

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29

Article type : Article

Suitability of Laurentian Great Lakes for invasive species based on global species distribution models and local habitat

Andrew M. Kramer^{1,†}, Gust Annis², Marion E. Wittmann^{3,4}, William L. Chadderton², Edward S. Rutherford⁵, David M. Lodge^{6,7}, Lacey Mason⁸, Dmitry Beletsky⁹, Catherine Riseng⁸, and John M. Drake¹

¹Odum School of Ecology, University of Georgia, Athens, Georgia 30602 USA

²The Nature Conservancy, Lansing, Michigan 48906 USA

³Department of Biological Sciences, University of Notre Dame, Notre Dame, Indiana 46556 USA

⁴ Natural Reserve System, University of California, Santa Barbara, California 93106 USA

⁵NOAA Great Lakes Environmental Research Laboratory, Ann Arbor, Michigan 48108-9719 USA

⁶Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, New York 14853 USA

⁷Environmental Change Initiative, University of Notre Dame, Notre Dame, Indiana 46556 USA

⁸School of Natural Resources and Environment, University of Michigan, Ann Arbor, Michigan 48109-1041 USA

⁹Cooperative Institute for Limnology and Ecosystems Research, School of Natural Resources and Environment University of Michigan, Ann Arbor, Michigan 48108 USA

Received 21 April 2017; accepted 25 May 2017. Corresponding Editor: Wyatt Cross.

†E-mail: kramera3@uga.edu

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1002/ecs2.1883](https://doi.org/10.1002/ecs2.1883)

This article is protected by copyright. All rights reserved

30 **Abstract**

31 Efficient management and prevention of species invasions requires accurate prediction of where
32 species of concern can arrive and persist. Species distribution models provide one way to
33 identify potentially suitable habitat by developing the relationship between climatic variables and
34 species occurrence data. However, these models when applied to freshwater invasions are
35 complicated by two factors. The first is that the range expansions that typically occur as part of
36 the invasion process violate standard species distribution model assumptions of data stationarity.
37 Second, predicting potential range of freshwater aquatic species is complicated by the reliance on
38 terrestrial climate measurements to develop occurrence relationships for species that occur in
39 aquatic environments. To overcome these obstacles, we combined a recently developed
40 algorithm for species distribution modeling—range bagging—with newly available aquatic
41 habitat-specific information from the North American Great Lakes region to predict suitable
42 habitat for three potential invasive species: golden mussel, killer shrimp and Northern snakehead.
43 Range bagging may more accurately predict relative suitability than other methods because it
44 focuses on the limits of the species environmental tolerances rather than central tendency or
45 “typical” cases. Overlaying the species distribution model output with aquatic habitat-specific
46 data then allowed for more specific predictions of areas with high suitability. Our results
47 indicate there is suitable habitat for Northern snakehead in the Great Lakes, particularly shallow
48 coastal habitats in the lower four Great Lakes where literature suggests they will favor areas of
49 wetland and submerged aquatic vegetation. These coastal areas also offer the highest suitability
50 for golden mussel, but our models suggest they are marginal habitats. Globally, the Great Lakes
51 provide the closest match to the currently invaded range of killer shrimp, but they appear to pose
52 an intermediate risk to the region. Range bagging provided reliable predictions when assessed
53 either by a standard test set or tests for spatial transferability, with golden mussel being the most
54 difficult to accurately predict. Our approach illustrates the strength of combining multiple
55 sources of data, while reiterating the need for increased measurement of freshwater habitat at
56 high spatial resolutions to improve the ability to predict potential invasive species.

57 **Key words:** environmental niche; golden mussel; habitat suitability; killer shrimp;
58 nonindigenous species; northern snakehead.

59

60 **Introduction**

61 Aquatic invasive species (AIS) have imposed substantial ecological damage on freshwater
62 ecosystems (Ricciardi and MacIsaac 2000, Cucherousset and Olden 2011), prompting a more
63 proactive, holistic approach to invasive species management (Leung et al. 2002, Pagnucco et al.
64 2014). The identification of high risk species, transport pathways (e.g. Ricciardi and Rasmussen
65 1998, Keller et al. 2009, Gantz et al. 2015), surveillance sites, and opportunities for
66 implementation of slow-the-spread strategies all rely upon an accurate prediction of locations
67 suitable for nonindigenous species to establish and persist (Gormley et al. 2011, Jiménez-
68 Valverde et al. 2011, Václavík et al. 2012, Gallien et al. 2012).

69
70 Species distribution models (SDMs) estimate the statistical relationship between species
71 occurrence and environmental conditions (Elith and Leathwick 2009) and applications of these
72 models have been used to identify suitable habitat outside of the current range (Barve et al.
73 2011), predict range shifts in response to climate change (Austin and Van Niel 2011, VanDerWal
74 et al. 2013) and predict the spread of invasive species (Kulhanek et al. 2011). However, these
75 applications involve violation of a key assumption of SDM methods, namely stationarity in
76 species occurrence (Barve et al. 2011, Václavík and Meentemeyer 2012, Pagel and Schurr 2012).
77 A spreading species is not in equilibrium with environmental conditions and has not had previous
78 opportunity to sample novel environments in which it may thrive. The result is limited
79 information on the suitability of novel combinations of environmental conditions (Veloz et al.
80 2012). For actively invading species, using data from the invaded and native ranges provides the
81 strongest basis for extrapolation and is the most practical approach given physiological
82 information needed for a mechanistic model is generally lacking (Araújo and Peterson 2012).
83 Nevertheless, there is still a risk of underestimating the extent of suitable habitat.

84
85 Predicting suitable habitat for freshwater aquatic species using SDMs faces an additional
86 challenge. SDMs are typically fit to measured or interpolated climatic or environmental
87 measurements of the terrestrial environment. This is because the most reliably available climate
88 variables come from global data sets (Hijmans et al. 2005, Tyberghein et al. 2012) or remote
89 sensing (Cord and Rödder 2011, Bisrat et al. 2012) that primarily measure terrestrial climate
90 variables. These data are often only an indirect indicator of freshwater environmental conditions,
91 and several physical characteristics of water bodies can strongly influence the correlation

92 between atmospheric and aquatic conditions (Mohseni and Stefan 1999). For example, bottom
93 waters of deep, seasonally stratified lakes are to a large extent decoupled from lake surface (and
94 adjacent atmosphere) for a substantial portion of the year due to the presence of a thermocline
95 (Boyce et al. 1989). Moreover, aquatic species predictions are also sensitive to aquatic
96 environment-specific conditions such as hydrological and substrate stability, wave action or flow,
97 water chemistry or clarity (Leathwick et al. 2005, 2011, Snelder et al. 2006, Brenden et al. 2008,
98 McKenna Jr. and Castiglione 2010).

99
100 Identifying which nonindigenous species are most likely to be introduced, establish, and to result
101 in negative impacts is a complex endeavor that complicates management of biological invasions.
102 This not only requires knowledge about the pathways of introduction and propagule pressure, but
103 also the suitability of the receiving environment and likelihood of establishment (Leung et al.
104 2012). One reliable predictor that a non-native species will cause environmental damage is
105 invasiveness in other locations (Kolar and Lodge 2001). In addition, species that have high rates
106 of introduction or are established in adjacent regions are also more likely to become established
107 (Lockwood et al. 2005). For some species with known invasion histories and observations of
108 widespread impact, information may exist on life history, environmental tolerance and
109 persistence outside the native range. The ability to assess invasion risk can be increased by
110 combining SDMs with this existing information, even when wide-spread measurements are not
111 available to create mechanistic predictions from laboratory-measured environmental tolerances.

112
113 We performed a two part evaluation of the potential range of three nonindigenous species which
114 could invade the North American Great Lakes: golden mussel (*Limnoperna fortunei*), killer
115 shrimp (*Dikerogammarus villosus*) and Northern Snakehead (*Channa argus*). These species have
116 been identified by the United States Aquatic Nuisance Species Task Force (USACE 2011) as
117 species of concern to US waters and have been predicted as probable future invaders of the Great
118 Lakes (Ricciardi and Rasmussen 1998, Council of Great Lakes Governors 2013). First we
119 estimated the habitat suitability throughout the Great Lakes for each species using a novel SDM
120 algorithm called "range bagging" (Drake 2015). This algorithm was designed to estimate species
121 range limits based on climate variables measured at all precisely known occurrences world-wide.
122 To assess ecosystem-specific suitable habitat, these model predictions were then merged with

123 high resolution spatially explicit data representing the localized aquatic environment in terms of
124 established species tolerances (Wittmann et al. 2017). The results provide species range
125 predictions for a set of molluscan, crustacean and vertebrate nonindigenous aquatic species,
126 illustrating the value of the two-part evaluation and showing that inclusion of measurements of
127 aquatic conditions (or more reliable terrestrial surrogates) is key to providing models for the
128 potential ranges of invasive aquatic species that are most relevant to management decisions (US
129 Environmental Protection Agency 2008, Kilroy et al. 2008).

130

131 **Methods**

132 Species

133 Three species were chosen for this analysis based on their likelihood of invasion and concern
134 from scientists and the regional management agencies. The golden mussel (*Limnoperna fortunei*;
135 Dunker, 1857) is an epifaunal bivalve, native to mainland China. Since the mid 1960s, golden
136 mussel has been unintentionally dispersed across the globe through fouling of shipping vessels
137 and established populations are now present in Hong Kong, Taiwan, Japan, Brazil, Paraguay,
138 Uruguay, Bolivia and Argentina (Ricciardi 1998). Spawning occurs at temperatures between 16
139 and 28°C (Xu et al. 2013) and after external fertilization, a free-living straight-hinged larva
140 develops, subsequently evolving into an actively swimming veliger larva capable of survival in a
141 wide variety of habitats (Karatayev et al. 2007). Golden mussel are thought to have life histories
142 and habitat preferences similar to Dreissenid mussels, which have a widespread distribution and
143 ecological impacts in the Great Lakes watershed (Karatayev et al. 2007, Fahnenstiel et al. 2010,
144 Kerfoot et al. 2010, Vanderploeg et al. 2010).

145

146 Killer shrimp (*Dikerogammarus villosus*) is an amphipod native to the Ponto-Caspian region of
147 Eastern Europe and Ukraine. Since the 1980s, it has spread to at least 17 European countries
148 along the complex European canal-river systems (Pöckl 2009) and to the United Kingdom via
149 maritime shipping (Gallardo and Aldridge 2012). Killer shrimp is expected to continue its spread
150 in Europe and eventually to North America (Ricciardi and Rasmussen 1998). Many studies show
151 that *D. villosus* can prey upon many macroinvertebrate species, including native and other
152 nonnative amphipods, as well as fish eggs and larvae (Dick and Platvoet 2000, Dick et al. 2002,
153 Casellato et al. 2006). *D. villosus* range expansion is influenced by hydrological regime,

154 temperature, salinity, water quality, substrate and food availability (Bruijs et al. 2001, Devin et
155 al. 2003, Josens et al. 2005, MacNeil et al. 2010, Boets et al. 2010).

156

157 Northern snakehead (*Channa argus*) is a fish native to China, Russia and Korea (Courtenay and
158 Williams 2004). It is established in regions of the eastern and central US (Potomac River,
159 Chesapeake Bay, Lower Mississippi River (Arkansas) and individuals have been observed in
160 California, New Jersey, New York and in watersheds adjacent to the Great Lakes
161 (<http://nas.er.usgs.gov>). This species is capable of survival in poorly oxygenated waters and has
162 been found inhabiting shallow (<2 m) ponds or swamps, canals, reservoirs, lakes, and rivers
163 (Courtenay and Williams 2004). The presence of submersed aquatic vegetation can provide a
164 benefit for its reproduction, but is not necessary. *C. argus* has a wider latitudinal range and
165 temperature tolerance (0 to >30°C) compared with other snakehead species (Courtenay and
166 Williams 2004) and earlier SDM models using global climatic data suggest much of North
167 America is suitable (Herborg et al. 2007). It is an aggressive predator, with a broad diet including
168 fish, invertebrate and amphibian species (Courtenay and Williams 2004). Northern snakehead is
169 listed as injurious wildlife under the US Lacey Act (Lacey Act 1900).

170

171 Occurrence data

172 Worldwide occurrence records for each study species were obtained from the primary literature,
173 the Global Biodiversity Information Facility (<http://www.gbif.org/>), Fishbase (Froese and Pauly
174 2011), and USGS (<http://nas.er.usgs.gov>) databases. All records with position uncertainty >50
175 km were removed to ensure accurate match with the environmental data. All presence locations
176 in both the native and introduced ranges were included in this analysis, yielding 81 unique
177 locations for *L. fortunei* (22 native and 59 introduced; Appendix S1: Fig. S1), 233 unique
178 locations for *D. villosus* (16 native and 217 introduced; Appendix S1: Fig. S2), and 198 unique
179 locations for *C. argus* (47 native and 151 introduced; Appendix S1: Fig. S3).

180

181 Environmental data

182 Nineteen global climate surfaces were obtained at 5 min resolution from the WorldClim dataset
183 (Hijmans et al. 2005, <http://www.worldclim.org>). These variables are derivations of monthly
184 rainfall and climate data that have been interpolated from weather stations, and are designed to

185 have increased relevance to species physiological limits (Hijmans et al. 2005, Graham and
186 Hijmans 2006). To reduce bias and improve model stability, climate variables were rescaled and,
187 in some cases, transformed. Variables BIO1-BIO11 and BIO15 were rescaled by subtracting the
188 global mean and dividing by the global standard deviation. Variables BIO18 and BIO19 were
189 rescaled similarly after log transformation to improve symmetry. Variables BIO12-BIO14,
190 BIO16 and BIO17 were transformed via the empirical cumulative distribution function to obtain
191 a uniform distribution as a normal distribution could not be approximated. (See archived R code
192 for further details.) All climate variables were included in the analysis in order to estimate the
193 best predictive model given the limited a priori understanding of the correlations between
194 atmospheric and aquatic climate and of which variables are most influential for each species.

195
196 Data on the distribution of submerged aquatic vegetation (SAV) at 30 m resolution in the
197 optically shallow areas of lakes Huron, Michigan, Erie, and Ontario were obtained from the
198 Michigan Tech Research Institute (MTRI). These data were generated using an MTRI-
199 developed depth invariant algorithm and depend on Landsat satellite data collected during the
200 vegetative growing season (Michigan Tech Research Institute 2012, Shuchman et al. 2013).
201 Vegetative growing season and years varied by lake: Lake Erie, May–September 2006–2011;
202 Lake Huron, March–September 2007–2011; Lake Michigan, April–May 2008–2011; Lake
203 Ontario, April–September 2008–2011. Some portions of these lakes could not be classified due
204 to high turbidity. Submersed vegetation data did not exist for Lake Superior or Lake St. Clair.
205 Data for the remaining four lakes were combined using the Mosaic to New Raster tool process in
206 ArcGIS Version 10.2 (Esri 2014) with a cell size of 30 m. This study used classes 1 (light
207 submerged aquatic vegetation) and 7 (dense submerged aquatic vegetation).

208
209 Data on the spatial distribution of wetlands were compiled by the Great Lakes Coastal Wetland
210 Inventory (Great Lakes Coastal Wetland Consortium 2004). The inventory utilized the most
211 comprehensive coastal wetlands data available for the Great Lakes and connecting channels and
212 was derived from multiple sources. Coastal wetlands polygonal data were rasterized with a cell
213 size of 30 m.

214
215 Benthic temperature data acquired from the NOAA Great Lakes Coastal Forecasting System

216 (GLCFS) nowcasts were summarized by the Great Lakes Aquatic Habitat Framework (GLAHF).
217 The GLCFS uses a 3D hydrodynamic model (Schwab and Bedford 1994, Chu et al. 2011,
218 Beletsky et al. 2013) with a horizontal resolution ranging from 2 km (Lakes Erie, Huron and
219 Michigan) to 5km (Lake Ontario) and 10 km (Lake Superior) to nowcast lake temperatures
220 (among other physical variables) at 20 vertical levels in all lakes except Lake Erie (that has 21
221 level). Averages of August monthly data from the years 2006-2012 were used and were
222 combined for individual lakes using a mosaic process with an output cell size of 2000 meters
223 (Esri 2014).

224

225 Modeling potential distribution

226 To determine the habitat suitability in novel environments we applied SDMs using the range
227 bagging algorithm (Drake 2015). This algorithm estimates species range limits in multi-
228 dimensional climate space using bootstrap aggregation. Range bagging has comparable accuracy
229 to the widely used MaxEnt approach on high quality validation data sets (Drake 2015) and also
230 performs similarly to widely used species distribution models for traditional approaches (Drake
231 and Richards unpublished data) and invasive species (Wittmann et al. 2017, Cope et al. in
232 review) including the three species considered here (Kramer et al. unpublished data). Range
233 bagging was chosen over some more commonly used algorithms (e.g. Maxent) for two primary
234 reasons. First, range bagging estimates the environmental limits of species habitat, giving it an
235 ecologically relevant interpretation (Drake 2015, Cope et al. in review). This quantity more
236 closely matches the concept of the ecological niche advanced by Hutchinson (Hutchinson 1957).
237 Consideration of environmental tolerances—rather than the central tendency—may offer a more
238 conservative (i.e. larger) estimate of the ecological niche relevant to invasive species risk
239 assessment. Further, range bagging uses only presence points, removing the need for selecting a
240 suitable area from which to sample background points, as this choice has recently been shown to
241 have substantial effects on model reliability (Barve et al. 2011, Kramer et al unpublished data).

242

243 Range bagging models were constructed by fitting convex hulls to 256 random combinations of
244 two environmental variables from the global climate surfaces (Drake 2015). We specified the
245 parameter identifying the proportion of points sampled in each bootstrapped combination to be 1.
246 While smaller subsets may give better performance on the holdout test points if presences are

247 sampled from sink habitat (Drake 2015), such limitations could be restrictive relative to novel
248 environmental conditions experienced by non-native species. Along with all approaches to
249 species distribution modeling, range bagging models may still be biased due to the issues of non-
250 stationarity and novel environmental combinations. For validation, therefore, models were fit to
251 80% of the data and performance (area under the receiver operating curve, or AUC) was
252 evaluated on a 20% hold-out test set (Appendix S1: Fig S4). For golden mussel, a 60-40 training
253 test split was used due to fewer data points. The models fit to the training data are presented in
254 the maps, such that performance on the 20% (or 40% for golden mussel) hold-out represents
255 provides the measure of accuracy of the displayed models. As a second measure of performance,
256 we estimated the continuous Boyce index (Hirzel et al. 2006, Petitpierre et al. 2012). The Boyce
257 index was designed for presence-only data and was calculated with the R package “ecospat”. The
258 Boyce index varies from -1 to 1 with values greater than zero indicating agreement between the
259 prediction and the presences in the test data (Hirzel et al. 2006). The AUC and the Boyce index
260 were calculated with the same model output and data for each run of the model. Variance in
261 model performance was assessed with 10-fold cross-validation on the training data (Appendix
262 S1: Fig. S4). Additionally, we estimated the transferability of the model by performing 5-fold
263 cross-validation on data that was divided into longitudinal bins (Wenger and Olden 2012,
264 Appendix S1: Fig. S4). This test measures the ability of the model to predict occurrence in
265 distinct geographical areas, with longitudinal bins being appropriate for the occurrence of these
266 species on multiple continents. Code used to fit models and display results is archived on Dryad
267 (<http://dx.doi.org/10.5061/dryad.d4144>).

268

269 The outcome of a range bagging model is an estimate of niche centrality for each species at each
270 point on a map. After fitting the model, niche centrality was estimated for each grid cell in the
271 global climate data set and mapped. Niche centrality refers to the tendency of an environment to
272 be centered within the environmental range of a species across multiple environmental variables.
273 We also estimated variable importance for each species by permuting each predictor variable and
274 measuring the reduction in accuracy on the withheld test set. Because data reflect only
275 occurrence records, a set of random background points was necessarily selected to function as
276 absence points in the AUC calculation (Elith and Leathwick 2009). These points were taken from
277 a large buffer (2000 km surrounding the known occurrence records), consistent with the fact that

278 these species are well known to be entrained in long distance transport pathways and therefore
279 not greatly limited by dispersal.

280

281 Aquatic habitat suitability

282 Species-specific information on environmental limits to occurrence, survival and spawning was
283 collected from extensive review of the primary literature, which included but was not limited to
284 field-based and laboratory experiments, surveillance efforts, and review and synthesis
285 publications. Published limits in any measured environmental variable, including temperature,
286 substrate type, water chemistry, and aquatic vegetation were recorded. Cases for which reliable
287 Great Lakes-wide data was unavailable were discarded (primarily water quality data and
288 zooplankton species composition). When sources differed in their reported limits the extrema
289 from the set of values were considered to be the limit. This resulted in surprisingly limited
290 concrete knowledge about environmental limits to establishment and persistence for the three
291 species considered here. These environmental conditions were then used to restrict the projected
292 range bagging model to the suitable areas of the Great Lakes, treating the environmental limits as
293 thresholds beyond which habitat was unsuitable for that aspect of species life-history.

294

295 As spawning temperature $>16^{\circ}\text{C}$ is required for golden mussel reproduction (Xu et al. 2013),
296 available benthic temperature data on the Great Lakes were used to delineate the potential
297 spawning habitat. Constraints based solely on depth were not included, although it is thought that
298 golden mussel may colonize the same habitats as Dreissenid mussels in the Great Lakes
299 (Ricciardi 1998). Further, the existing measured depth limits for golden mussel are constrained
300 by sampling and their current largely riverine distribution.

301

302 Published field and laboratory studies of killer shrimp did not identify any useful environmental
303 limitations that could be addressed with the Great Lakes dataset. The reported temperature range
304 for survival is large (Bruijs et al. 2001, Wijnhoven et al. 2003, Velde et al. 2009), as is substrate
305 usage (Boets et al. 2010) and dissolved oxygen (Gallardo and Aldridge 2012). Killer shrimp has
306 only been reported at shallow depths (Lods-Crozet and Reymond 2006), but as with golden
307 mussel this limit appears to depend on sampling, the limited depth range of colonized habitats
308 and other factors, rather than a well-understood biological limitation.

309

310 The minimum temperature requirement for Northern snakehead spawning is 18°C (Amanov
311 1974) and survival has been observed at temperatures 0–30°C (Okada 1960, Courtenay and
312 Williams 2004). *C. argus* are known to survive long periods in low oxygen environments (Frank
313 1970, Courtenay and Williams 2004) and prefer habitats with mud and aquatic vegetation (Okada
314 1960, Courtenay and Williams 2004). Thus, Great Lakes regions with wetlands and submerged
315 aquatic vegetation are of particular concern, even though the fish is able to occur outside of these
316 areas (Amanov 1974). To identify areas where population densities and ecological impacts may
317 be highest, the snakehead niche map was restricted by the combination of coastal wetlands and
318 submerged aquatic vegetation.

319

320 **Results**

321 Golden mussel

322 There was relatively little climate overlap between the Great Lakes basin and native and current
323 non-native golden mussel distribution (Fig. 1). Niche centrality was highest in Lake Erie, but
324 never exceeded 0.39, indicating most of the marginal niche models did not include
325 environmental conditions observed in the Great Lakes basin. At the global scale, several regions
326 where golden mussel is not established had much higher relative suitability, including the south-
327 east U.S. and eastern Australia (Fig. 2). Model AUC on a balanced set of withheld test points and
328 random background points was 0.89 (Appendix S1: Fig. S5) and the Boyce index was 0.78.
329 Random cross-validation showed a similar average AUC and slightly lower Boyce index
330 (Appendix S1: Fig. S6, Fig. S7). The spatial cross-validation for golden mussel had much weaker
331 performance, indicating the invaded and native ranges experience distinct climates (Appendix
332 S1: Fig. S6, Fig. S7). The most influential variable for model performance was seasonality of
333 precipitation (Appendix S1: Fig. S8), however, the loss of performance from permuting any
334 single climate value was small and of similar magnitude, indicating redundancy due to high
335 correlation among variables and/or a multivariate niche.

336

337 The suitable niche for golden mussel was further restricted when limitations on reproduction
338 were taken into account. Areas with average summer benthic temperatures warm enough to
339 facilitate spawning were limited to Lake Erie, southern Lake Michigan and shallow bays of all

340 lakes (Fig. 3). The range bagging model and water temperature data indicated that Lake Superior
341 and colder areas are unlikely to support viable populations.

342

343 Killer shrimp

344 Killer shrimp displayed variable, intermediate niche centrality across the Great Lakes basin, with
345 the highest match with current occurrences occurring in the southern regions (Fig. 4). There were
346 no areas fully within the climate envelope of known killer shrimp populations. Indeed, viewed at
347 the global scale, the climate niche only estimates high niche overlap for the native and currently
348 invaded areas in the Ponto-Caspian region and Europe (Fig. 5). Model AUC was high at 0.96
349 (Appendix S1: Fig. S9) with the most influential variables including the amount of precipitation
350 in the wettest month and the wettest quarter (Appendix S1: Fig. S10). The Boyce index was 0.56,
351 with comparable performance for both random and spatial cross-validation (Appendix S1: Fig.
352 S6, Fig. S7). Alternative models using only temperature covariates had lower (killer shrimp and
353 snakehead) or equivalent performance (golden mussel) for these species and did not provide
354 increased ability to differentiate importance among the correlated climate predictors. As
355 explained above, the few directly measured biotic limitations to killer shrimp persistence were so
356 broad as to include the entirety of the Great Lakes.

357

358 Northern snakehead

359 The snakehead niche model estimated high climate overlap in the Great Lakes basin. The
360 majority of Lake Erie, Lake Michigan and Lake Ontario, as well as large areas of Lakes Huron
361 and Superior had niche centrality exceeding 0.8 (Fig. 6). This was unsurprising given the high
362 niche centrality of much of eastern North America (consistent with the frequency of known
363 persistent populations in this region; Appendix S1: Fig. S8). The model predicted a large area of
364 Europe and part of South America as equally suitable environments for this species (Fig. 7).
365 Model AUC was 0.93 (Appendix S1: Fig. S11) and Boyce index was 0.73. Cross-validation
366 showed consistently positive AUC and Boyce indices (Appendix S1: Fig. S6, Fig. S7). Several
367 climate variables were influential in model performance, including the temperature of the wettest
368 quarter and the diurnal range in temperature (Appendix S1: Fig. S12).

369

370 Areas with surface water temperatures suitable for Northern snakehead spawning were relatively

371 limited, but they occurred in all lakes other than Lake Superior (Fig. 8). Areas with existing SAV
372 and wetlands comprised only 2% of the lake area, but constitute suitable habitat for this species
373 (Fig. 9). In most cases the spawning temperatures and aquatic vegetation coincide, indicating the
374 potential for significant ecological impact if snakehead were to establish in the lakes.

375

376 **Discussion**

377 Integrating habitat specific information that characterizes the underwater environment with
378 species distribution models improved the delineation of potential suitable habitat for non-
379 indigenous species. As a result, we were able to use an SDM with a range bagging algorithm to
380 provide informative estimates of relative climate suitability for all three species at the regional
381 scale. While these suitability estimates cannot directly predict the probability of persistence
382 without additional information (Phillips and Elith 2013), the high AUC values indicated these
383 models did successfully predict occurrence of the three species in their native and current
384 introduced ranges. The broad temperature range of killer shrimp and lack of information on other
385 habitat requirements precluded the production of a habitat specific delineation for this species.
386 However, high resolution aquatic environmental data from within the Great Lakes enabled
387 prediction at finer scale by identifying specific areas with highest chance of establishment and
388 persistence for golden mussel and snakehead, as well as areas likely to enable population growth
389 and spread. To generalize across species, Lake Erie and southern Lake Michigan contained the
390 areas most similar to the current climate niche of these species, while Lake Superior was more
391 peripheral. Further, shallow coastal habitats appear to be most at risk if these species become
392 established in the Great Lakes, whereas deeper, colder benthic habitat of all lakes appears to be
393 unsuitable.

394

395 The high AUC values for the test data and the random cross-validation showed that the range
396 bagging models were effective at determining the relative suitability of these species for their
397 native and introduced range. These high AUC values were supported by positive Boyce indices,
398 indicating the models did better than random at predicting suitability for presences in the test and
399 random cross-validation datasets. These findings were also validated for killer shrimp and
400 snakehead using the more challenging approach of spatially subdividing the datasets suggested
401 by Wenger and Olden (2012). Models for golden mussel performed more poorly when fit to

402 spatially defined subsets of the data than to all the data. In fact, the Boyce index suggested
403 performance no better than random under this cross-validation. This highlights one of the
404 difficulties of fitting species distribution models to species in the process of invading multiple
405 regions. This is unsurprising for golden mussel, given the distance and differences between the
406 native range in east Asia and invaded areas in South America. One implication is that a model fit
407 only to the native range would misrepresent and underestimate the suitability of various habitats,
408 indicating that using all available occurrences from the native and introduced ranges are likely to
409 produce the most reliable predictions. At the same time, this case reminds us that species
410 distribution models can perform poorly with non-analog climates. Accordingly, the analyses
411 presented herein models represent an attempt to estimate risk as effectively as possible, but they
412 may underestimate risk for particular locations.

413

414 To our knowledge, this is the first global niche model of golden mussel, an invader having strong
415 impacts in its introduced ranges (Ricciardi 1998). While the Great Lakes climate is relatively
416 distinct in seasonality and temperature from the observed niche, several other areas appear to be
417 highly suitable for establishment of this species, including uninvaded parts of South America, the
418 Gulf of Mexico, southeastern United States, the east coast of Australia and parts of southeast
419 Africa. Our analysis suggests that increasing surveillance in these areas may be warranted. If this
420 species could establish in the relatively novel environment of the Great Lakes, its ability to
421 reproduce would probably be limited to Lake Erie and the warmer and shallower parts of Lakes
422 Michigan, Huron and Ontario. However, this does allow that the lakes could still act as a
423 beachhead for invasion (Rothlisberger and Lodge 2013) of the more suitable lower Mississippi
424 River via the Chicago Area Waterway System, potentially mirroring the historic spread of
425 dreissenid mussels across North America. There is also potential for suitable spawning area to
426 grow given expected increases of surface water temperatures in the Great Lakes (Trumpickas et
427 al. 2009, Kao et al. 2015). The ability to incorporate these underwater specific habitat data layers
428 representing the golden mussel's temperature limitation improves the understanding of where
429 surveillance efforts can be focused. Relevant environmental layers on water chemistry or
430 substrate type are not yet available for the entirety of the lakes, but the tolerances of golden
431 mussel are broad enough that few parts of the lake would fall outside their tolerance (Ricciardi
432 1998, Boltovskoy et al. 2006).

433

434 Our analysis provides a less clear prediction for killer shrimp in the Great Lakes. The niche
435 centrality for the lakes is ambiguous, generally between 0.2 and 0.5, with Lake Superior again
436 being the most peripheral. This could mean that the Great Lakes may be marginal but possibly
437 suitable habitat for killer shrimp. However, the current distribution of killer shrimp is
438 environmentally restricted by a contiguous geographic and climatic area, and our knowledge of
439 this species is primarily derived from canal and riverine habitats (Pöckl 2009) which limits the
440 ability to develop alternative environmental limits that can be applied within the lakes. Thus, it
441 could be either that these are the only suitable environments or that the species simply has not
442 yet been transported to other novel habitats, making non-stationarity a contributor to the
443 uncertainty of predicting killer shrimp habitat. Moreover, given the existing distribution, the
444 Great Lakes had the highest niche centrality of potential introduction regions globally, consistent
445 with a previous analysis of European ports (Keller et al 2010). Since killer shrimp can persist
446 across the range of water temperatures observed in the Great Lakes and seem to have a broad
447 diet we were unable to further pinpoint high risk areas based on local environmental conditions.
448 For example, killer shrimp have been shown to successfully colonize hard substrates (i.e., stones)
449 and low density zebra mussel habitats (Kobak et al. 2015). While consistent spatial data on the
450 distribution of these habitats are not yet available for the whole of the lakes, information on these
451 habitats at the local scale could further inform assessments of risk for killer shrimp
452 establishment.

453

454 Climate conditions throughout much of the Great Lakes significantly overlap with the estimated
455 niche of northern snakehead. This is consistent with predictions of Herborg et al. (2007) as well
456 as observations of occurrences of this species elsewhere in North America, including watersheds
457 adjacent to the Great Lakes. However, our inclusion of within-lake environmental conditions
458 produced a refinement on previous models based on surface temperature, and helped identify
459 specific habitats vulnerable to this species. Required spawning temperatures indicate that specific
460 and, outside of Lake Erie, relatively limited areas would be available for reproduction.

461 SAV/wetland regions in the Great Lakes overlapped with the suitable spawning temperatures,
462 suggesting that these habitats may be vulnerable to snakehead establishment.

463

464 Modeling these three species reinforces the benefits and challenges of relying on climatic
465 variables to apply SDMs in aquatic systems, particularly large lakes and rivers that are poorly
466 coupled to proximal air temperature and precipitation (Boyce et al. 1989, Gronewold and Stow
467 2013). This is especially relevant for the Great Lakes where the bottom temperature changes
468 little throughout the year in areas deeper than about 30 m. Range bagging SDMs effectively
469 estimated the intensity of occurrence in the current range, providing important information on
470 relative suitability of locations at the regional and global scales. The analysis suggests the
471 relative importance of precipitation at a global scale as surrogates for riverine hydrologic regime
472 (Leathwick et al. 2011), although these variables are likely to have limited relevance within the
473 waters of the Great Lakes themselves and are correlated with each other. Further, the
474 microhabitat variations that are known to be important for many aquatic species are often
475 unrelated to surface climatic measurements. Here we build on previous Great Lakes case studies
476 (EPA 2008) that have combined satellite derived data layers with climatic data, and show how
477 well-characterized subsurface aquatic habitat variables can also be used to improve spatial risk
478 assessment.

479
480 The visualizations of habitat suitability produced here could be an important tool for natural
481 resource managers, and advance the understanding of the risk of invasion by these three species.
482 This process of combining information from species locations and studies of environmental
483 tolerances will have value for many other aquatic species. This process also highlights the need
484 for the development of a standardized set of global or regional aquatic habitats data layers and
485 measurements. By associating additional data on variables such as pH, nitrogen, phosphorus,
486 calcium, substrate, and water temperature with species presences, more direct, high resolution
487 models could be produced. The substantial ecological and economic costs posed by aquatic
488 invasive species should motivate the collection of such data sets and continued advances in
489 integrating different types of information.

490 491 Data Availability

492 Data and code are archived at Dryad (<http://dx.doi.org/10.5061/dryad.d4144>).

493

494 **Acknowledgments**

495 This research was funded by Environmental Protection Agency Great Lakes Restoration
496 Initiative and NOAA Center for Sponsored Coastal Ocean Research: NA09NOS4780192,
497 NA10NOS4780218. Color producing agent (CPA) retrieval images produced by the Michigan
498 Tech Research Institute (MTRI) funded by the Great Lakes Observing System (GLOS) and
499 NASA. This is NOAA GLERL contribution #1852.

500

501 **Literature Cited**

502 Amanov, A. A. 1974. Morphology and mode of life in the Amur snakehead (*Ophiocphalu argus*
503 *warpachowskii*) in Chimkurgan reservoir. Journal of Ichthyology 14:713–717.

504 Araújo, M. B., and A. T. Peterson. 2012. Uses and misuses of bioclimatic envelope modeling.
505 Ecology 93:1527–1539.

506 Austin, M. P., and K. P. Van Niel. 2011. Improving species distribution models for climate
507 change studies: variable selection and scale. Journal of Biogeography 38:1–8.

508 Barve, N., V. Barve, A. Jiménez-Valverde, A. Lira-Noriega, S. P. Maher, A. T. Peterson, J.
509 Soberón, and F. Villalobos. 2011. The crucial role of the accessible area in ecological niche
510 modeling and species distribution modeling. Ecological Modelling 222:1810–1819.

511 Beletsky, D., N. Hawley, and Y. R. Rao. 2013. Modeling summer circulation and thermal
512 structure of Lake Erie. Journal of Geophysical Research: Oceans 118:6238–6252.

513 Bisrat, S. A., M. A. White, K. H. Beard, and D. Richard Cutler. 2012. Predicting the distribution
514 potential of an invasive frog using remotely sensed data in Hawaii. Diversity and
515 Distributions 18:648–660.

516 Boets, P., K. Lock, M. Messiaen, and P. L. M. Goethals. 2010. Combining data-driven methods
517 and lab studies to analyse the ecology of *Dikerogammarus villosus*. Ecological Informatics
518 5:133–139.

519 Boltovskoy, D., N. Correa, D. Cataldo, and F. Sylvester. 2006. Dispersion and ecological impact
520 of the invasive freshwater bivalve *Limnoperna fortunei* in the Río de la Plata watershed and
521 beyond. Biological Invasions 8:947–963.

522 Boyce, F. M., M. A. Donelan, P. F. Hamblin, C. R. Murthy, and T. J. Simons. 1989. Thermal
523 structure and circulation in the great lakes. Atmosphere-Ocean 27:607–642.

- 524 Brenden, T. O., L. Wang, and P. W. Seelbach. 2008. A river valley segment classification of
525 Michigan streams based on fish and physical attributes. *Transactions of the American*
526 *Fisheries Society* 137:1621–1636.
- 527 Bruijs, M. C. M., B. Kelleher, G. Van Der Velde, and A. B. De Vaate. 2001. Oxygen
528 consumption, temperature and salinity tolerance of the invasive amphipod *Dikerogammarus*
529 *villosus*: indicators of further dispersal via ballast water transport. *Archiv für Hydrobiologie*
530 152:633–646.
- 531 Chu, P. Y., J. G. W. Kelley, G. V. Mott, A. Zhang, and G. A. Lang. 2011. Development,
532 implementation, and skill assessment of the NOAA/NOS Great Lakes Operational Forecast
533 System. *Ocean Dynamics* 61:1305–1316.
- 534 Cord, A., and D. Rödder. 2011. Inclusion of habitat availability in species distribution models
535 through multi-temporal remote-sensing data? *Ecological Applications* 21:3285–3298.
- 536 Council of Great Lakes Governors. 2013. Communication from the Council of Great Lakes
537 Governors (2013) Governors and Premiers unite to block “least wanted” aquatic invasive
538 species. USA.
- 539 Courtenay, W. R., and J. D. Williams. 2004. Snakeheads (Pisces, Channidae): a biological
540 synopsis and risk assessment. U.S. Department of the Interior, U.S. Geological Survey,
541 Reston, Virginia, USA.
- 542 Cucherousset, J., and J. D. Olden. 2011. Ecological impacts of nonnative freshwater fishes.
543 *Fisheries* 36:215–230.
- 544 Devin, S., C. Piscart, J. N. Beisel, and J. C. Moreteau. 2003. Ecological traits of the amphipod
545 invader *Dikerogammarus villosus* on a mesohabitat scale. *Archiv für Hydrobiologie* 158:43–
546 56.
- 547 Drake, J. M. 2015. Range bagging: a new method for ecological niche modelling from presence-
548 only data. *Journal of The Royal Society Interface* 12:20150086.
- 549 Elith, J., and J. R. Leathwick. 2009. Species distribution models: ecological explanation and
550 prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*
551 40:677–697.
- 552 ESRI. 2014. ArcMap 10.2. Environmental Systems Research Institute, Redlands, California,
553 USA.

- 554 Fahnenstiel, G., T. Nalepa, S. Pothoven, H. Carrick, and D. Scavia. 2010. Lake Michigan lower
555 food web: Long-term observations and *Dreissena* impact. *Journal of Great Lakes Research* 36
556 Supplement 3:1–4.
- 557 Frank, S. 1970. Acclimatization experiments with amur snake head, *Ophiocephalus argus*
558 *warpachowskii* Berg, 1909 in Czechoslovakia. *Acta Societatis Zoologicae Bohemoslovacae*
559 34:277–283.
- 560 Froese, R., and D. Pauly. 2011. FishBase. <http://www.fishbase.org>
- 561 Gallardo, B., and D. C. Aldridge. 2012. Priority setting for invasive species management: risk
562 assessment of Ponto-Caspian invasive species into Great Britain. *Ecological Applications*
563 23:352–364.
- 564 Gallien, L., R. Douzet, S. Pratte, N. E. Zimmermann, and W. Thuiller. 2012. Invasive species
565 distribution models: how violating the equilibrium assumption can create new insights. *Global*
566 *Ecology and Biogeography* 21:1126–1136.
- 567 Gantz, C. A., D. R. Gordon, C. L. Jerde, R. P. Keller, W. L. Chadderton, P. D. Champion, and D.
568 M. Lodge. 2015. Managing the introduction and spread of non-native aquatic plants in the
569 Laurentian Great Lakes: a regional risk assessment approach. *Management of Biological*
570 *Invasions* 6:45-55.
- 571 Gormley, A. M., D. M. Forsyth, P. Griffioen, M. Lindeman, D. S. L. Ramsey, M. P. Scroggie, and
572 L. Woodford. 2011. Using presence-only and presence–absence data to estimate the current
573 and potential distributions of established invasive species. *Journal of Applied Ecology* 48:25–
574 34.
- 575 Graham, C. H., and R. J. Hijmans. 2006. A comparison of methods for mapping species ranges
576 and species richness. *Global Ecology and Biogeography* 15:578–587.
- 577 Great Lakes Coastal Wetland Consortium. 2004. Great Lakes Coastal Wetland Inventory.
578 <http://glc.org/projects/habitat/coastal-wetlands/cwc-inventory/>
- 579 Gronewold, A. D., and C. A. Stow. 2013. Unprecedented seasonal water level dynamics on one
580 of the Earth's largest lakes. *Bulletin of the American Meteorological Society* 95:15–17.
- 581 Guisan, A., and W. Thuiller. 2005. Predicting species distribution: offering more than simple
582 habitat models. *Ecology Letters* 8:993–1009.

583 Herborg, L.-M., N. E. Mandrak, B. C. Cudmore, and H. J. MacIsaac. 2007. Comparative
584 distribution and invasion risk of snakehead (Channidae) and Asian carp (Cyprinidae) species
585 in North America. *Canadian Journal of Fisheries and Aquatic Sciences* 64:1723–1735.

586 Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 2005. Very high resolution
587 interpolated climate surfaces for global land areas. *International Journal of Climatology*
588 25:1965–1978.

589 Hirzel, A. H., G. Le Lay, V. Helfer, C. Randin, and A. Guisan. 2006. Evaluating the ability of
590 habitat suitability models to predict species presences. *Ecological Modelling* 199:142–152.

591 Hutchinson, G. E. 1957. Concluding remarks. *Cold Spring Harbor Symposium* 22:415–427.

592 Jiménez-Valverde, A., A. T. Peterson, J. Soberón, J. M. Overton, P. Aragón, and J. M. Lobo.
593 2011. Use of niche models in invasive species risk assessments. *Biological Invasions*
594 13:2785–2797.

595 Josens, G., A. B. de Vaate, P. Usseglio-Polatera, R. Cammaerts, F. Chérot, F. Grisez, P.
596 Verboonen, and J.-P. V. Bossche. 2005. Native and exotic Amphipoda and other Peracarida in
597 the River Meuse: new assemblages emerge from a fast changing fauna. *Hydrobiologia*
598 542:203–220.

599 Karatayev, A. Y., D. Boltovskoy, D. K. Padilla, and L. E. Burlakova. 2007. The invasive bivalves
600 *Dreissena polymorpha* and *Limnoperna fortunei*: parallels, contrasts, potential spread and
601 invasion impacts. *Journal of Shellfish Research* 26:205–213.

602 Keller, R. P., D. M. Lodge, M. A. Lewis, and J. F. Shogren. 2009. Bioeconomics of invasive
603 species: integrating ecology, economics, policy, and management. Oxford University Press,
604 Oxford, UK.

605 Kerfoot, W. C., F. Yousef, S. A. Green, J. W. Budd, D. J. Schwab, and H. A. Vanderploeg. 2010.
606 Approaching storm: disappearing winter bloom in Lake Michigan. *Journal of Great Lakes*
607 *Research* 36 Supplement 3:30–41.

608 Kilroy, C., T. H. Snelder, O. Floerl, C. C. Vieglais, and K. L. Dey. 2008. A rapid technique for
609 assessing the suitability of areas for invasive species applied to New Zealand's rivers.
610 *Diversity and Distributions* 14:262–272.

611 Kobak, J., Ł. Jermacz, and A. Dzierżyńska-Białończyk. 2015. Substratum preferences of the
612 invasive killer shrimp *Dikerogammarus villosus*. *Journal of Zoology* 297:66–76.

613 Kolar, C. S., and D. M. Lodge. 2001. Progress in invasion biology: predicting invaders. Trends in
614 Ecology & Evolution 16:199–204.

615 Kulhanek, S. A., B. Leung, and A. Ricciardi. 2011. Using ecological niche models to predict the
616 abundance and impact of invasive species: application to the common carp. Ecological
617 Applications 21:203–213.

618 Lacey Act. 1900. An act to enlarge the powers of the Department of Agriculture, prohibit the
619 transportation by interstate commerce of game killed in violation of local laws, and for other
620 purposes. chapter 553, 187-89. US Government, Washington, D.C., USA.

621 Leathwick, J. R., D. Rowe, J. Richardson, J. Elith, and T. Hastie. 2005. Using multivariate
622 adaptive regression splines to predict the distributions of New Zealand’s freshwater
623 diadromous fish. Freshwater Biology 50:2034–2052.

624 Leathwick, J. R., T. Snelder, W. L. Chadderton, J. Elith, K. Julian, and S. Ferrier. 2011. Use of
625 generalised dissimilarity modelling to improve the biological discrimination of river and
626 stream classifications. Freshwater Biology 56:21–38.

627 Leung, B., et al. 2012. TEASIng apart alien species risk assessments: a framework for best
628 practices. Ecology Letters 15:1475–1493.

629 Lockwood, J. L., P. Cassey, and T. Blackburn. 2005. The role of propagule pressure in explaining
630 species invasions. Trends in Ecology & Evolution 20:223–228.

631 Lods-Crozet, B., and O. Reymond. 2006. Bathymetric expansion of an invasive gammarid
632 (*Dikergammarus villosus*, Crustacea, Amphipoda) in Lake Léman. Journal of Limnology
633 65:141.

634 MacNeil, C., et al. 2010. The Ponto-Caspian ‘killer shrimp’, *Dikergammarus villosus* (Sowinsky,
635 1894), invades the British Isles. Aquatic Invasions 5:441–445.

636 McKenna, Jr., J. E., and C. Castiglione. 2010. Hierarchical multi-scale classification of nearshore
637 aquatic habitats of the Great Lakes: Western Lake Erie. Journal of Great Lakes Research
638 36:757–771.

639 Michigan Tech Research Institute. 2012. Satellite-derived lake submerged aquatic vegetation
640 (SAV). <http://www.mtri.org/cladophora.html>.

641 Mohseni, O., and H. G. Stefan. 1999. Stream temperature/air temperature relationship: a physical
642 interpretation. Journal of Hydrology 218:128–141.

643 Okada, Y. 1960. Studies of the freshwater fishes of Japan, II, Special part. Journal of the Faculty
644 of Fisheries Prefectural University of Mie 4:31–860.

645 Pagel, J., and F. M. Schurr. 2012. Forecasting species ranges by statistical estimation of
646 ecological niches and spatial population dynamics. *Global Ecology and Biogeography*
647 21:293–304.

648 Petitpierre, B., C. Kueffer, O. Broennimann, C. Randin, C. Daehler, and A. Guisan. 2012.
649 Climatic niche shifts are rare among terrestrial plant invaders. *Science* 335:1344–1348.

650 Phillips, S. J., and J. Elith. 2013. On estimating probability of presence from use–availability or
651 presence–background data. *Ecology* 94:1409–1419.

652 Pöckl, M. 2009. Success of the invasive Ponto-Caspian amphipod *Dikerogammarus villosus* by
653 life history traits and reproductive capacity. *Biological Invasions* 11:2021–2041.

654 Ricciardi, A. 1998. Global range expansion of the Asian mussel *Limnoperna fortunei*
655 (Mytilidae): Another fouling threat to freshwater systems. *Biofouling* 13:97–106.

656 Ricciardi, A., and H. J. MacIsaac. 2000. Recent mass invasion of the North American Great
657 Lakes by Ponto–Caspian species. *Trends in Ecology & Evolution* 15:62–65.

658 Ricciardi, A., and J. B. Rasmussen. 1998. Predicting the identity and impact of future biological
659 invaders: a priority for aquatic resource management. *Canadian Journal of Fisheries and*
660 *Aquatic Sciences* 55:1759–1765.

661 Rothlisberger, J. D., and D. M. Lodge. 2013. The Laurentian Great Lakes as a beachhead and a
662 gathering place for biological invasions. *Aquatic Invasions* 8:361–374.

663 Schwab, D. J., and K. W. Bedford. 1994. Initial implementation of the great lakes forecasting
664 system: a real-time system for predicting lake circulation and thermal structure. *Water Quality*
665 *Research Journal of Canada* 292–3:203–220.

666 Shuchman, R. A., M. J. Sayers, and C. N. Brooks. 2013. Mapping and monitoring the extent of
667 submerged aquatic vegetation in the Laurentian Great Lakes with multi-scale satellite remote
668 sensing. *Journal of Great Lakes Research* 39 Supplement 1:78–89.

669 Snelder, T. H., et al. 2006. Development of an ecologic marine classification in the New Zealand
670 region. *Environmental Management* 39:12–29.

671 Tyberghein, L., H. Verbruggen, K. Pauly, C. Troupin, F. Mineur, and O. De Clerck. 2012. Bio-
672 ORACLE: a global environmental dataset for marine species distribution modelling. *Global*
673 *Ecology and Biogeography* 21:272–281.

674 US Environmental Protection Agency. 2008. Predicting future introductions of nonindigenous
675 species to the Great Lakes, EPA/600/R-08/066F. National Center for Environmental
676 Assessment, Washington, D.C., USA.

677 USACE. 2011. Great Lakes and Mississippi Interbasin Study: non-native species of concern and
678 dispersal risk for the Great Lakes and Mississippi River Interbasin Study. U.S. Army Corps of
679 Engineers, Chicago, Illinois, USA.

680 Václavík, T., J. A. Kupfer, and R. K. Meentemeyer. 2012. Accounting for multi-scale spatial
681 autocorrelation improves performance of invasive species distribution modelling (iSDM).
682 *Journal of Biogeography* 39:42–55.

683 Václavík, T., and R. K. Meentemeyer. 2012. Equilibrium or not? Modelling potential distribution
684 of invasive species in different stages of invasion. *Diversity and Distributions* 18:73–83.

685 Vanderploeg, H. A., J. R. Liebig, T. F. Nalepa, G. L. Fahnenstiel, and S. A. Pothoven. 2010.
686 *Dreissena* and the disappearance of the spring phytoplankton bloom in Lake Michigan.
687 *Journal of Great Lakes Research* 36 Supplement 3:50–59.

688 VanDerWal, J., H. T. Murphy, A. S. Kutt, G. C. Perkins, B. L. Bateman, J. J. Perry, and A. E.
689 Reside. 2013. Focus on poleward shifts in species' distribution underestimates the fingerprint
690 of climate change. *Nature Climate Change* 3:239–243.

691 Velde, G. van der, R. S. E. W. Leuven, D. Platvoet, K. Bacela, M. A. J. Huijbregts, H. W. M.
692 Hendriks, and D. Kruijt. 2009. Environmental and morphological factors influencing
693 predatory behaviour by invasive non-indigenous gammaridean species. *Biological Invasions*
694 11:2043–2054.

695 Veloz, S. D., J. W. Williams, J. L. Blois, F. He, B. Otto-Bliesner, and Z. Liu. 2012. No-analog
696 climates and shifting realized niches during the late quaternary: implications for 21st-century
697 predictions by species distribution models. *Global Change Biology* 18:1698–1713.

698 Wenger, S. J., and J. D. Olden. 2012. Assessing transferability of ecological models: an
699 underappreciated aspect of statistical validation. *Methods in Ecology and Evolution* 3:260–
700 267.

701 Wijnhoven, S., M. C. van Riel, and G. van der Velde. 2003. Exotic and indigenous freshwater
702 gammarid species: physiological tolerance to water temperature in relation to ionic content of
703 the water. *Aquatic Ecology* 37:151–158.

704 Wittmann, M. E., G. Annis, A. M. Kramer, L. Mason, C. Riseng, E. S. Rutherford, W. L.
705 Chadderton, D. Beletsky, J. M. Drake, and D. M. Lodge. 2017. Refining species distribution
706 model outputs using landscape-scale habitat data: forecasting grass carp and *Hydrilla*
707 establishment in the Great Lakes region. *Journal of Great Lakes Research* 43:298–307.
708 Xu, M., Z. Wang, C. C. Lin, B. Pan, and N. Zhao. 2013. Experimental study of invasion and
709 biofouling of freshwater mussel *Limnoperna fortunei*. *International Journal of Geosciences*
710 04:1–7.

Author Manuscript

711 **Figure Legends**

712 Figure 1: Golden mussel niche centrality in the North American Great Lakes basin. Low values
713 of niche centrality indicate climate conditions in the Great Lakes basin are often outside of the
714 predicted niche.

715

716 Figure 2: Golden mussel global niche. Map of niche centrality values with higher values
717 indicating climate conditions falling within the modeled niche.

718

719 Figure 3: Golden mussel habitat suitability restricted by spawning habitat. Visualization of the
720 climate-based niche in areas of the Great Lakes warm enough for golden mussel spawning.

721 Spawning can occur at benthic temperatures greater than 16°C (Xu et al. 2013).

722

723 Figure 4: Killer shrimp niche centrality in the North American Great Lakes basin. Intermediate
724 values of niche centrality indicate climate conditions in the Great Lakes basin often, but not
725 completely, overlap the predicted niche.

726

727 Figure 5: Killer shrimp global niche. Map of niche centrality values with higher values indicating
728 climate conditions falling within the modeled niche.

729

730

731 Figure 6: Northern snakehead niche centrality in the North American Great Lakes basin. Map of
732 niche centrality for northern snakehead in the Great Lakes basin. High values of niche centrality
733 indicate climate conditions in the Great Lakes basin fall generally within the predicted niche.

734

735 Figure 7: Snakehead global niche. Map of niche centrality values with higher values indicating
736 climate conditions falling within the modeled niche.

737

738 Figure 8: Northern snakehead habitat suitability restricted by spawning habitat. Visualization of
739 the climate-based niche in areas of the Great Lakes warm enough for snakehead spawning.

740 Spawning can occur at benthic temperatures greater than 18°C.

741

742 Figure 9: Northern snakehead habitat suitability restricted by aquatic vegetation. Visualization of
743 the climate-based niche in wetlands and areas of submerged aquatic vegetation. These are
744 preferred habitat for northern snakehead. Note that submerged aquatic vegetation data is missing
745 for Lake St. Clair.

Author Manuscript

















