

Research Article

Ensemble forecasting of potential habitat for three invasive fishes

Helen M. Poulos^{1*}, Barry Chernoff^{1,2}, Pam L. Fuller³ and David Butman⁴¹ College of the Environment, Wesleyan University, Middletown, CT 06457, USA² Departments of Biology and Earth and Environmental Studies, Wesleyan University, Middletown, CT 06547, USA³ Nonindigenous Aquatic Species Program, Southeast Ecological Science Center, Gainesville, FL 32653, USA⁴ Yale School of Forestry and Environmental Studies, New Haven, CT 06511, USAE-mail: hpoulos@wesleyan.edu (HMP), bchernoff@wesleyan.edu (BC), pfuller@usgs.gov (PLF), david.butman@yale.edu (DB)

*Corresponding author

Received: 9 May 2011 / Accepted: 15 August 2011 / Published online: 14 October 2011

Abstract

Aquatic invasive species pose major ecological and economic threats to aquatic ecosystems worldwide via displacement, predation, or hybridization with native species and the alteration of aquatic habitats and hydrologic cycles. Modeling the habitat suitability of alien aquatic species through spatially explicit mapping is an increasingly important risk assessment tool. Habitat modeling also facilitates identification of key environmental variables influencing invasive species distributions. We compared four modeling methods to predict the potential continental United States distributions of northern snakehead *Channa argus* (Cantor, 1842), round goby *Neogobius melanostomus* (Pallas, 1814), and silver carp *Hypophthalmichthys molitrix* (Valenciennes, 1844) using maximum entropy (Maxent), the genetic algorithm for rule set production (GARP), DOMAIN, and support vector machines (SVM). We used inventory records from the USGS Nonindigenous Aquatic Species Database and a geographic information system of 20 climatic and environmental variables to generate individual and ensemble distribution maps for each species. The ensemble maps from our study performed as well as or better than all of the individual models except Maxent. The ensemble and Maxent models produced significantly higher accuracy individual maps than GARP, one-class SVMs, or DOMAIN. The key environmental predictor variables in the individual models were consistent with the tolerances of each species. Results from this study provide insights into which locations and environmental conditions may promote the future spread of invasive fish in the US.

Key words: invasive fishes, ensemble mapping, *Channa argus*, *Neogobius melanostomus*, *Hypophthalmichthys molitrix*

Introduction

Aquatic invasive species pose major ecological and economic threats to lakes and waterways worldwide through the displacement of native species and the alteration of hydrologic cycles (Wilcove et al. 1998; Ricciardi and MacIsaac 2000; Pimentel et al. 2005). Early detection and prevention of species' invasion are paramount for maintaining the ecological integrity of uninvaded habitats (Simberloff 2003; Vander Zanden and Olden 2008). Yet, forecasting the future spread of nonindigenous species to new locations remains difficult due to the complex interactions among exotics, natives, humans, and local environmental conditions (Ricciardi and Rasmussen 1998; Rejmánek 2000; Moles et al. 2008).

Predicting invasive species establishment and spread is an essential component of developing management guidelines for early detection and eradication. Species distribution modeling

(SDM) offers a method for generating spatially explicit information for prioritizing invasive species management by identifying potential invasive habitat. Recent advances in geographic information systems (GIS), data mining methods, and data availability have led to the proliferation of a wide array of SDM approaches. The majority of SDM methods rely on the concept of the fundamental niche (Grinnell 1917), or the constraints of a species range, as a means of estimating a species' potential distribution given a limited number of known occurrences. SDM models use empirical data to describe and predict species distributions through statistical examination of the species-environment relationship (Guisan and Zimmermann 2000; Guisan and Thuiller 2005).

The performance of individual SDMs varies widely among methods and species (Elith et al. 2006; Marmion et al. 2009; Diniz-Filho et al. 2009). Consensus methods that integrate multiple individual models provide robust

estimates of potential species' distributions (Araújo and New 2007; Marmion et al. 2009). Ensemble maps that highlight areas of agreement among model predictions offer a way to reduce the uncertainty of basing management activities on solely one SDM model. However, the degree to which ensemble models outperform individual SDM models is still widely debated (Marmion et al. 2009; Stohlgren et al. 2010). Nor has the ensemble approach been widely applied to mapping the potential habitat of invasive species (but see Roura-Pascual et al. 2009; Stohlgren et al. 2010).

Our goal was to explore the utility of the ensemble mapping approach for modeling the potential habitat of three high-profile fish invaders across the continental United States including the northern snakehead *Channa argus* (Cantor, 1842), round goby *Neogobius melanostomus* (Pallas, 1814), and silver carp *Hypophthalmichthys molitrix* (Valenciennes, 1844). We compared four individual SDM methods including maximum entropy (Maxent; Phillips et al. 2006; Phillips and Dudík 2008), DOMAIN; Carpenter et al. 1993), the genetic algorithm for rule-set production (GARP; Stockwell and Noble 1992; Stockwell and Peters 1999), one-class support vector machines (SVM; Scholkopf et al. 2001; Guo et al. 2005) and an ensemble model (Stohlgren et al. 2010) to estimate the potential distribution of each of the species in our study. Our specific objectives included: 1) identifying regions with a high risk of invasion, and 2) describing the environmental settings that were most susceptible to fish invasion.

Species accounts

The northern snakehead is native to rivers of eastern China, southern Siberia, and Korea (Courtenay and Williams 2004). It is an air breathing piscivorous fish that can survive for up to four days out of water. Their temperature tolerance ranges from 0 to over 30°C and they can survive under ice (Okada 1960; Courtenay and Williams 2004). Snakeheads prefer stagnant shallow ponds, swamps and rivers with muddy substrates and emergent vegetation (Dukravets and Machulin 1978; Dukravets 1992). The northern snakehead is a prominent aquaculture species in China where it is grown in ponds, rice paddies, and reservoirs (Atkinson 1977; Sifa and Senlin 1995). The species was most likely introduced into the United States through live-

food fish markets (Courtenay and Williams 2004).

Silver carp is a large-bodied filter-feeder native to large rivers, and warm water lakes, ponds and backwaters that flood frequently in subtropical and temperate Asia (Kolar et al. 2005). Silver carp requires rivers that are large and long enough for eggs to drift downstream for several days before they hatch (Schofield et al. 2005). Silver carp are frequently associated with eutrophic open areas of standing or slow-flowing water (Rassmussen 2002) in the upper and middle layers of the water column in both their native and introduced ranges (Kolar et al. 2005). Silver carp are adapted to a wide range of temperatures (16–40°C), although their optimal range falls between 26–33.5°C (Opuszyński et al. 1989). The species was intentionally introduced into the southern United States for eutrophication control in lakes, reservoirs, and sewage treatment facilities (Williamson and Garvey 2004; Kolar et al. 2005). They subsequently escaped and established breeding populations throughout much of the Mississippi basin after major flooding events.

The round goby is a small benthic fish native to the Black and Caspian Seas of Eastern Europe (Charlebois et al. 2001). Round gobies feed primarily on arthropods as juveniles, but they shift feeding to mollusks in adulthood (Jude and DeBoe 1996; French and Jude 2001; Carman et al. 2006). They feed principally on zebra mussels (*Dreissena polymorpha* Pallas, 1771) in both their native and introduced habitats. Round gobies are euryhaline and tolerant of low dissolved oxygen content and a wide range of temperatures (-1–30°C) (Moskal'kova 1996; Weimer and Keppner 2000). This species occupies a variety of habitats including coarse gravel, shell, and sandy inshores in its native range, but prefers cobble substrates within its introduced range in the United States (Charlebois et al. 1997; Ray and Corkum 2001). Initial round goby records in the Great Lakes were in harbor locations experiencing heavy inter-lake and international shipping, suggesting that the introduction of this species occurred from discharged ship ballast water (Jude and Smith 1992). The species spread prolifically after introduction due to its aggressive behavior, ability to spawn repeatedly throughout the spring and summer, and parental care by males (Charlebois et al. 1997; MacInnis and Corkum 2000).

Methods

Species distribution data

We used spatial species occurrence data from the Nonindigenous Aquatic Species Database (NAS 2011) for silver carp ($n = 270$), round goby ($n = 639$), and northern snakehead ($n = 189$). Presence data included observations from 1993 to 2010 for round goby, from 1975 to 2010 for silver carp, and from 2002 to 2010 for northern snakehead. While the NAS database maintains distribution information from across the United States for over 1000 aquatic species outside their native ranges, we chose these species because they represented aggressive invaders with more than 100 occurrence records in the database. NAS data were derived from a range of sources including literature, museum specimens, monitoring programs, state and federal agencies, and individual observation reports. The NAS database harbors the most complete record of nonindigenous aquatic invaders for the United States, however sources some potential sampling bias may have existed in our datasets due to species inventories that were close to roads, population centers, and universities.

Environmental data

We derived a set of 20 raster-based climatic, topographic, and spectral habitat predictor variables from a range of raster data providers (Table 1) to explore the environmental influences on species distribution patterns across the continental United States. We chose these candidate environmental predictors based on their relevance to riparian systems and their biological influences on the fish in this study. Grids were clipped to the extent of the USGS hydrography dataset to avoid modeling fish distributions outside riparian areas (USGS 2003). We also resampled each grid to 1 km from its native spatial resolution, which ranged from 30 m to 1 km using ArcMap v9.3 (ESRI 2008). We chose this resolution based on our intent to model national-scale species distributions and because it was the coarsest native resolution of any dataset.

The entire dataset of raster predictor variables was reduced through pairwise evaluation for each species to reduce multicollinearity among the predictors as suggested by Elith et al. (2010). We followed methods outlined by Stohlgren et al. (2010) which used the correlation matrix as a means of identifying highly correlated pairs of

habitat predictors ($r > 0.7$). For correlated pairs, we removed the less interpretable predictor variable which reduced the dataset to approximately 15 variables for each species. For example, if January minimum temperature and mean minimum temperature were highly correlated, we kept mean minimum temperature since it captured a longer record of winter temperature as a whole.

Species distribution modeling

We fit four well-known species distribution models to produce ensemble prediction maps for each species. Maximum entropy (Maxent) (Phillips et al. 2006; Phillips and Dudík 2008), DOMAIN (Carpenter et al. 1993), the open modeler implementation of the genetic algorithm for rule-set production (GARP-OM) (Anderson et al. 2003), and one-class support vector machines (SVM) (Scholkopf et al. 2001, Guo et al. 2005) were chosen for inclusion based on their good performance with presence-only data and because they differ both conceptually and statistically (Elith et al. 2006; Kelly et al. 2007). Maxent was run using the Maxent software for species habitat modeling v3.3.3e (Phillips et al. 2006). The DOMAIN and one-class SVMs models were run in ModEco (Guo et al. 2010), and GARP-OM was run using OpenModeller Software (Muñoz et al. 2011).

Maxent uses a deterministic algorithm that finds the optimal probability distribution (potential distribution) of a species across a study area based on a set of environmental constraints. Maxent determines the best potential distribution by selecting the most uniform distribution subject to the constraint that each environmental variable in the modeled distribution matches its empirical average over the known distributional data (i.e. presence data). Maxent is sensitive to sampling biases in clustered or disparate datasets such as the ones included in this study. We followed the methods employed by Jarnevich and Reynolds (2011) to limit the spatial extent from which Maxent could select background points to locations within 50 kilometers of an existing presence record. DOMAIN is a presence-only nearest neighbor method that uses the Gower metric to estimate the environmental similarity between a site of interest and the nearest presence record in environmental space. The GARP-OM algorithm uses a set of rule types (logistic, regression, range rules, negated range rules, and atomic

Table 1. Environmental predictor variables evaluated for inclusion in the Red Shiner species distribution models, their native spatial resolution and data sources.

Predictor variables	Description	Native spatial resolution	Source
Climatic variables			
ppt	mean precipitation 1980-1997	1 km	Daymet ¹
Tmax	mean minimum temperature 1980-1997	1 km	Daymet ¹
Tmin	mean maximum temperature 1980-1997	1 km	Daymet ¹
jantmin	mean minimum temperature of the coldest month	1 km	Daymet ¹
augtmax	mean maximum temperature of the hottest month	1 km	Daymet ¹
sumprecip	average summer precipitation	1 km	Daymet ¹
summheat	summer heat: mean warmest month temperature)/(mean summer precipitation/1000)	1 km	Daymet ¹
annheat	annual heat: (mean annual temperature + 10)/mean annual precipitation/1000)	1 km	Daymet ¹
ppt	mean annual precipitation	1 km	Daymet ¹
remotely sensed variables			
NDVI	MODIS Normalized Difference Vegetation Index	500 m	MODIS ²
LAI	MODIS July Leaf Area Index	500 m	MODIS ²
topographic and environmental variables			
elevation	elevation in meters	30 m	National Hydrography Dataset ³
% flat	percent flat elevation in watershed	30 m	National Hydrography Dataset ³
flow accumulation	flow accumulation from ArcHydro extension of ArcGIS	30 m	National Hydrography Dataset ³
watershed slope	slope of the watershed in degrees	30 m	National Hydrography Dataset ³
downslope	calculated as the difference between a cell's elevation in meters and the lowest elevation	30 m	National Hydrography Dataset ³
elevation change			
sand (%)	percent sand in soil		National Hydrography Dataset ³
stream power index	erosive power of overland flow measured as the area of the catchment area multiplied by the tangent of the slope calculated using Whitebox Geospatial Analysis Tools (Lindsay 2009)	30 m	National Hydrography Dataset ³
upslope neigh	number of upslope neighbors	30 m	National Hydrography Dataset ³
sed trans	sediment transport capacity index (Moore and Burch 1986)	30 m	National Hydrography Dataset ³
baseflow	baseflow index grid from USGS stream gauges, expressed as percentage of baseflow relative to total flow	30 m	USGS ⁴

Data sources:

¹Daymet: <http://www.daymet.org>

²MODerate resolution Imaging Spectroradiometer (MODIS) vegetation Indices: <http://mrtweb.cr.usgs.gov>

³National Hydrography Dataset: <http://www.horizon-systems.com/nhdplus>

⁴United States Geological Survey (USGS): <http://landcover.usgs.gov/usgslandcover.php>

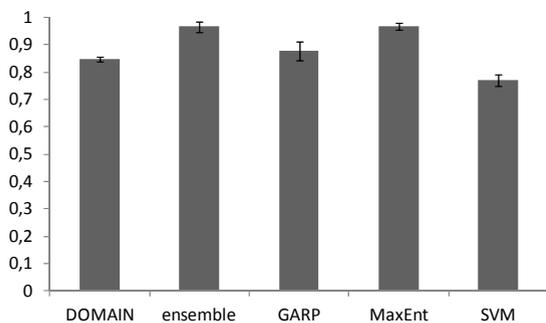
rules), or if-then statements that describe a species' niche. It operates by scanning broadly across the environmental space and iteratively refines solutions that show high optimization values to produce a prediction map. We used the Open Modeller best subsets implementation of the algorithm and ran 100 replicate models with a convergence limit of 0.01. The final model resulted from a summation of the top 10 of the 100 model runs. One-class SVMs are an adaptation of traditional two-class SVMs that seek to partition multidimensional data using

kernel-based hyperplanes. The one-class algorithm essentially maps the presence data into a feature space using a kernel density function and then separates the mapped vectors from the origin (the original member of the absence class).

We ran each model by randomly dividing our data into training and validation datasets comprising 70% and 30% of each dataset, respectively. We produced the final individual models by including only those predictor variables that explained >1% of the model

Table 2. The area under the curve (AUC) values for each species and model type calculated using the ROC calculator by Eng (2006).

species	model	AUC
Silver carp	ensemble	0.99
	MaxEnt	0.95
	GARP	0.91
	DOMAIN	0.86
	SVM	0.80
Round goby	ensemble	0.93
	MaxEnt	0.96
	GARP	0.81
	DOMAIN	0.85
	SVM	0.78
Northern snakehead	ensemble	0.98
	MaxEnt	0.99
	GARP	0.91
	DOMAIN	0.83
	SVM	0.73

**Figure 1.** Accuracy (area under the curve, AUC) comparisons among the different modeling methods used in the study. Models differ significantly according to a one-way ANOVA. Significant pairwise differences in performance are shown in lower case letters from Tukey's Pairwise comparisons.

variation which reduced the number of predictors to approximately 10 data layers. We converted the continuous maps to binary presence/absence maps by selecting thresholds where sensitivity was equal to specificity following Liu et al. (2005).

We generated ensemble maps for each species using a simple weighted summation approach for combining the four binary maps from each modeling method and multiplying it by the individual model performance (area under the receiver operating characteristic curve) described in the following paragraph. We chose the weighted summation approach based on its straightforward and easily interpretable output. In the final species maps, the value of each pixel

represented the degree of model agreement. A pixel with a value of zero indicated that none of the models predicted species presence at that location. Larger values indicated higher prediction agreement among the three models, which suggested that these locations comprised suitable potential habitat.

Model accuracy was assessed using the test data to calculate the area under the receiver operating characteristic (ROC) curve (AUC) (Fielding et al. 1997). AUC values typically range from zero to one. Values > 0.9 indicate high accuracy, values of 0.7–0.9 indicate good accuracy, and values below 0.7 indicate low accuracy (Swets 1988). AUC values were calculated using the web-based calculator for ROC curves (Eng 2006). This web-based calculator was chosen because it can handle a range of data types including binary, ordinal, and continuous scales. We generated 200 pseudo-absence points for each species to perform the AUC analysis. Pseudoabsence points were generated for each species in ArcMap by randomly generating points using a 100 km buffer around species presence records. We tested for significant differences among the models using a one-way ANOVA to compare AUCs with species as the replicate followed by Tukey's pairwise comparisons to test for significant differences among the individual and ensemble modeling techniques. Data were log-transformed prior to analysis to fulfill ANOVA assumptions of linearity.

Results

Model performance

All of the individual habitat suitability models reached AUC values above 0.7 (Table 2), indicating good overall prediction accuracy. The individual and ensemble model accuracies differed significantly ($P < 0.0001$) (Figure 1). The Maxent models consistently reached AUC values above 0.9 for all species. Maxent displayed significantly superior performance over all other individual models. GARP was the second highest performing model, reaching AUC values above 0.9 for all species except round goby (AUC = 0.82). However, it did not perform significantly better than DOMAIN. The DOMAIN and SVM models were the poorest performing algorithms in this study (AUC values ranging from 0.73 to 0.86).

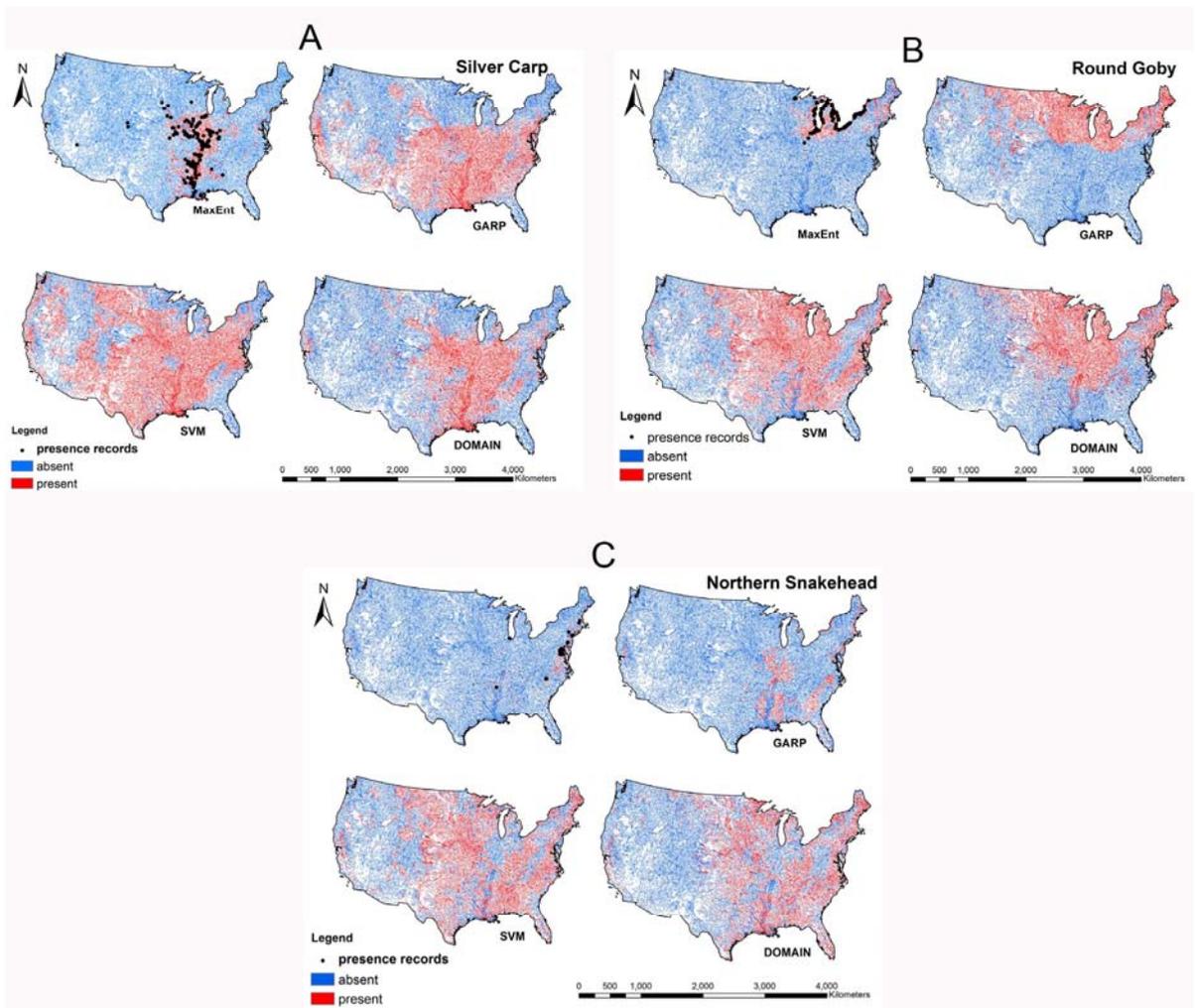


Figure 2. Species occurrences and individual model variation for A) Silver carp, B) Round goby, and C) Northern snakehead.

The individual habitat suitability maps varied considerably by species and by algorithm (Figure 2). Maxent was the most conservative model (i.e., it produced a smaller area of suitable habitat), followed in order by GARP, DOMAIN, and SVM. Maxent tended to restrict the predicted habitat to locations close to existing presence records, while SVM projected the potential habitat of all species across much of the United States. Differences among these models in the number of predicted habitat pixels

followed the individual model predictive performance noted in Table 2. More conservative models like Maxent and GARP displayed higher predictive performance, while models that predicted more extensive habitat areas including DOMAIN and SVM resulted in acceptable, but lower accuracy maps.

The ensemble models demonstrated high prediction accuracy (AUC 0.93–0.99) that surpassed all of the individual models except Maxent (Table 2; Figures 3–5). The silver carp

Ensemble forecasting of potential habitat for three invasive fishes

Figure 3. Ensemble Map for silver carp. Ensemble values were calculated using a weighted summation approach where each model was weighted by its performance and then summed together. Values range from 0-4.

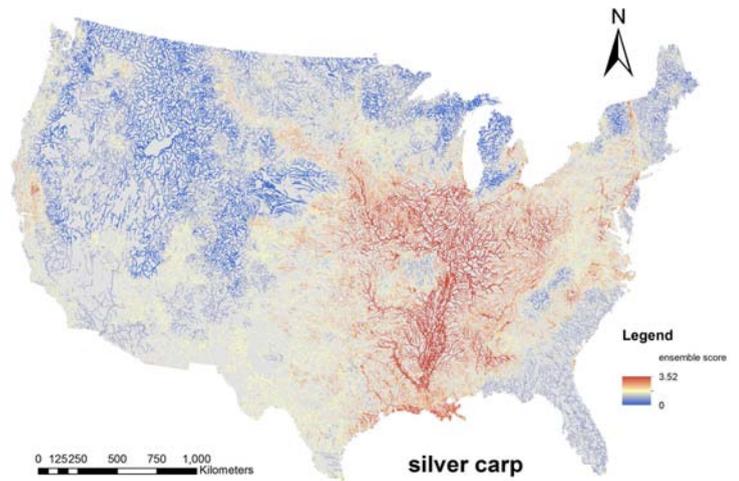


Figure 4. Ensemble Map for round goby. Ensemble values were calculated using a weighted summation approach where each model was weighted by its performance and then summed together. Values range from 0-4.

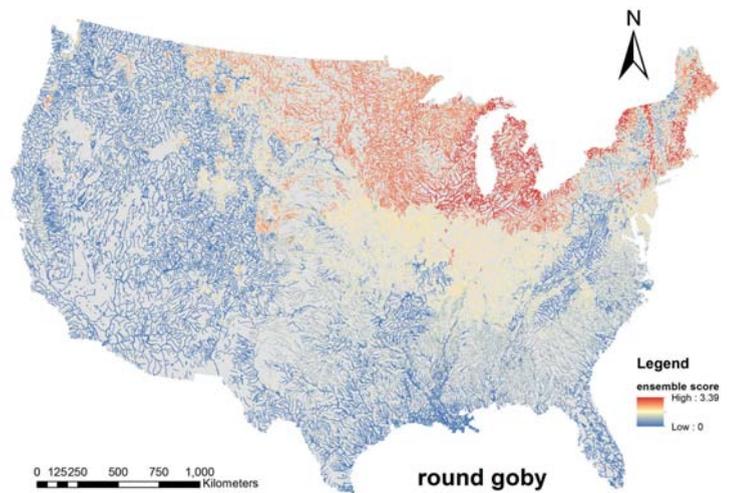
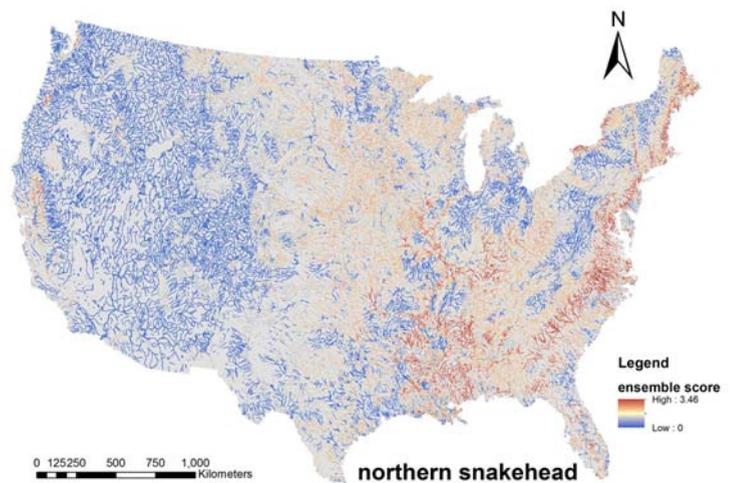


Figure 5. Ensemble Map for northern snakehead. Ensemble values were calculated using a weighted summation approach where each model was weighted by its performance and then summed together. Values range from 0-4.



ensemble model predicted that it was capable of spreading beyond its current distribution in the Mississippi watershed to the Ohio River and the eastern United States. All four of the round goby models suggested that this species could spread to portions of New England and to inland rivers surrounding its present distribution in the Great Lakes. Three of the four models (GARP, DOMAIN, and SVM) also indicated that the majority of the upper Midwest represented potential round goby habitat. The snakehead ensemble model demonstrated high agreement across the northeastern United States. The ensemble map suggested that this species could spread to portions of the southeastern United States including the Mississippi River system.

Environmental influences on potential habitat

We used the Maxent output to explore how the environmental variables shaped species distribution patterns. We chose Maxent for evaluating the influence of the environmental predictors on fish habitat since it was the highest performing individual model in this study, and because it was the only algorithm we used that evaluated species' responses to the environmental predictors.

Elevation, percent sand, the normalized difference vegetation index (NDVI), and baseflow were the main variables influencing potential silver carp habitat (Figure 6). Elevation contributed 27.5% to the model followed by percent sand (15%), NDVI (10.3%), and baseflow (10%). The silver carp species response curves suggested that it could invade elevations near 100 m that had slow-flowing water, low sand content, and little vegetative cover.

Round goby habitat was restricted to sites with maximum air temperatures ranging from 13 to 15°C (northerly latitudes) with intermediate precipitation (Figure 7). High leaf area index (LAI) was another indicator of round goby habitat, which suggested that rivers with dense forest canopy cover and sites within 1 km of forested shorelines at high latitudes were most susceptible to round goby colonization. Slope, mean minimum air temperature, and annual heat (mean annual air temperature: mean precipitation) were other minor round goby predictors. Mean maximum air temperature contributed 29% to the model, followed by LAI (17.9%), precipitation (13.9%), slope (6.4%), mean minimum air temperature (5.8%), and annual heat (5.5%).

Northern snakehead habitats were characterized by low elevation, slow-flowing waters with emergent vegetation (Figure 8). Snakehead habitat was also predicted across middle latitudes where maximum temperatures ranged between 18 and 20°C. High NDVI and flat watersheds were other minor predictors of snakehead habitat. Elevation contributed 36% to the northern snakehead model, followed by runoff (14%), mean maximum air temperature (9%), and NDVI (8%).

Discussion

Consensus methods

Our results illustrate the utility of the ensemble approach for modeling the potential spread of invasive fishes across the United States using multiple lines of evidence for guiding management. However, the ensemble maps in this study do not significantly outperform the individual Maxent model. So while consensus methods that combine multiple algorithms have the potential to provide more realistic species distribution simulations because they highlight areas of agreement among models and include a range of modeling techniques (Araújo and New 2007; Marmion et al. 2009; Stohlgren et al. 2010), using ensemble maps that incorporate lower fidelity SDMs (i.e. DOMAIN and SVM) may be less desirable than basing management on a single, high-performing model.

Moreover, the Maxent modeling approach is the only method that provides key information about the environmental tolerances of the fishes in our study that can be used for protecting susceptible habitats from future invasion. The potential habitats of our study species are consistent with their site preferences within both the species' native and invaded ranges. All of the fishes in our study demonstrate high tolerance to extreme habitats (i.e. stagnant, turbid waters), which is a typical characteristic of successful alien invaders (Kolar and Lodge 2002; Sakai et al. 2001).

Northern snakehead

The ensemble forecast for the northern snakehead predicts its highest likelihood occurrence around population centers on the mid and northeast Atlantic slope which are areas likely to have human populations that traditionally

Figure 6. Environmental response curves for variables contributing greater than 10% to the silver carp model.

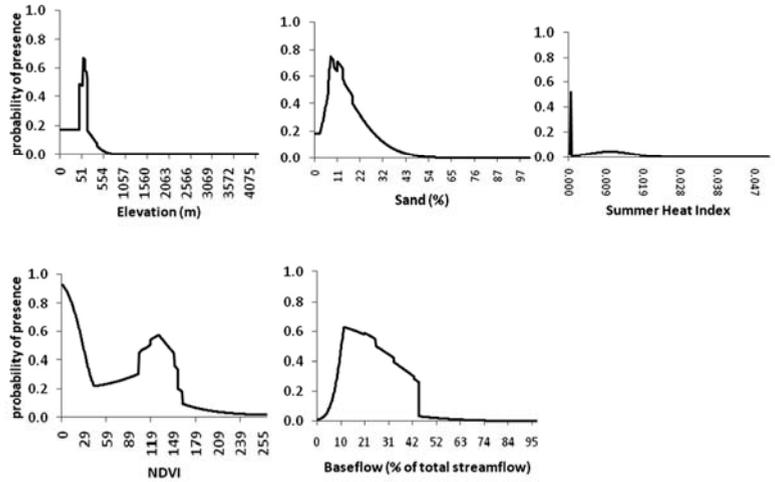


Figure 7. Environmental response curve for variables contributing greater than 10% to the round goby model.

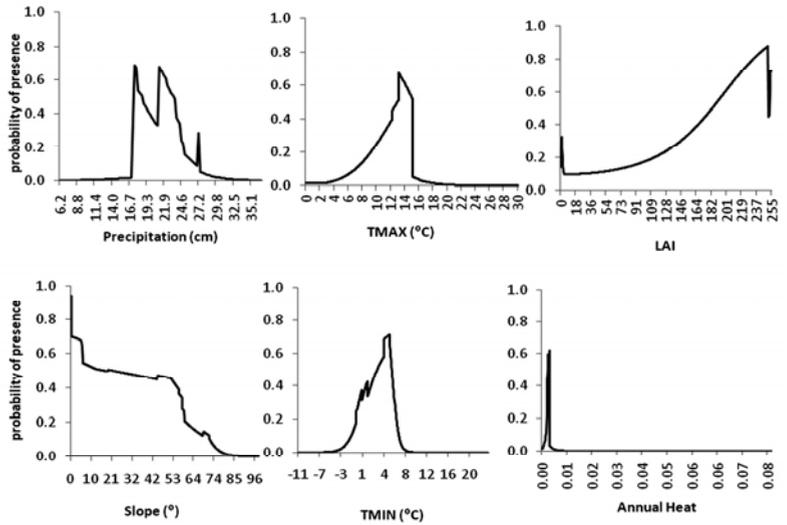
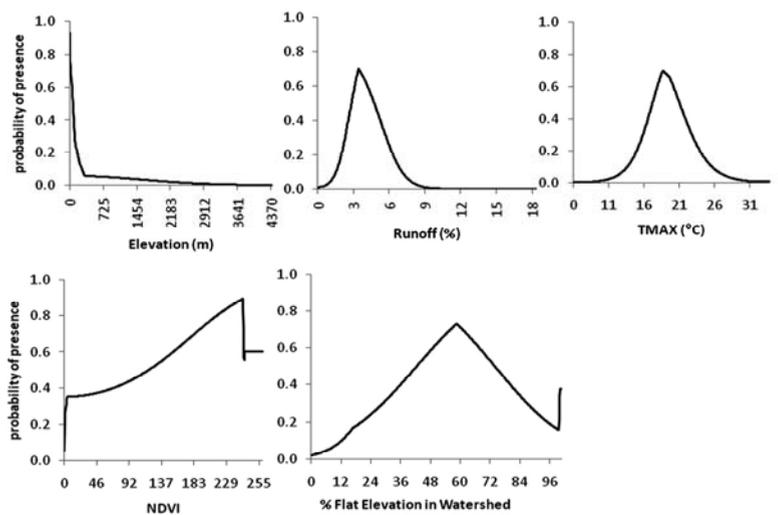


Figure 8. Environmental response curves for variables contributing greater than 10% to the northern snakehead model.



use northern snakeheads as a protein source (Courtenay and Williams 2004). Our model predictions are consistent with northern snakehead preference of shallow waters along stagnant channel margins and embayments with submersed or floating aquatic vegetation (Courtenay and Williams 2004; Odenkirk and Owens 2005). Middle latitudes (maximum air temperatures of 18–20°C) in the United States are also susceptible to northern snakehead invasion, which matches the latitudinal range of its native distribution in China.

The prediction map indicates that this species could also spread to warmer parts of the southeastern United States and Florida. Although other, more tropical snakehead species maintain reproductive populations in the southeastern United States, the more temperate northern snakehead has failed to do so to date. Yet, our models suggest that this region has environmental conditions that could promote the success of this species with repeated future introductions.

Silver carp

The Silver carp model predictions are consistent with this species' preference of disturbed sites with slow-flowing, silty waters like those in shallow channel borders and inland side channels with low dissolved oxygen content (De Grandchamp 2006). However, the movement of this species into new regions may be limited by silver carp reproductive tolerances and food availability. Risk assessment based on temperature, food, and predators suggests that silver carp spread may be limited to waters that are warmer than 12°C and have sufficient phytoplankton food sources (Cooke et al. 2009; G. R. Smith, personal communication, July 2010).

The northern snakehead and silver carp ensemble maps improve upon prior predictive maps by Herborg et al. (2007) and Chen et al. (2007). The 1 km resolution datasets used in this study versus the 0.5 degree data used by Herborg et al. (2007) may explain some of the differences between our results and previous work. The high Maxent, GARP, and ensemble map AUC values relative to other existing GARP-based risk maps for these two species may also be related to our inclusion of a wide range of climatic, topographic and hydrologic variables. These factors could explain why our models reached AUC values above 0.9, while Herborg et al.

(2007) and Chen et al. (2007) use a limited suite of climatic predictors and generally report lower AUC values.

These results differ from Herborg et al. (2007) who identify thermal tolerance as the major factor influencing the potential spread of snakeheads and carp. Our inclusion of a wide range of environmental parameters well beyond the limited climatic and topographic data used by Herborg et al. (2007) may have allowed us to identify other key silver carp habitat requirements. While Kolar and Lodge (2001) suggest that the Great Lakes Region was not susceptible to invasion by silver carp, our ensemble map, those of Mandrak and Cudmore (2004), Herborg et al. (2007) and Chen et al. (2007) and recently documented carp spread into the Mississippi and Illinois Rivers demonstrate the vulnerability of these region to silver carp spread. The risk of silver carp invasion into this region could have major impacts on the \$7.1 billion dollar sport-fishing and tourism industries, especially in light of the dangers associated with angler collisions with jumping silver carp from boat motor noise and the potential negative effects on the food web by a large planktivorous fish (Ricciardi 2001). Tactics for preventing the spread of silver carp from the Illinois River and Lake Michigan shipping canals northward into the Great Lakes are costly and widely debated, e.g., \$18 million electric fish barriers, closure of the Chicago Sanitary and Ship Canal, and banning importation of live food fish and bait fish. Preventative strategies include the establishment of several electroshock barriers and proposed lock closures of the Chicago and Sanitary Ship Canal (Brammeier et al. 2008, Hansen 2010). However, Silver Carp may not spread to the main waters of the Great Lakes because over 99% of the waters are too cold (colder than 12°C) and may have insufficient food due to zebra and quagga mussel consumption of plankton (Cooke et al. 2009; G. R. Smith, personal communication, July 2010). Moreover, if they do establish in the shallow littoral zone of the southern Great Lakes, they will encounter more species of large predators than these carp have ever encountered, anywhere in the world.

Round goby

Round goby is a denizen of large cold water lakes and rivers in its native habitat and it is

predicted to spread to similar environments within the United States. Round goby has the potential to move from the Great Lakes to surrounding Midwestern tributaries and lakes. Although the Maxent round goby model shows that its spread may be limited to locations with high precipitation and maximum air temperatures below 15°C (i.e. higher latitudes), the ensemble map indicates that round goby could spread across much of the Midwestern United States via the Illinois and Missouri Rivers, and to parts of New England including the Connecticut River. Moreover, its thermal threshold represents its lower limit meaning that it could move northward into Canadian tributaries of the Great Lakes Region. Its dependence as an adult on freshwater mollusks may limit its spread to areas where there are abundant food sources. However, the rapid spread of zebra and quagga Mussels across the United States means that there is abundant food for this species in many waterways (Drake and Bossenbroek 2004; Carman et al. 2006; Stokstad 2007).

In their native range, Round Gobies occur predominantly in large lakes and they are often associated with rooted aquatic macrophytes (Ray and Corkum 2001). However, recent collection data show that the gobies commonly occur without rooted macrophytes along the western edge of Lake Michigan or in large closed-canopy tributaries to the Great Lakes (Chernoff, pers. obs., July 2006), which may explain the importance of LAI as a predictor of round goby habitat. This is another instance of the well-documented characteristics of invader species that they have the ability to rapidly and repeatedly adapt to newly colonized environments (e.g., Lee 1999).

Caveats

Although our habitat suitability models performed well, the model error highlights the limitations of our approach for extrapolating species distributions geographically and the difficulties associated with reducing ecological processes to numerical models. Additional model evaluation using independent data or map validation over time could improve map reliability.

Risk management and decision support

Predicting the potential occurrence of exotic invaders is a useful tool for resource managers,

especially in national parks, fish and wildlife refuges, and for areas having high aquatic biodiversity or high imperilment. Organizations like the Nonindigenous Aquatic Species Alert System, the Invasive Plant Atlas of New England, the Global Invasive Species Database, and the National Invasive Species Information Center provide key resources for rapid response and early mitigation of invasive threats. The current National Invasive Species Management Plan (NISC 2008) acknowledges the value of predictive mapping as a decision support tool for stopping the spread of alien species by identifying high risk habitats for focusing spread prevention efforts. Prohibiting the sale of invasive fishes as live food or bait in highly susceptible areas is one potentially effective method for reducing the risk of new introductions to sensitive habitats (Keller and Lodge 2007). Mitigating the risk of spread through interconnected waterways by establishing electrical or physical barriers is another potential, albeit contentious way of reducing the risk of future invasive spread. While commerce remains a top policy priority for lawmakers (NISC 2008), restricting the movement of known aggressive invaders in highly vulnerable regions could protect uninvaded habitats from experiencing major shifts in ecosystem structure as a result of new species introductions.

Developing abundance-based models for these invaders would build upon this research by identifying locations that are vulnerable to the proliferation of round goby, silver carp, and northern snakehead. Such maps would provide a higher level of decision support to managers by pinpointing areas that are likely to host high numbers of invaders in the future. Other models that predict the rate of spread or the leading edge of invasion could similarly offer insights into the direction of the invasion front to bolster rapid response measures.

Acknowledgements

The authors wish to thank Amy Benson from the U.S. Geological Survey for her role in entering data for round goby, northern snakehead, and silver carp into the NAS database. We also thank Dr. Noel Burkhead and one other anonymous reviewer from USGS for helpful comments on this manuscript. Support for this project was provided by a grant from the Mellon Foundation and support for Environmental Studies by Robert Schumann.

References

- Araújo MB, New M (2007) Ensemble forecasting of species distributions. *Trends Ecology and Evolution* 22(1): 42–47, <http://dx.doi.org/10.1016/j.tree.2006.09.010>
- Anderson RP, Lew D, Peterson AT (2003) Evaluating predictive models of species' distributions: criteria for selecting optimal models. *Ecological Modelling* 162: 211–232
- Atkinson CE (1977) People's Republic of China, In Brown EE (ed), *World Fish Farming: Cultivation and Economics*: Westport, Connecticut, AVI Publishing Co., pp 321–344
- Brammeier J, Polls I, Mackey S (2008) Preliminary feasibility of ecological separation of the Mississippi River and the Great Lakes to prevent the transfer of aquatic invasive species. Great Lakes Fishery Commission and Great Lakes Fishery Trust, Ann Arbor and Lansing
- Carman SM, Janssen J, Jude DJ, Berg MB (2006) Diel interactions between prey behaviour and feeding in an invasive fish, the Round Goby, in a North American river. *Freshwater Biology* 51: 742–755, <http://dx.doi.org/10.1111/j.1365-2427.2006.01527.x>
- Carpenter G, Gillison AN, Winter J (1993) DOMAIN: a flexible modelling procedure for mapping potential distributions of plants and animals. *Biological Conservation* 2: 667–680
- Charlebois PM, Marsden JE, Goettl RG, Wolfe RK, Jude DJ, Rudnicka S (1997) The round goby, *Neogobius melanostomus* (Pallas), a review of European and North American literature. Illinois-Indiana Sea Grant Program and Illinois Natural History Survey. INHS Special Publication No. 20
- Charlebois PM, Corkum LD, Jude DJ, Knight C (2001) The Round Goby (*Neogobius melanostomus*) invasion: current research and future needs. *Journal of Great Lakes Research* 27(3):263–266, [http://dx.doi.org/10.1016/S0380-1330\(01\)70641-7](http://dx.doi.org/10.1016/S0380-1330(01)70641-7)
- Chen P, Wiley EO, Mcynset KM (2007) Ecological niche modeling as a predictive tool: Silver Carp and bighead carps in North America. *Biological Invasions* 9: 43–51, <http://dx.doi.org/10.1007/s10530-006-9004-x>
- Courtenay WR, Williams JD (2004) Snakeheads (Pisces, Channidae) - a biological synopsis and risk assessment. United States Geological Survey Circular 1251, 143 pp (Available from: http://fl.biology.usgs.gov/Snakehead_circ_1251/circ_1251_courtenay.pdf)
- Cooke SL, Hill WR, Meyer KP (2009) Feeding at different plankton densities alters invasive bighead carp (*Hypophthalmichthys nobilis*) growth and zooplankton species composition. *Hydrobiologia* 625: 185–193, <http://dx.doi.org/10.1007/s10750-009-9707-y>
- Cremer MC, Smitherman RO (1890) Food and growth of silver and bighead carp in cages and ponds. *Aquaculture* 20: 57–64, [http://dx.doi.org/10.1016/0044-8486\(80\)90061-7](http://dx.doi.org/10.1016/0044-8486(80)90061-7)
- DeGrandchamp KL (2006) Habitat selection and movement of bighead carp and Silver Carp in the lower Illinois River. Master's Thesis. Southern Illinois University, 57 pp
- Diniz-Filho JAF, Bini LM, Rangel TF, Loyola RD, Hof C, Bravo DN, Araújo MB (2009) Partitioning and mapping uncertainties in ensembles of forecasts of species turnover under climate change. *Ecography* 32: 897–906, <http://dx.doi.org/10.1111/j.1600-0587.2009.06196.x>
- Drake JM, Bossenbroek JM (2004) The potential distribution of Zebra Mussels in the United States. *Bioscience* 54(10): 931–941, [http://dx.doi.org/10.1641/0006-3568\(2004\)054\[0931:TPD OZM\] 2.0.CO;2](http://dx.doi.org/10.1641/0006-3568(2004)054[0931:TPD OZM] 2.0.CO;2)
- Dukravets GM (1992) The Amur snakehead, *Channa argus warpachowskii*, in the Talas and Chu River drainages: *Journal of Ichthyology* 31(5):147–151, <http://dx.doi.org/10.1111/j.2006.0906-7590.04596.x>
- Dukravets GM, Machulin AI (1978) The morphology and ecology of the Amur snakehead, *Ophiocephalus argus warpachowskii*, acclimatized in the Syr Dar'ya basin: *Journal of Ichthyology* 18(2): 203–208
- Elith J, Graham CH, Anderson RP, Dudík M, Ferrier S, Guisan A, Hijmans RJ, Huettmann F, Leathwick J R, Lehmann A, Li J, Lohmann LG, Loisell BA, Manion G, Moritz C, Nakamura M, Nakazawa Y, Mc Overton J, Peterson AT, Phillips SJ, Richardson K, Scachetti-Pereira R, Shapire RE, Soberón J, Williams S, Wisz MS, Zimmermann NE (2006) Novel methods improve species' distributions from occurrence data. *Ecography* 29: 129–151, <http://dx.doi.org/10.1111/j.2006.0906-7590.04596.x>
- Elith J, Kearney M, Phillips SJ (2010) The art of modeling range-shifting species. *Methods in Ecology and Evolution* 1: 330–342, <http://dx.doi.org/10.1111/j.2041-210X.2010.00036.x>
- Eng J (2006) ROC analysis: web-based calculator for ROC curves. Baltimore: Johns Hopkins University, <http://www.jrocfits.org> (Accessed April 1, 2010)
- Earth Systems Research Institute (ESRI) (2008) ArcMap version 9.3. Redlands, CA
- Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environment Conservation* 24(1): 38–49, <http://dx.doi.org/10.1017/S037689299700088>
- French JRP, Jude DJ (2001) Diets and diet overlap of nonindigenous gobies and small benthic native fishes co-inhabiting the St Clair River, Michigan. *Journal of Great Lakes Research* 27: 300–311, [http://dx.doi.org/10.1016/S0380-1330\(01\)70645-4](http://dx.doi.org/10.1016/S0380-1330(01)70645-4)
- Grinnell J (1917) Field tests of theories concerning distributional control. *American Naturalist* 51:115–128, <http://dx.doi.org/10.1086/279591>
- Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology. *Ecological Modelling* 135: 147–186, [http://dx.doi.org/10.1016/S0304-3800\(00\)00354-9](http://dx.doi.org/10.1016/S0304-3800(00)00354-9)
- Guisan A, Thuiller W (2005) Predicting species distribution: offering more than simple habitat models. *Ecology Letters* 8: 993–1009, <http://dx.doi.org/10.1111/j.1461-0248.2005.00792.x>
- Guo Q, Kelley M, Graham CH (2005) Support vector machines for predicting distribution of sudden oak death in California. *Ecological Modelling* 182: 75–90, <http://dx.doi.org/10.1016/j.ecolmodel.2004.07.012>
- Guo Q, Liu Y (2010) ModEco: An integrated software package for ecological niche modeling. *Ecography* 33: 1–6, <http://dx.doi.org/10.1111/j.1600-0587.2010.06416.x>
- Hansen MJ (2010) The Asian carp threat to the Great Lakes. *Great Lakes Fishery Commission*, 8 pp
- Herborg L, Mandrak NE, Cudmore BC, MacIsaac HJ (2007) Comparative distribution and invasion risk of snakehead (Channidae) and Asian carp (Cyprinidae) species in North America. *Canadian Journal of Fisheries and Aquatic Science* 64: 1723–1735, <http://dx.doi.org/10.1139/f07-130>
- Jarnevich CS, Reynolds LV (2011) Challenges of predicting the potential distribution of a slow-spreading invader: a habitat suitability map for an invasive riparian tree. *Biological Invasions* 13: 153–163, <http://dx.doi.org/10.1007/s10530-010-9798-4>
- Jude RH, Smith GR (1992) Establishment of Gobiidae in the Great Lakes basin. *Canadian Journal of Fisheries and Aquatic Science* 49: 416–421, <http://dx.doi.org/10.1139/f92-047>
- Jude DJ, DeBoe S (1996) Possible impact of gobies and other introduced species on habitat restoration efforts. *Canadian Journal Fisheries and Aquatic Science* 53 (Suppl. 1): 136–141, <http://dx.doi.org/10.1139/f96-001>
- Keller RP, Lodge DM (2007) Species invasions from commerce in live aquatic organisms: problems and possible solutions. *Bioscience* 57(5): 428–436, <http://dx.doi.org/10.1641/B570509>

- Kelly M, Guo Q, Liu D, Shaari D (2007) Modeling the risk for a new invasive forest disease in the United States: An evaluation of five environmental niche models. *Computers, Environment and Urban Systems* 31: 689–710, <http://dx.doi.org/10.1016/j.compenvurbsys.2006.10.002>
- Kolar CS, Lodge DM (2002) Ecological Predictions and Risk Assessment for Alien Fishes in North America. *Science* 298(5596): 1233–1236, <http://dx.doi.org/10.1126/science.1075753>
- Kolar C, Courtenay WR Jr, Chapman D, Housel CM, Williams JD, Jennings DP (2005) Asian carps of the genus *Hypophthalmichthys* (Pices, Cyprinidae) – a biological synopsis and risk assessment. United States Geological Survey Report to U.S. Fish and Wildlife Service, 175 pp
- Lee CE (1999) Rapid and repeated invasions of fresh water by the saltwater copepod *Eurytemora affinis*. *Evolution* 53: 1423–1434, <http://dx.doi.org/10.2307/2640889>
- Lindsay J (2009) Whitebox Geospatial Analysis Tools v. 1.0.5. University of Guelph, Canada
- Liu CR, Berry PM, Dawson TP, Pearson RG (2005) Selecting thresholds of occurrence in the prediction of species distributions. *Ecography* 28(3): 385–393, <http://dx.doi.org/10.1111/j.0906-7590.2005.03957.x>
- MacInnis AJ, Corkum LD (2000) Fecundity and reproductive season of the Round Goby *Neogobius melanostomus* in the upper Detroit River. *Transactions of the American Fisheries Society* 129: 136–144, [http://dx.doi.org/10.1577/1548-8659\(2000\)129<0136:FARSOT>2.0.CO;2](http://dx.doi.org/10.1577/1548-8659(2000)129<0136:FARSOT>2.0.CO;2)
- Mandrak NE, Cudmore B (2004) Risk assessment for Asian carps in Canada. Canadian Science advisory Secretariat. Research Document 2001/103
- Marmion M, Parvainien M, Luoto M, Heikkinen RK, Thuiller W (2009) Evaluation of consensus methods in predictive species distribution modeling. *Diversity and Distributions* 15: 59–69, <http://dx.doi.org/10.1111/j.1472-4642.2008.00491.x>
- Moore ID, Burch GJ (1986) Sediment Transport Capacity of Sheet and Rill Flow: Application of Unit Stream Power Theory. *Water Resources Research* 22(8): 1350–1360
- Moskal'kova KI (1996) Ecological and morphophysiological prerequisites to range extension in the round goby *Neogobius melanostomus* under conditions of anthropogenic pollution. *Voprosy Ikhtologii* 36: 584–590
- Moles AT, Gruber MAM, Bonser SP (2008) A new framework for predicting invasive plant species. *Journal of Ecology* 96: 13–17
- Munoz MED, de Giovannir, de Siqueira MF, Sutton T, Brewer P, Pereira RS, Canhos DAL, Canhos VP (2011) OpenModeller: a generic approach to species' potential distribution modelling. *Geoinformatica* 15: 111–135, <http://dx.doi.org/10.1007/s10707-009-0090-7>
- NAS (2011) Nonindigenous Aquatic Species Database (NAS), <http://nas.er.usgs.gov> (Accessed on 15 February 2010)
- National Invasive Species Council (NISC) (2008) 2008-2012 National invasive species management plan. Department of the Interior: Washington (DC). NISC, 646 pp
- Odenkirk J, Owens S (2007) Northern Snakeheads in the Tidal Potomac River System. *Transactions of the American Fisheries Society* 134(6): 1605–1609, <http://dx.doi.org/10.1577/T05-085.1>
- Okada Y (1960) Studies of the freshwater fishes of Japan, II, Special part.—*Journal of the Faculty of Fisheries Prefectural University of Mie* 4(3): 1–860
- Opusynski K, Lirski A, Myszowski L, Wolnicki J (1989) Upper lethal and rearing temperatures for juvenile common carp, *Cyprinus carpio* L. and Silver Carp, *Hypophthalmichthys molitrix* (Valenciennes). *Aquaculture and Fish Management* 20: 287–294
- Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modelling of species geographic distributions. *Ecological Modelling* 190(3–4): 231–259, <http://dx.doi.org/10.1016/j.ecolmodel.2005.03.026>
- Phillips SJ, Dudik M (2008) Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation. *Ecography* 31(2): 161–175, <http://dx.doi.org/10.1111/j.0906-7590.2008.5203.x>
- Pimentel D, Zuniga R, Morrison D (2005) Update on the economic costs associated with alien-invasive species in the United States. *Ecological Economics* 52:273–288, <http://dx.doi.org/10.1016/j.ecolecon.2004.0.002>
- Rasmussen JL (2002) The Cal-Sag and Chicago sanitary and ship canal: a perspective on the spread and control of selected aquatic nuisance fish species. U. S. Fish and wildlife Service, Region 3 White Paper, Fisheries Resources Office, Rock Island, Illinois, 26 pp
- Ray WJ, Corkum LD (2001) Habitat and site affinity of the Round Goby. *Journal of Great Lakes Research* 27(3): 329–334, [http://dx.doi.org/10.1016/S0380-1330\(01\)70648-X](http://dx.doi.org/10.1016/S0380-1330(01)70648-X)
- Rejmánek M (2000) Invasive plants: approaches and predictions. *Austral Ecology* 25: 497–506, <http://dx.doi.org/10.1046/j.1442-9993.2000.01080.x>
- Ricciardi A (2001) Facilitative interactions among aquatic invaders: Is an 'invasional meltdown' occurring in the Great Lakes? *Canadian Journal of Fisheries and Aquatic Science* 58: 2513–2525, <http://dx.doi.org/10.1139/f01-178>
- Ricciardi A, Rasmussen JB (1998) Predicting the identity and impact of future biological invaders: a priority for aquatic resource management. *Canadian Journal of Fisheries and Aquatic Science* 55: 1759–1765, <http://dx.doi.org/10.1139/f98-066>
- Ricciardi A, MacIsaac HJ (2000) Recent mass invasion of the North American Great Lakes by Ponto-Caspian species. *Trends in Ecology and Evolution* 15: 62–65, [http://dx.doi.org/10.1016/S0169-5347\(99\)01745-0](http://dx.doi.org/10.1016/S0169-5347(99)01745-0)
- Roura-Pascual N, Brotons L, Peterson AT, Thuiller W (2009) Consensual predictions of potential distributional areas for invasive species: A case study of Argentine ants in the Iberian Peninsula. *Biological Invasions* 11(4):1017–1031, <http://dx.doi.org/10.1007/s10530-008-9313-3>
- Sakai AK (2001) The population biology of invasive species. *Annual Review of Ecological Systems* 32: 305–332, <http://dx.doi.org/10.1146/annurev.ecolsys.32.081501.114037>
- Schofield PJ, Williams JD, Nico LG, Fuller P, Thomas MR (2005) Foreign Nonindigenous Carp and Minnows (Cyprinidae) in the United States – A Guide to Their Identification, Distribution and Biology. U.S. Geological Survey Scientific Investigations Report 2005-5041, 103 pp. Available from: http://fl.biology.usgs.gov/sir2005_5041_report_LowRez.pdf
- Scholkopf B, Platt JC, Shawe-Taylor J, Smola AJ, Williamson RC (2001) Estimating the support of a high-dimensional distribution. *Neural Computation* 13: 1443–1471, <http://dx.doi.org/10.1162/089976601750264965>
- Sifa L, Senlin X (1995) Culture and capture of fish in Chinese reservoirs: Ottawa, Canada, International Development Research Centre, 128 pp
- Simberloff D (2003) How much information on population biology is needed to manage introduced species? *Conservation Biology* 17: 83–92, <http://dx.doi.org/10.1046/j.1523-1739.2003.02028.x>
- Stohlgren TJ, Ma P, Kumar S, Rocca M, Morisette JT, Jarnevich CS, Benson N (2010) Ensemble habitat mapping of invasive plant species. *Risk Analysis* 30(2): 224–235, <http://dx.doi.org/10.1111/j.1539-6924.2009.01343.x>
- Stockwell DRB, Noble IR (1992) Induction of sets of rules from animal distribution data: a robust and informative method of data analysis. *Mathematics and Computers in Simulation* 33:385–390, [http://dx.doi.org/10.1016/0378-4754\(92\)90126-2](http://dx.doi.org/10.1016/0378-4754(92)90126-2)

- Stockwell DRB, Peters D (1999) The GARP modelling system: problems and solutions to automated spatial prediction. *International Journal of Geographical Information Science* 13:143–158, <http://dx.doi.org/10.1080/136588199241391>
- Stokstad E (2007) Feared quagga mussel turns up in the western United States. *Science* 315(5811): 453, <http://dx.doi.org/10.1126/science.315.5811.453>
- Swets JA (1988) Measuring the accuracy of diagnostic systems. *Science* 240(3): 1285–1293, <http://dx.doi.org/10.1126/science.3287615>
- USGS (2003) Streams and water bodies of the United States. Available online at: <http://nationalatlas.gov/atlasftp.html>
- Vander Zanden MJ, Olden JD (2008) A management framework for preventing the secondary spread of aquatic invasive species. *Canadian Journal of Fisheries and Aquatic Science* 65(7): 1512–1522, <http://dx.doi.org/10.1139/F08-099>
- Weimer MT, Keppner SM (2000) The Round Goby (*Apollonia melanostomus*) in Lake Erie. *Great Lakes Research Review* 5:19–24
- Wilcove DS, Rothstein D, Dubow J, Phillips S, Losos E (1998) Quantifying threats to imperiled species in the United States. *BioScience* 448: 607–615, <http://dx.doi.org/10.2307/1313420>
- Williamson CJ, Garvey JE (2005) Growth, fecundity, and diets of newly established Silver Carp in the middle Mississippi River. *Trans American Fisheries Society* 134: 1423–1430, <http://dx.doi.org/10.1577/T04-106.1>