Lüönd, 1982). *Hydrilla* was first detected in Florida in the 1960s (Steward et al., 1984) and is now considered invasive and naturalized in the United States as well as much of temperate North America. Introduced populations also occur in Central and South America, Africa, Europe, and New Zealand (Langeland, 1996). *Hydrilla* often has unwanted impacts such as impeding water conveyance, impairment of recreation activities, displacement of native plants, and alteration of nearshore community structure (Gordon, 1998; Langeland, 1996). The monoecious form of *Hydrilla* has been found in waterways with a direct connection to the Great Lakes in New York and Ohio as recently as 2012 (Jacono et al., 2014). The Great Lakes has experienced a number of native aquatic macrophyte declines of ecologic and cultural importance, including wild celery (*Vallisneria americana*) and Wild rice (*Zizania palustris*) (Schloesser and Manny, 2007; Sierszen et al., 2012). Concern about sensitive wetland species like these in the Great Lakes continues to increase as *Hydrilla* is discovered in adjacent watersheds because studies have shown that *Hydrilla* can competitively exclude these and other native aquatic macrophytes when they are co-located (Chadwell and Engelhardt, 2008; Langeland, 1996; Rybicki and Carter, 2002). Due in part to its life history, *Hydrilla* is extremely difficult to eradicate (Rejmánek and Pitcairn, 2002). Grass carp prefers *Hydrilla* as a food source, and is commonly used as a biocontrol agent for *Hydrilla* in the southern US, Texas and other regions where the plant is a nuisance (Chilton et al., 2008; Pine and Anderson, 1991; Shireman and Maceina, 1981).

## Species Distribution Model

We predicted regions of suitable climate conditions for grass carp and *Hydrilla* by estimating statistical relationships between a widely-used set of global climate variables and species occurrence records using a new SDM method called “range bagging” (Drake, 2015). Range bagging is a form of boundary estimation, considering the limits of the environmental space where a species can persist. The range bagging algorithm efficiently estimates the range limits in a multi-dimensional space of environmental variables using bootstrap aggregation. By repeatedly defining the convex hull of occupied environments in 2 of  $n$  dimensions at a time it is possible to determine how often a given environment occurs inside these niche boundaries. The resulting measure, called “niche centrality”, refers to the proportion of times an environment occurs within the environmental range of a species across the bootstrapped combinations of environmental variables. Range bagging compares well to other species distribution models in traditional SDM contexts (Drake, 2015; Drake and Richards in review) and for invasive species (Cope et al. in review) with the distinct advantages of not requiring pseudo-absence points for fitting and having an ecologically relevant interpretation (Drake, 2015; Cope et al. in review).

The range bagging models were trained on a random partition of 80% of the occurrence data.

Performance was reported as the area under the ROC (receiver operating characteristic) curve, which is referred to as AUC, or “Area Under Curve” (Hanley and McNeil, 1982) on the remaining test data and an equal-sized set of randomly selected pseudo-absences (e.g., background localities to be used for model parameterization). The pseudo-absence points were taken from a 2000 km buffer around the presence points. Such use of background points is common practice for data that only contain occurrence records (Elith and Leathwick, 2009). As a complementary measure of performance, we estimated the continuous Boyce index (Hirzel et al., 2006; Petitpierre et al., 2012) that is designed for presence-only data using the R package “ecospat”. The Boyce index varies from -1 to 1 with values greater than zero indicating



agreement between the predictions and the presences in the test data (Hirzel et al., 2006). This was calculated alongside the AUC, with both metrics using the same model output and data for each run of the model. We assessed the variance in model performance by performing 10-fold cross-validation on the training data (Table S1). Further, we estimated the transferability of the model by conducting 5-fold cross-validation on data that was divided into longitudinal bins (Wenger and Olden, 2012). This test measures the ability of the model to predict occurrence in distinct geographical areas, with longitudinal bins being appropriate for sampling the occurrence of these species on multiple continents.

The niche centrality for each species was estimated as the fraction of times climate conditions were captured within the range of the observed occurrences. We also provided estimates of variable importance for each species. The importance of candidate predictor variables was measured by permuting each variable in sequence and calculating the average reduction in accuracy (as measured by changes in AUC, see Electronic Supplementary Material (ESM) Figures S1, S4), across 500 permutations. The marginal effect of each variable on model performance was then plotted by varying the predictor of interest while holding other variables at their median value (See ESM Figures S3, S6).

### **Grass carp and *Hydrilla* occurrence records and environmental climate variables**

Global grass carp and *Hydrilla* locality information (i.e., positive occurrence) used in the range bagging model was obtained from the primary literature and published databases accessed in December 2014 from the Global Biodiversity Information Facility (GBIF; gbif.org), Fishbase (fishbase.org), and United States Geological Survey (USGS; usgs.nas.gov) databases. Overall, 663 grass carp and 1017 *Hydrilla* occurrence records were collected and these were globally distributed on all continents except Antarctica, South America and Australia (grass carp only) (Figures 1, 3). grass carp and *Hydrilla* occurrence records were collected between 1934 – 2014 and 1953 – 2014, respectively. Each

georeferenced position was verified, and all localities with an uncertainty of position (e.g., a place described as “China” rather than a particular river reach or other waterbody location) larger than 50 km were removed to improve the accuracy of the model predictions.

The environmental climate variable dataset used in the range bagging model was comprised of 19 environmental climate variables (Hijmans et al., 2005; worldclim.org). These climate variables, in turn, are derived from global maps of temperature ( $n=11$ ) and precipitation ( $n=8$ ) interpolated from observed data (representative of 1950-2000) (see Hijmans et al., 2005 for a full description of the data used to compile the climate variable data) and have been used to previously represent current climate in species distribution modelling due to fine spatial resolution and global coverage (Hijmans et al., 2005).

Specifically, the 11 climate variables that concern temperature include annual mean temperature, mean diurnal range, isothermality, temperature seasonality, maximum temperature of the warmest month, minimum temperature of the coldest month, temperature annual range, mean temperature of the wettest quarter, mean temperature of the driest quarter, mean temperature of the warmest quarter, and mean temperature of the coldest quarter. The 8 climate variables that concern precipitation are annual precipitation, precipitation of the wettest month, precipitation of the driest month, precipitation seasonality, precipitation of the wettest quarter, driest quarter, warmest quarter, and coldest quarter. The data used was the standard download available from WorldClim, specifically we used the R package ‘dismo’ to access the data.

WorldClim and similar climate data have been shown to be relevant to construction of abiotic niche space (Araújo and Peterson, 2012). In past studies, both large-scale temperature and precipitation climate variables (e.g., WorldClim data) have been used as proxies for local aquatic variables and to model stream fish distributions and several studies suggest that models built using macroscale variables

perform similar to models based on local variables for aquatic species distributions (Frederico et al., 2014; Porter et al., 2000; Watson and Hillman, 1997). Population establishment and growth rate of grass carp (and other aquatic and terrestrial species) have been shown to change with habitat suitability measures derived from environmental niche models developed this collection of climate variables (Wittmann et al., 2016). We also included the full suite of variables to develop habitat suitability predictions for *Hydrilla*, noting that average annual temperature, average monthly temperature, diurnal temperature regime, and water availability have previously been used to model *Hydrilla* fitness, growth and establishment (Gu, 2006; Langeland, 1996; Rybicki and Carter, 2002; Spencer et al., 2000; Sutton, 1996; Zhang et al., 2013). Accordingly, it seems plausible that the full set of variables would be appropriate for consideration in this application given the potential effects of both temperature and precipitation on littoral zone conditions (e.g., where grass carp and *Hydrilla* will occur) and other hydrologic regimes (flow, turbidity, water temperature) in the Great Lakes ecosystem. See ESM Appendix S1 for a detailed description of each environmental climate variable, variable rescaling methodologies, and steps taken to reduce bias in model fitting.

### **Habitat-specific evaluation**

We used Great Lakes habitat data layers taken from the Great Lakes Aquatic Habitat Framework (GLAHF; <http://glahf.org/>; Wang et al., 2015) benthic growing degree days, wetland habitat) and the Michigan Tech Research Institute (submersed aquatic vegetation, photic zone) to create spatial data layers which were used to restrict the species distribution model outputs to suitable habitats in the Great Lakes region after converting all layers to a common projection and cell size. Species-specific information on environmental limits to occurrence, survival, and spawning was collected from extensive review of the primary literature. This included published limits of measured environmental variables, such as temperature, growing degree days, substrate type preference, and water chemistry. Cases for

which reliable Great Lakes-wide data were unavailable were discarded (e.g., primarily substrate type or water quality data). When literature sources differed in their reported limits, the extrema from the set of values were considered to be the limit. These environmental conditions were then used to restrict the projected range bagging model to the suitable areas of the Great Lakes for each species.

To identify local (non-climatological) restriction to *Hydrilla* and grass carp distribution we used several variables to identify Great Lakes regions likely to provide suitable habitat for these species based on location specific information. To restrict the *Hydrilla* model, we used a measure of heat accumulation, or the number of accumulated growing degree days (GDD;  $n \geq 500$ ), to determine regions of the Great Lakes that would allow a 50% sprouting rate of axillary turions formed by monoecious *Hydrilla* plants (after Spencer et al., 2000). We also restricted the *Hydrilla* model by depth of the euphotic zone to represent the limitation of water transparency and the maximum depth of *Hydrilla* colonization within a lake (Canfield and Langeland, 1985).

To restrict the grass carp model, we used a combined data layer representative of submersed aquatic vegetation (SAV) and wetland regions of the Great Lakes. Grass carp reproduction is largely based on processes that occur in riverine habitats, namely water velocity, flow regime and temperature (Shireman and Smith, 1983; Stanley et al., 1978). Adjacent tributaries to the Great Lakes appear to have suitable thermal and hydrologic conditions for successful recruitment and habitation (Chapman et al., 2013; Kocovsky et al., 2012). Thus, we sought to delineate the distribution of this species by SAV and coastal wetland layers in part because of its likelihood for direct utilization of these habitats as a food resource and refuge. Through literature review, we found that the thermal conditions in the Great Lakes were suitable for grass carp physiology and could not designate any other Great Lakes specific habitat variables by which to restrict grass carp habitat by. We also created a second assessment of the grass

carp niche by restricting the Great Lakes region projection to a combined SAV, wetland, and projected *Hydrilla* (including GDD and photic zone restriction) data layer.

### **Great Lakes Habitat Variables**

Growing degree day (GDD) data were calculated using benthic temperature data (available in the GLAHF database) derived from the NOAA Great Lakes Coastal Forecasting System's (GLCFS) vertical temperature models (NOAA-GLERL, 2013). The GLCFS employs a 3D hydrodynamic model (Beletsky et al., 2013; Chu et al., 2011; Schwab and Bedford, 1994) with a horizontal resolution ranging from 2 km (Lakes Erie, Huron and Michigan) to 5km (Lake Ontario) to 10 km (Lake Superior) to determine lake temperatures (among other physical variables) at 20 vertical levels in all lakes except Lake Erie, which has 21 levels. Benthic temperature was derived using the bottom layer of the hydrodynamic model. Benthic GDD was computed using a lower bound of 8°C based on requirements of *Hydrilla* (Barko and Smart, 1986; Spencer et al., 2000) and was calculated separately for each of the Great Lakes using the formula,  $GDD = T_{avg} - T_{base}$ , when  $T_{avg} \geq T_{base}$ . Data for all five lakes (with the exception of Lake St. Clair for which data were not available) were combined by year using a mosaic process in ArcGIS Version 10.2 with an output cell size of 2000 m. We then computed an average value for each grid cell using the years 2006-2012.

Great Lakes euphotic zone depth ( $z_{1\%}$ ), or the depth where only 1% of the surface photosynthetic available radiation (PAR) remains, was calculated directly from the diffuse attenuation coefficient for downwelling irradiance at 490 nm ( $Kd_{490}$ ), in  $m^{-1}$  (NASA, 2015). This algorithm was evaluated using an empirical relationship derived from in situ measurements of  $Kd_{490}$  and blue-to-green band ratios of remote sensing reflectances near 490 nm and between 547 and 565 nm. PAR data have been validated through comparisons between two satellite systems (MODIS and SeaWiFS PAR) and *in situ* PAR at three

Great Lakes locations (Chatham in Lake Superior, Muskegon in Lake Michigan, and Gaylord in Lake Huron) (Yousef et al., in review).

Wetland data were compiled by the Great Lakes Coastal Wetland Inventory (GLCWC, 2004). This inventory utilized the most comprehensive coastal wetlands data available for the Great Lakes and connecting channels and was derived from multiple sources. Additional information about these data can be found at <http://glc.org/projects/habitat/coastal-wetlands/cwc-inventory/>. Coastal wetlands polygonal data were rasterized for use with a cell size of 30 m.

Submerged aquatic vegetation (SAV) data were produced and provided by the Michigan Tech Research Institute (Brooks et al., 2015; MTRI, 2012; Shuchman et al., 2013). The data have a 30 m resolution and represent the extent of SAV in the optically shallow areas of lakes Huron, Michigan, Erie, and Ontario. The data were generated using an MTRI-developed, depth-invariant algorithm applied to Landsat satellite data. The satellite data were collected during the vegetative growing season during various year ranges. Specifically, vegetative growing seasons were determined by temperature and years varied by lake; Lake Erie, May-September 2006-2011; Lake Huron, March-September 2007-2011; Lake Michigan, April-May 2008-2011; Lake Ontario, April-September 2008-2011. Some portions of these lakes could not be classified due to high turbidity (Shuchman et al., 2013). SAV data did not exist or were not available for Southern Green Bay and Lake Superior and thus were excluded from the SAV restriction for grass carp. Data for the remaining four lakes were combined using the "Mosaic to New Raster" tool in ArcToolbox in ESRI ArcGIS Desktop Version 10.2 (ESRI, 2014) with a cell size of 30 meters. We utilized classes 1 (light submerged aquatic vegetation) and 7 (dense submerged aquatic vegetation). More information about the SAV data can be found at <http://mtri.org/cladophora.html>.

Each of these Great Lakes habitat layers was first re-projected and resampled to match the projection and cell size of the model, i.e., the geographic coordinate system (WGS 1984) with a cell size of

0.000833 x 0.000833 degrees (approximately 65 x 93 m). Model restrictions were performed using the habitat layers as a 'mask' within ArcGIS's Raster ToolBox to limit the model to the spatial areas defined by the physiological parameters as determined through the literature review described above.

## Results

*Hydrilla* ( $n=1017$ ) has been observed on all continents except Antarctica; however most observations occurred outside of its native range (i.e., the Indian subcontinent and regions of Korea; (Jacono et al., 2014) (Figure 1a). Niche centrality of *Hydrilla* was generally  $>0.5$  worldwide, indicating that its environmental tolerances are wide enough that most of the global land mass is contained within its estimated climate niche. Most suitable habitats for *Hydrilla* were identified in North America south of Canada, South America, Europe, Southern Africa, Australia, and in eastern Asia (Figure 1b). The model demonstrated considerable predictive accuracy (AUC = 0.922, ESM Figure S1, Boyce index = 0.785), and mean diurnal temperature range and precipitation of the warmest quarter were the two most important predictor variables (Table 1, ESM Figure S2). Results from the random cross-validation and spatial cross-validation support the predictive value of this model (ESM Table S1).

Overall, niche centrality, or suitable climate condition that falls inside the ecological niche, for *Hydrilla* in the Great Lakes was generally high ( $>0.8$ ) for 86% of the basin (Figure 2a). Clipping the niche centrality for the Great Lakes by GDD indicates that most of the shallow habitats of Lakes Michigan, Huron and Ontario may provide suitable habitat for axillary turion growth of this species. When restricted only by GDD, almost the entirety of Lake Erie has the potential for *Hydrilla* establishment, and in contrast, Lake Superior generally cannot support axillary turion growth of *Hydrilla* due to its benthic thermal profile (Figure 2b). However, incorporating photic zone with the GDD restriction decreased the amount of suitable *Hydrilla* habitat by 62% relative to the GDD clip, and *Hydrilla* is predicted to occur only along the southerly shorelines of Michigan and Huron, and in western and central Lake Erie (Figure 2c).

Similar to *Hydrilla*, grass carp occurrences ( $n=663$ ) were distributed on six continents (except Antarctica), and most observations occurred outside of the native range (i.e., Eastern Asia) (Figure 3a). There was considerable overlap in predicted suitable climate for grass carp and *Hydrilla*, with high niche centrality for grass carp occurring in most regions of North and South America, Europe, Southern Africa, and Australia (Figure 3b), although to a lesser extent than *Hydrilla* (Figure 1b). The model demonstrated considerable predictive accuracy (AUC = 0.877, ESM Figure S4, Boyce index = 0.79). The mean temperatures of the wettest, driest, and warmest quarters and the precipitation of the driest quarter were the four most important predictor variables (Table 1, ESM Figure S5). Results from the random cross-validation and spatial cross-validation show similar predictive performance as for *Hydrilla* (ESM Table S1).

Most of the Great Lakes region had a high niche centrality ( $>0.8$ ), indicating suitable climate conditions for grass carp (Figure 4a). However, when restricted by SAV and wetland layers, suitable habitat was reduced by 98% (Figure 4b). The inclusion of predicted *Hydrilla* niche greatly increased (+633%) the amount of estimated available suitable habitat for grass carp, relative to the SAV and wetland layer, in the Great Lakes (Figure 4c).

## Discussion

Our study predicted regions of suitable climate and habitat for two NAS that currently threaten the Great Lakes. We also estimated the spatial extent of a potential joint invasion of these species. Species distribution model outputs predicted suitable climate conditions within the ecological niche for both species on all continents (excluding Antarctica). Further, in agreement with previously published SDMs for grass carp and *Hydrilla* (Barnes et al., 2014; DeVaney et al., 2009; Herborg et al., 2007; Wittmann et al., 2014), most regions of the Great Lakes were estimated to contain highly suitable climate conditions, with significant spatial overlap of these conditions in all five lakes. By restricting SDM projections with



Great Lakes specific habitat layers, we were able to refine the spatial distribution of environments that pose the highest establishment likelihood given each species' physiological limitations. By combining these predictions with information on the potential for introduction, likelihood of impacts, and secondary spread, natural resource managers may better identify potential hotspots for NAS establishment, develop strategies to prevent new introductions, and prioritize sites for surveillance, containment, or control (Vander Zanden and Olden, 2008).

While the range bagging SDMs predicted that most of the Great Lakes contain suitable climate conditions to support our study organisms, the restricted models indicated that *Hydrilla* and grass carp habitats would occur primarily in the southerly nearshore zones of lakes Michigan and Huron and in the western and central basins of Lake Erie. That these predicted regions occurred generally in the littoral zone was largely driven by photic zone depth and temperature (e.g., GDD) limitations of *Hydrilla* and other SAV. Subsequently, these SAV limits determined the potential suitable habitat for grass carp, which we discuss the details of this interaction in greater detail below. These limiting factors are dynamic variables, often owing to anthropogenic impacts such as climate change, land-use change, and the establishment of other invasive species (Gronewold et al., 2013; Trumpickas et al., 2009; Vanderploeg et al., 2010; Wiley et al., 2010). As these drivers continue to change, and the availability of effective mitigation options may also develop, there could be expansion or contraction of the nearshore ranges of these species. The dynamic nature of environmental change--including novel species interactions--emphasizes the need for iterative risk assessment and the value of long-term datasets to inform and update modeling efforts such as those presented here.

How nonindigenous species may interact with one another in invaded ecosystems has been difficult to predict (Grosholz et al., 2000). This difficulty can be attributed to the variable and hard-to-predict outcomes of species interactions (Cope and Winterbourn, 2004; Johnson et al., 2009; Michelan et al.,

2014). We explored the possibility of the establishment of *Hydrilla* influencing the distribution of grass carp. Relative to model predictions without *Hydrilla*, grass carp range increased due to the assumed presence of *Hydrilla*. While *Hydrilla* is known to be a preferred food source for grass carp (Shireman and Smith, 1983), the specific interaction between these two species, should they become established within the lakes, is uncertain and highly dependent on the unique characteristics of the ecosystems in which they may co-exist. For example, the ecosystem effects when grass carp occurs in the same environment as *Hydrilla* vary based on the population densities of both species, community composition, and ecosystem type and size (e.g., small ponds versus large reservoir systems) (Bain, 1993; Chilton et al., 2008; Dibble and Kovalenko, 2009; Noble et al., 1986).

Our modeling effort may have over- or underestimated potential distributions of *Hydrilla* and grass carp in Great Lakes habitats. Uncertainties or absence of available habitat data at appropriate spatial scales may have impacted the amount of predicted grass carp or *Hydrilla* habitat. For example, limitations associated with the satellite-derived SAV data, such as maximum optical depth, which ranges from 7 m for Lake Erie to > 20 m for Lake Michigan (no data were available for Lake Superior), might cause underestimation of the true distribution of vegetation in the basin (Ashraf et al., 2010). Additionally, nearshore areas known to have extensive macrophyte beds include southern Green Bay (Albert and Minc, 2004); however to our knowledge, no spatial data layers for SAV were available for this area. The absence of these SAV data suggest an underestimation of potential grass carp habitat in this, and possibly other areas.

We used the range bagging methodology in this study for four reasons. First, because range bagging estimates the environmental limits of species habitat, it more closely matches the concept of the ecological niche (Drake, 2015). As a result, its consideration of boundaries rather than the central tendency may offer a more conservative (i.e., broader) niche relevant for invasive species risk

assessment, while also assisting with interpretability. Second, range bagging uses only presence points for model fitting, removing the need for selecting a suitable area from which to sample background points. Third, range bagging is shown to perform as well as widely used alternatives, such as MaxEnt and boosted regression trees on validation data sets and aquatic invasive species, including the two considered here (Drake, 2015; Elith et al., 2011; Kramer et al., *in prep*; Cope et al. *in review*; Drake and Richards *in review*). To support this, we showed that range bagging provided informative predictions even when conducted on spatially-segregated partitions of the data, indicating transferability (Wenger and Olden, 2012). Fourth, range bagging is computationally feasible even when the number of environmental dimensions is large (Drake, 2015).

The implementation of habitat data layers in conjunction with range bagging, or any SDM algorithm, is not always straightforward, especially as environmental tolerances, habitat requirements and species interactions are usually poorly documented (Kilroy et al., 2008). Observations of spatial or physical limitations, such as maximum or minimum depth distribution of a species within a water column are not necessarily representative of the limitation for that species' distribution. This type of information may be misleading without a fuller understanding of what mechanisms related to depth truly limit species establishment. As such, we did not restrict *Hydrilla* distribution by observations of maximum depth from the literature, because the information relating depth to survival in the field was not available. However, we believe it appropriate to utilize the known relationship between light extinction and *Hydrilla* survival because it captures the mechanism by which depth may be a limiting factor. Future efforts could include using correlative studies such as those presented in Gallardo and Aldridge (2013), to better understand the relationship between habitat and species occurrence.

One way to potentially further improve upon range bagging is through the direct integration of experimental physiological data, rather than the model restrictions we utilized in this analyses. Incorporating a prior probability on the minimum or maximum conditions for survival within the range bagging algorithm would be particularly useful in cases where a species is suspected of not reaching the extents of its environmental tolerances for contingent biogeographic or ecological reasons (e.g. an invasive species not at equilibrium with its expected range). Further, the synergistic effects of these complex interactions are likely key variables in the prediction of sustained populations. The understanding of these relationships may be further complicated by potential temporal mismatches between species occurrence records and climatological data used as environmental input variables for SDMs. Herein, for example, the climatological data were collected from 1950 – 2000, and the habitat-specific data were collected after 2005. Arguably, there may be significant differences between the climatological data after 2005, thus affecting the relationship between SDM outputs for grass carp and *Hydrilla*, and the habitat-specific data used to restrict the predicted ranges.

### **Management Implications and Conclusions**

The methods developed in this study may provide a strategy for deriving a scientifically-informed prioritization of Great Lakes regions for the management of existing and future invasions of NAS. The integrated approach presented here relies upon the development and availability of high resolution and broadly scaled habitat data layers that can be applied to assess the ecological risk of NAS or the potential habitat for native species. By combining species-relevant habitat layers with SDM predictions, we sought to develop a useful scientific result for managers who may not only have interest in *Hydrilla* and grass carp establishment in the Great Lakes, but also have an interest in analyses that can be used to form the basis for surveillance or control programs of other NAS. Combining habitat data layers with SDMs not only reduces the uncertainty about where NAS may establish, but these types of multi-species

analyses also serve to identify vulnerable regions of the Great Lakes. Important next steps can include: (1) utilizing habitat layers and known species occurrences to determine and quantify the relationships between them, and (2) taking advantage of existing information (e.g., physiological limitations, occurrences, predictive models based on climate matching) on NAS, particularly those established in adjacent watersheds, to explore the potential consequences of not just one, but multiple invasions.

## **Acknowledgments**

This research was funded by Environmental Protection Agency Great Lakes Restoration Initiative and NOAA Center for Sponsored Coastal Ocean Research: NA09NOS4780192, NA10NOS4780218. We thank M. Barnes for *Hydrilla* data. We would also like to acknowledge the Great Lakes Aquatic Habitat Framework (GLAHF) and the Michigan Tech Research Institute (MTRI). Color producing agent (CPA) retrieval images were produced by the Michigan Tech Research Institute (MTRI) and funded by the Great Lakes Observing System (GLOS) and NASA. This is NOAA GLERL contribution number 1830.

## References

- Albert, D.A., Minc, L.D., 2004. Plants as regional indicators of Great Lakes coastal wetland health. *Aquat. Ecosyst. Health Manag.* 7, 233–247.
- Araújo, M.B., Peterson, A.T., 2012. Uses and misuses of bioclimatic envelope modeling. *Ecology* 93, 1527–1539.
- Ashraf, S., Brabyn, L., Hicks, B.J., Collier, K., 2010. Satellite remote sensing for mapping vegetation in New Zealand freshwater environments: A review. *N. Z. Geog.* 66, 33–43.
- Bain, M., 1993. Assessing Impacts of Introduced Aquatic Species: Grass Carp in Large Systems. *Environ. Manage.* 17, 211–224.
- Barnes, M.A., Jerde, C.L., Wittmann, M.E., Chadderton, W.L., Ding, J., Zhang, J., Purcell, M., Budhathoki, M., Lodge, D.M., 2014. Geographic selection bias of occurrence data influences transferability of invasive *Hydrilla verticillata* distribution models. *Ecol. Evol.* 4, 2584–2593.
- Beletsky, D., Hawley, N., Rao, Y.R., 2013. Modeling summer circulation and thermal structure of Lake Erie. *J. Geophys. Res. Ocean.* 118, 6238–6252.
- Brooks, C., Grimm, A., Shuchman, R., Sayers, M., Jessee, N., 2015. A satellite-based multi-temporal assessment of the extent of nuisance *Cladophora* and related submerged aquatic vegetation for the Laurentian Great Lakes. *Remote Sens. Environ.* 58–71.
- Canfield, D., Langeland, K., 1985. Relations between water transparency and maximum depth of macrophyte colonization in lakes. *J. Aquat. Plant Manag.* 23, 25–28.
- Carlson Mazur, M.L., Kowalski, K.P., Galbraith, D., 2014. Assessment of suitable habitat for *Phragmites australis* (common reed) in the Great Lakes coastal zone. *Aquat. Invasions* 9, 1–19.
- Chadwell, T.B., Engelhardt, K.A.M., 2008. Effects of pre-existing submersed vegetation and propagule pressure on the invasion success of *Hydrilla verticillata*. *J. Appl. Ecol.* 45, 515–523.
- Chapman, D.C., Davis, J.J., Jenkins, J.A., Kocovsky, P.M., Miner, J.G., Farver, J., Jackson, R.P., 2013. First evidence of grass carp recruitment in the Great Lakes Basin. *J. Great Lakes Res.* 39, 547–554.
- Chilton, E., Webb, M., Ott, R., 2008. Hydrilla Management in Lake Conroe, Texas: A Case History. *Am. Fish. Soc. Symp.* 62, 247–257.
- Chu, P.Y., Kelley, J.G.W., Mott, G.V., Zhang, A.J., Lang, G.A., 2011. Development, implementation, and skill assessment of the NOAA/NOS Great Lakes Operational Forecast System. *Ocean Dyn.* 61, 1305–1316.
- Cook, C.D.K., Lüönd, R., 1982. A revision of the genus *Hydrilla* (Hydrocharitaceae). *Aquat. Bot.* 13, 485–504.

- Cope, N.J., Winterbourn, M.J., 2004. Competitive interactions between two successful molluscan invaders of freshwaters: an experimental study. *Aquat. Ecol.* 38, 83–91.
- DeVaney, S.C., McNyset, K.M., Williams, J.B., Peterson, A.T., Wiley, E.O., 2009. A Tale of Four “Carp”: Invasion Potential and Ecological Niche Modeling. *PLoS One* 4, e5451.
- Dibble, E.D.E., Kovalenko, K., 2009. Ecological impact of grass carp: a review of the available data. *J. Aquat. Plant Manag.* 47, 1–15.
- Drake, J.M., 2015. Range bagging: a new method for ecological niche modeling from presence-only data. *J. R. Soc. Interface* 12, 20150086.
- Elith, J., Leathwick, J.R., 2009. Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annu. Rev. Ecol. Evol. Syst.* 40, 677–697.
- Elith, J., Leathwick, J.R., Hastie, T., 2008. A working guide to boosted regression trees. *J. Anim. Ecol.*
- Esri, 2011. ArcGIS Desktop: Release 10. Redlands CA.
- Fitzpatrick, M.C., Gotelli, N.J., Ellison, A.M., 2013. MaxEnt versus MaxLike: empirical comparisons with ant species distributions. *Ecosphere* 4, 1–37.
- Frederico, R.G., Marco, P.D., Zuanon, J., 2014. Evaluating the use of macroscale variables as proxies for local aquatic variables and to model stream fish distributions. *Freshw. Biol.* 59, 2303–2314.
- Fuller, P.L., Nico, L.G., Williams, J.D., 1999. Nonindigenous fishes introduced into inland waters of the United States, American Fisheries Society Special Publication 27, Bethesda, Maryland. 622 pp.
- Gallardo, B., Aldridge, D.C., 2013. Priority setting for invasive species management: risk assessment of Ponto-Caspian invasive species into Great Britain. *Ecol. Appl.* 23, 352–64.
- Gies, M., Sondermann, M., Hering, D., Feld, C.K., 2015. Are species distribution models based on broad-scale environmental variables transferable across adjacent watersheds? A case study with eleven macroinvertebrate species. *Fundam. Appl. Limnol. für Hydrobiol.* 186, 63–97.
- GLCWC, 2004. Great Lakes Coastal Wetland Consortium [WWW Document]. Gt. Lakes Coast. Wetl. Consort. Invent. [Digital Data]. URL <http://glc.org/projects/habitat/coastal-wetlands/cwc-inventory/>
- Gordon, D.R., 1998. Effects of invasive, non-indigenous plant species on ecosystem processes: Lessons from Florida. *Ecol. Appl.* 8, 975–989.
- Gronewold, A.D., Fortin, V., Lofgren, B., Clites, A., Stow, C.A., Quinn, F., 2013. Coasts, water levels, and climate change: A Great Lakes perspective. *Clim. Change* 120, 697–711.
- Grosholz, E.D., Ruiz, G.M., Dean, C.A., Shirley, K.A., Maron, J.L., Connors, P.G., 2000. The impacts of a nonindigenous marine predator in a California bay. *Ecology* 81, 1206–1224.



- Gu, B., 2006. Environmental conditions and phosphorus removal in Florida lakes and wetlands inhabited by *Hydrilla verticillata* (Royle): Implications for invasive species management. *Biol. Invasions* 8, 1569–1578.
- Guisan, A., Thuiller, W., 2005. Predicting species distribution: offering more than simple habitat models. *Ecol. Lett.* 8, 993–1009.
- Guisan, A., Zimmermann, N.E., 2000. Predictive habitat distribution models in ecology. *Ecol. Modell.* 135, 147–186.
- Hanley, J.A., McNeil, B.J., 1982. The Meaning and Use of the Area under a Receiver Operating ( ROC ) Curvel Characteristic. *Radiology* 143, 29–36.
- Havel, J.E., Kovalenko, K.E., Thomaz, S.M., Amalfitano, S., Kats, L.B., 2015. Aquatic invasive species: challenges for the future. *Hydrobiologia* 750, 147–170.
- Herborg, L., Mandrak, N.E., Cudmore, B.C., Macisaac, H.J., 2007. Comparative distribution and invasion risk of snakehead (Channidae) and Asian carp (Cyprinidae) species in North America. *Can. J. Fish. Aquat. Sci.* 64, 1723–1735.
- Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* 25, 1965–1978.
- Hirzel, A.H., Le Lay, G., Helfer, V., Randin, C., Guisan, A., 2006. Evaluating the ability of habitat suitability models to predict species presences. *Ecol. Modell.* 199, 142–152.
- Hutchinson, G.E., 1958. Concluding remarks, in: *Cold Spring Harbor Symposia on Quantitative Biology*. pp. 415–427.
- Jacono, C.C., Richerson, M.M., Morgan, V.H., 2014. *Hydrilla verticillata* [WWW Document]. USGS Nonindigenous Aquat. Species Database. URL [www.nas.usgs.gov](http://www.nas.usgs.gov) (accessed 5.10.14).
- Jiménez-Valverde, A., Peterson, A.T., Soberón, J., Overton, J.M., Aragón, P., Lobo, J.M., 2011. Use of niche models in invasive species risk assessments. *Biol. Invasions* 13, 2785–2797.
- Johnson, P.T.J., Olden, J.D., Solomon, C.T., Vander Zanden, M.J., 2009. Interactions among invaders: Community and ecosystem effects of multiple invasive species in an experimental aquatic system. *Oecologia* 159, 161–170.
- Kilroy, C., Snelder, T.H., Floerl, O., Vieglais, C.C., Dey, K.L., 2008. A rapid technique for assessing the suitability of areas for invasive species applied to New Zealand’s rivers. *Divers. Distrib.* 14, 262–272.
- Kocovsky, P.M., Chapman, D.C., McKenna, J.E., 2012. Thermal and hydrologic suitability of Lake Erie and its major tributaries for spawning of Asian carps. *J. Great Lakes Res.* 38, 159–166.
- Langeland, K.A., 1996. *Hydrilla verticillata* (L.F.) Royle (Hydrocharitaceae), “The Perfect Aquatic Weed.”

Castanea 61, 293–304.

- Lockwood, J.L., Hoopes, M.F., Marchetti, M.P., 2013. *Invasion Ecology*. John Wiley & Sons, West Sussex.
- Maclsaac, H., Grigorovich, I., Ricciardi, A., 2001. Reassessment of Species Invasions Concepts: The Great Lakes Basin as a Model. *Biol. Invasions* 3, 405–416.
- Michelan, T.S., Silveira, M.J., Petsch, D.K., Pinha, G.D., Thomaz, S.M., 2014. The invasive aquatic macrophyte *Hydrilla verticillata* facilitates the establishment of the invasive mussel *Limnoperna fortunei* in Neotropical reservoirs. *J. Limnol.* 73, 598–602.
- Mills, E.L., Leach, J.H., Carlton, J.T., Secor, C.L., 1993. Exotic Species in the Great Lakes: A History of Biotic Crises and Anthropogenic Introductions. *J. Great Lakes Res.* 19, 1–54.
- MTRI, 2012. Satellite-Derived Lake Submerged Aquatic Vegetation (SAV) [Digital Data] [WWW Document]. Michigan Tech Res. Inst. URL <http://www.mtri.org/cladophora.html>
- NOAA-GLERL, 2013. Great Lakes Coastal Forecasting System vertical temperature models [Digital Data] [WWW Document]. Gt. Lakes Coast. Forecast. Syst. URL <http://www.glerl.noaa.gov/res/glcfs/>
- Noble, R.L., Bettoli, P.W., Betsill, R.K., 1986. Considerations for the Use of Grass Carp in Large, Open Systems. *Lake Reserv. Manag.* 2, 46–48.
- Petitpierre, B., Kueffer, C., Broennimann, O., Randin, C., Daehler, C., Guisan, A., 2012. Climatic Niche Shifts Are Rare Among Terrestrial Plant Invaders. *Science*. 335, 1344–1348.
- Pine, R., Anderson, L., 1991. Plant preferences of triploid grass carp. *J. Aquat. Plant Manag.* 29, 80–82.
- Porter, M.S., Rosenfeld, J., Parkinson, E.A., 2000. Predictive Models of Fish Species Distribution in the Blackwater Drainage, British Columbia. *North Am. J. Fish. Manag.* 20, 349–359.
- Pulliam, H.R., 2000. On the relationship between niche and distribution. *Ecol. Lett.* 3, 349–361.
- Raibley, P.T., Blodgett, D., Sparks, R.E., 1995. Evidence of grass carp (*Ctenopharyngodon idella*) reproduction in the Illinois and Upepr Mississippi Rivers. *J. Freshw. Ecol.* 10, 65–74.
- Rejmánek, M., Pitcairn, M.J., 2002. When is eradication of exotic pest plants a realistic goal, in: *Turning the Tide: The Eradication of Invasive Species*. pp. 249–253.
- Ricciardi, A., Maclsaac, H.J., 2000. Recent mass invasion of the North American Great Lakes by Ponto-Caspian species. *Trends Ecol. Evol.* 15, 62–65.
- Romanuk, T.N., Zhou, Y., Brose, U., Berlow, E.L., Williams, R.J., Martinez, N.D., 2009. Predicting invasion success in complex ecological networks. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 364, 1743–1754.
- Rothlisberger, J., Lodge, D., 2013. The Laurentian Great Lakes as a beachhead and a gathering place for biological invasions. *Aquat. Invasions* 8, 361–374.

- Rothlisberger, J.D., Finnoff, D.C., Cooke, R.M., Lodge, D.M., 2012. Ship-borne Nonindigenous Species Diminish Great Lakes Ecosystem Services. *Ecosystems* 15, 462–476.
- Rybicki, N.B., Carter, V., 2002. Light and temperature effects on the growth of wild celery and hydrilla. *J. Aquat. Plant Manag.* 40, 92–99.
- Schloesser, D.W., Manny, B.A., 2007. Restoration of Wild Celery, *Vallisneria americana* Michx., in the Lower Detroit River of the Lake Huron-Lake Erie Corridor. *J. Great Lakes Res.* 33, 8–19.
- Schwab, D.J., Bedford, K.W., 1994. Initial implementation of the Great Lakes forecasting system: A real-time system for predicting lake circulation and thermal structure. *Water Pollut. Res. J. Canada* 29, 203–220.
- Shireman, J., Maceina, M., 1981. The utilization of grass carp, *Ctenopharyngodon idella* Val., for Hydrilla control in Lake Baldwin, Florida. *J. Fish Biol.* 19, 629–636.
- Shireman, J., Smith, C., 1983. Synopsis of biological data on the grass carp, *Ctenopharyngodon idella* (Cuvier and Valenciennes, 1844), in: *FAO Fisheries Synopsis No. 135. Food and Agriculture Organization of the United Nations, Rome*, pp. 1–86.
- Shuchman, R.A., Sayers, M.J., Brooks, C.N., 2013. Mapping and monitoring the extent of submerged aquatic vegetation in the Laurentian Great Lakes with multi-scale satellite remote sensing. *J. Great Lakes Res.* 39, 78–89.
- Sierszen, M.E., Morrice, J.A., Trebitz, A.S., Hoffman, J.C., 2012. A review of selected ecosystem services provided by coastal wetlands of the Laurentian Great Lakes. *Aquat. Ecosyst. Health Manag.* 15, 92–106.
- Spencer, D.F., Ksander, G.G., Madsen, J.D., Owens, C.S., 2000. Emergence of vegetative propagules of *Potamogeton nodosus*, *Potamogeton pectinatus*, *Vallisneria americana*, and *Hydrilla verticillata* based on accumulated degree-days. *Aquat. Bot.* 67, 237–249.
- Stanley, J.G., Miley, W.W., Sutton, D.L., 1978. Reproductive requirements and likelihood for naturalization of escaped grass carp in the United States. *Trans. Am. Fish. Soc.* 107, 119–128.
- Steward, K.K., Van, T.K., Carter, V., Pieterse, A.H., 1984. Hydrilla invades Washington, D.C. and the Potomac. *Am. J. Bot.* 71, 162–163.
- Sutton, D.L., 1996. Depletion of turions and tubers of *Hydrilla verticillata* in the North New River Canal, Florida. *Aquat. Bot.* 53, 121–130.
- Trumpickas, J., Shuter, B.J., Minns, C.K., 2009. Forecasting impacts of climate change on Great Lakes surface water temperatures. *J. Great Lakes Res.* 35, 454–463.
- USACE, 2011. *Aquatic Herbicides*. Vicksburg, Mississippi.
- Vander Zanden, M.J., Olden, J.D., 2008. A management framework for preventing the secondary spread

- of aquatic invasive species. *Can. J. Fish. Aquat. Sci.* 65, 1512–1522.
- Vanderploeg, H.A., Liebig, J.R., Nalepa, T.F., Fahnenstiel, G.L., Pothoven, S.A., 2010. Dreissena and the disappearance of the spring phytoplankton bloom in Lake Michigan. *J. Great Lakes Res.* 36, 50–59.
- Wang, L., Riseng, C.M., Mason, L.A., Wehrly, K.E., Rutherford, E.S., McKenna, Jr., J.E., Castiglione, C., Johnson, L.B., Infante, D.M., Sowa, S., Robertson, M., Schaeffer, J., Khoury, M., Gaiot, J., Hollenhorst, T., Brooks, C., Coscarelli, M., 2015. A spatial classification and database for management, research and policy making: the Great Lakes Aquatic Habitat Framework. *J. Great Lakes Res.* 41, 584–596.
- Watson, G., Hillman, T.W., 1997. Factors Affecting the Distribution and Abundance of Bull Trout: An Investigation at Hierarchical Scales. *North Am. J. Fish. Manag.* 17, 237–252.
- Wenger, S.J., Olden, J.D., 2012. Assessing transferability of ecological models: An underappreciated aspect of statistical validation. *Methods Ecol. Evol.* 3, 260–267.
- Wiley, M.J., Hyndman, D.W., Pijanowski, B.C., Kendall, A.D., Riseng, C., Rutherford, E.S., Cheng, S.T., Carlson, M.L., Tyler, J.A., Stevenson, R.J., Steen, P.J., Richards, P.L., Seelbach, P.W., Koches, J.M., Rediske, R.R., 2010. A multi-modeling approach to evaluating climate and land use change impacts in a Great Lakes River Basin. *Hydrobiologia* 657, 243–262.
- Wittmann, M.E., Barnes, M.A., Jerde, C.L., Jones, L.A., Lodge, D.M., 2016. Confronting species distribution model predictions with species functional traits. *Ecol. Evol.* 6, 873–879.
- Wittmann, M.E., Jerde, C.L., Howeth, J.G., Maher, S.P., Deines, A.M., Jenkins, J.A., Whitley, G.W., Burbank, S.R., Chadderton, W.L., Mahon, A.R., Tyson, J.T., Gantz, C.A., Keller, R.P., Drake, J.M., Lodge, D.M., 2014. Grass carp in the Great Lakes region: Establishment potential, expert perceptions, and re-evaluation of experimental evidence of ecological impact. *Can. J. Fish. Aquat. Sci.* 71, 992–999.
- Yousef, F., Shuchman, R., Sayers, M., Fahnenstiel, G., In Review. Water Clarity of the Upper Great Lakes: Dramatic Changes Between 1998-2012. *J. Great Lakes Res.*
- Zhang, X., Liu, X., Ding, Q., 2013. Morphological responses to water-level fluctuations of two submerged macrophytes, *Myriophyllum spicatum* and *Hydrilla verticillata*. *J. Plant Ecol.* 6, 64–70.

## Tables

Species	Records (N)	AUC	Bioclim variables of greatest importance	Relative Importance of Bioclim variable (Proportion)	Great Lakes niche restriction scenarios	Reference supporting restriction variable
<i>Hydrilla</i> ( <i>H. verticillata</i> )	1017	0.896	Mean diurnal temp	0.06	(1) Accumulated Growing Degree days ( $n = 500$ )	Spencer et al. 2000
			Isothermality	0.06		
			Precip of warmest quarter	0.10	(2) Accumulated Growing Degree days + Photic Zone	Canfield and Langeland 1985
grass carp ( <i>C. idella</i> )	663	0.877	Mean temp of wettest quarter	0.14	(1) Submersed Aquatic Vegetation + Wetlands	Nixon and Miller, 1978
			Mean temp of driest quarter	0.09		
			Mean temp of warmest quarter	0.09	(2) Submersed Aquatic Vegetation + Wetlands + <i>Hydrilla</i> Niche	Bain, 1993
			Precip of driest quarter	0.07		

Table 1. Summary table for range bagging species distribution model outputs and restriction variables.

Two species were modeled using global occurrences (2nd column). AUC = Area Under Curve evaluation and Bioclim variables of greatest importance included in columns 3 and 4. Column 5 indicates the relative importance of each Bioclim variable for each species prediction. Columns 6 and 7 indicate the habitat data layers used to restrict the niche predictions for each species, and associated literature references used to motivate the restriction layer.

## Figure legends

Figure 1. Global occurrences ( $n=1017$ ) of *Hydrilla verticillata* (a; top panel) and niche centrality based on global occurrence records (b; bottom panel). Niche centrality refers to the tendency of an environment to be within the environmental range of a species across multiple environmental variables.

Figure 2. Niche centrality for *Hydrilla verticillata* for the comprehensive Great Lakes watershed region (a; top panel) and clipped using accumulated growing degree days (GDD) based on benthic temperature observations (b; middle panel), and clipped using GDD and photic zone (c; bottom panel). High values of niche centrality indicate climate conditions in the Great Lakes basin fall generally within the predicted niche.

Figure 3. Global occurrences ( $n=663$ ) of grass carp (*Ctenopharyngodon idella*) (a; top panel) and niche centrality based on global occurrence records (b; bottom panel). Niche centrality refers to the tendency of an environment to be within the environmental range of a species across multiple environmental variables.

Figure 4. Niche centrality for grass carp for the comprehensive Great Lakes watershed region (a; top panel) and clipped using a submersed aquatic vegetation (SAV) and wetlands data layer (b; middle panel) and a combined SAV, Wetlands and predicted *Hydrilla verticillata* niche (c; bottom panel). High values of niche centrality indicate climate conditions in the Great Lakes basin fall generally within the predicted niche.









