



**QUEEN'S
UNIVERSITY
BELFAST**

Combining global climate and regional landscape models to improve prediction of invasion risk

Kelly, R., Leach, K., Cameron, A., Maggs, C. A., & Reid, N. (2014). Combining global climate and regional landscape models to improve prediction of invasion risk. *Diversity and Distributions*, 20(8), 884–894. <https://doi.org/10.1111/ddi.12194>

Published in:
Diversity and Distributions

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
[Link to publication record in Queen's University Belfast Research Portal](#)

Publisher rights
Copyright 2014 John Wiley & Sons Ltd

This is the accepted version of the following article: Kelly, R, Leach, K, Cameron, A, Maggs, CA & Reid, N 2014, 'Combining global climate and regional landscape models to improve prediction of invasion risk' *Diversity and Distributions*, vol 20, no. 8, pp. 884–894., 10.1111/ddi.12194, which has been published in final form at <http://onlinelibrary.wiley.com/doi/10.1111/ddi.12194/abstract>.

General rights
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Open Access
This research has been made openly available by Queen's academics and its Open Research team. We would love to hear how access to this research benefits you. – Share your feedback with us: <http://go.qub.ac.uk/oa-feedback>

1 **Title: Combining global climate and regional landscape models to improve prediction of**
2 **invasion risk**

3

4 **Running title:** Multi-scale models of invasion risk

5

6 Ruth Kelly^{1*}, Katie Leach¹, Alison Cameron², Christine A. Maggs² & Neil Reid¹

7

8 ¹ *Quercus, School of Biological Sciences, Queen's University Belfast, Belfast BT9 7BL, UK*

9 ² *School of Biological Sciences, Queen's University Belfast, Belfast, BT9 7BL, UK*

10

11 ***Corresponding author:** *Tel.: +44 28 9097 2281, +44 77 07627488, Fax: +44 28 9097*

12 *5877;*

13 *E-mail address: ruth.kelly@qub.ac.uk*

14

15 Word count: 5,235

16

17 **(A) Abstract**

18

19 **Aim**

20 It is widely acknowledged that species distributions result from a variety of biotic and abiotic
21 factors operating at different spatial scales. Here, we have aimed to: i) determine the extent to
22 which global climate niche models can be improved by the addition of fine-scale regional
23 data; ii) examine climatic and environmental factors influencing the range of 15 invasive
24 aquatic plant species; and iii) provide a case study for the use of such models in invasion
25 management on an island.

26 **Location**

27 Global, with a case study of species invasions in Ireland.

28 **Methods**

29 Climate niche models of global extent (including climate only) and regional environmental
30 niche models (with additional factors such as human influence, land use and soil
31 characteristics) were generated using MaxEnt for 15 invasive aquatic plants. The
32 performance of these models within the invaded range of the study species in Ireland was
33 assessed, and potential hotspots of invasion suitability were determined. Models were
34 projected forward up to 2080 based on two climate scenarios.

35 **Results**

36 While climate variables are important in defining the global range of species, factors related
37 to land use and nutrient level were of greater importance in regional projections. Global
38 climatic models were significantly improved at the island scale by the addition of fine-scale
39 environmental variables (Area Under the Curve values increased by 0.18 and True Skill
40 Statistic by 0.36), and projected areas decreased from an average of 86% to 36% of the
41 island.

42 **Conclusion**

43 Refining climate niche models with regional data on land use, human influence and landscape
44 may have a substantial impact on predictive capacity, providing greater value for
45 prioritisation of conservation management at sub-regional or local scales.

46

47 **Keywords:** Aquatic, climate change, freshwater, invasive, macrophyte, MaxEnt.

48 (A) Introduction

49 Projections of the future range of invasive species based on climate niche models have
50 become increasingly prevalent (Thuiller *et al.*, 2005; Jiménez-Valverde *et al.*, 2011). Niche
51 models are a type of species distribution model that employs correlative techniques to
52 estimate the environmental niche of species based on species distribution data and maps of
53 environmental variables. This estimated species niche is then used to infer species occurrence
54 or habitat suitability. In invasion ecology, niche models are most commonly used to predict
55 'risk' of invasion based on suitability rather than species occurrence per se. Niche models
56 have identified likely source regions of 'high risk' invaders (e.g. Thuiller *et al.*, 2005), as well
57 as 'hotspots' where multiple invasions are most likely to occur (Ibáñez *et al.*, 2009;
58 O'Donnell *et al.*, 2012) and, conversely, areas where invasive species may decline, creating
59 potential for restoration of native communities (e.g. Bradley *et al.*, 2009).

60 Species distributions result from a variety of biotic and abiotic factors operating at
61 different spatial scales (Gaston, 2003). In invasion biology, this is often conceptualised as a
62 set of abiotic filters (Theohairdes & Dukes, 2007; Milbau *et al.*, 2009). Hence, while climate
63 limits species distributions at global scales, the prediction of invasion risk on regional scales
64 is likely to be improved by information on factors such as land use, soil or water nutrient
65 concentrations, pH and human influence. The relative importance of climate versus other
66 environmental variables in niche modelling has been investigated for various invasive species
67 within regions or continents (e.g. Thuiller *et al.*, 2006; Nielsen *et al.*, 2008; Ibanéz *et al.*,
68 2009; Capinha & Anastácio, 2011; Jarnevich *et al.*, 2011; Compton *et al.*, 2012). Further
69 studies have shown the importance of topography, habitat, microclimate and water
70 availability in determining species distributions at smaller spatial scales (e.g. Bossenbroek *et*
71 *al.*, 2004; Evangelista *et al.*, 2006; Gillingham *et al.*, 2012). However, the optimal method for
72 selecting predictor variables and their appropriate spatial scales in niche models remains

73 uncertain (Austin & Kimberley, 2011), and is likely to depend on the organism being
74 modelled (Araújo & Guisan, 2006).

75 The appropriate spatial extent of a study region depends largely on the aims of the study
76 and the nature of the study species (Elith & Leathwick, 2009). Climate niche models with a
77 global extent are most likely to be useful for modelling species with broad climatic niches
78 and high dispersal capacity such as invasive species. Niche models that include climate data
79 from only either the native or the invaded range of species tend to underestimate potential
80 niche space (Beaumont *et al.*, 2009; Jiménez-Valverde *et al.*, 2011). For this reason, several
81 recent authors have used global distribution data to create niche models of invasion risk based
82 on global climate. However, these models rarely incorporate fine-scale data, for example, on
83 local nutrient availability, pH or soil characteristics. This is probably due to the lack of
84 reliable data on such characteristics globally.

85 Another challenge in constructing global niche models may be the considerable
86 uncertainty regarding the precision of species distribution records in global datasets, hence
87 larger grid cells may be more appropriate for building global models. Conversely, national
88 and regional species databases usually have high precision. Here we demonstrate how it may
89 be useful to combine the outputs of global climate niche models with regional niche models
90 in order to create integrated models which make better use of all available datasets.

91 We examined the global climatic factors influencing the distribution of 15 invasive
92 freshwater plants considered 'high risk' in Europe. Aquatic plants are primary components of
93 freshwater ecosystems and invasive species have the capacity to fundamentally alter
94 ecosystem functioning. However, to date climate niche models of freshwater ecosystems
95 have focused mainly on animals (e.g. Liu *et al.*, 2011; Reshetnikov & Ficetola, 2011) and few
96 studies have used climate niche models to examine the global distributions of invasive
97 aquatic plants.

98 We provide a simple method for the amalgamation of global climate niche models with
99 regional data on land use, human influence, soil and water characteristics for the invaded
100 range, and show how this significantly changes predictive capacity at the regional scale.
101 Specifically, we aimed to i) determine the extent to which global climate niche models can be
102 improved by the addition of fine-scale regional data, ii) examine the factors governing the
103 range of key invasive aquatic plant species, and iii) provide a case study for the use of such
104 models in invasion management on an island.

105

106 **(A) Methods**

107

108 **(B) Species selection**

109 We modelled the distributions of eight non-native aquatic plants which are currently
110 established invasives in Ireland (*Azolla filiculoides* Lam., *Crassula helmsii* (Kirk) Cockayne,
111 *Elodea canadensis* Michx., *Elodea nuttallii* H. St. John, *Lagarosiphon major* (Ridl.) Moss,
112 *Lemna minuta* (Kunth), *Myriophyllum aquaticum* (Vell.) Verdc and *Nymphoides peltata*
113 (S.G. Gmel.) Kuntze). These species represent a range of introduction dates from 1836 to
114 1993. In addition we modelled seven potentially high-risk species which are either not
115 currently present in Ireland (*Cabomba caroliniana* A. Gray, *Eichhornia crassipes* (Mart.)
116 Solms, and *Salvinia molesta* D.S. Mitch) or present at less than 5 locations in Ireland (*Egeria*
117 *densa* Planch., *Hydrilla verticillata* (L.f.) Royle, *Hydrocotyle ranunculoides* L.f., *Ludwigia*
118 *grandiflora* (Michx) Greuter & Burdet). These were chosen based on the European and
119 Mediterranean Plant Protection Organisation (EPPO) lists of high risk species
120 (http://www.eppo.int/invasive_plants/ias_plants.htm). *H. verticillata* is a cryptogenic species
121 in Europe and its origin and status in Ireland are unclear. The native status of the species in
122 Ireland has been questioned due to its isolated distribution, distance from its native range,
123 sub-optimal growth conditions and the occurrence of female plants only (Minchin, 2007).
124 The decision to include this species reflects this uncertainty.

125

126 **(B) Niche models**

127 Firstly, niche models fitted with global climate data (Climate Niche Models (CNMs)) were
128 constructed for all species based on global distribution data. Secondly, for all 8 species which
129 are established in the study region (Ireland), niche models fitted with regional land cover
130 data, human influence, soil characteristics and climate suitability outputs from global CNMs

131 were constructed using species distributions within Ireland (hereafter regional Environmental
132 Niche Models (ENMs)). The projected output raster from the global CNM, representing the
133 climatic suitability of each cell for the target species, was included as an explanatory variable
134 in regional ENMs.

135 This two-stage integrated modelling approach was adopted for a number of reasons.
136 Firstly, it was not appropriate to build a climate suitability model based solely on Ireland or
137 Europe as the potential climatic range of each species was likely to be larger than either
138 region. Niche models that incorporate global climate data and global species distributions are
139 likely to be important when projecting species ranges into future climates which are novel to
140 the study region, but similar to current climates elsewhere. Secondly, reliable fine-scale
141 regional data were sometimes available only for Ireland (e.g. water nutrient content) or
142 Europe (e.g. CORINE land cover maps) and were, therefore, best suited to regional scale
143 modelling. Thirdly, uncertainty regarding the precision of records obtained from global
144 databases made them unsuitable for regional model building. Our two-stage approach
145 allowed for the incorporation of global climatic data into regional ENMs. Models were built
146 using MaxEnt 3.3.3k (Phillips *et al.*, 2010).

147

148 **(B) Global Climate Niche Models (CNMs)**

149 Data on the recorded distribution of each species were downloaded from the Global
150 Biodiversity Information Facility data portal (www.gbif.org). Records with no location
151 coordinates were excluded from the dataset; >97% of included records had a precision of
152 greater than 0.1 decimal degrees. Records prior to 1950 were excluded to match the
153 timeframe for the 'current' climate data (1950-2000). Species had between 122 and 4,474
154 records (mean = 1,395). Records from both native and invaded ranges were used to avoid
155 underestimation of climatic niche space. All species included had invaded ranges outside the

156 island of Ireland, and all species present in Ireland were also invasive in other climatically
157 similar regions (e.g. the UK). Plant records from Ireland were not included in the global
158 CNMs as these records would be used in training and testing of subsequent regional models.
159 Ireland is small and has a narrow range of climatic conditions which were not novel when
160 compared with global training data. This assumption was validated after model fitting by
161 examining the multivariate environmental similarity surface (MESS) output from MaxEnt
162 (following Elith *et al.*, 2010).

163 Standard climatic variables were downloaded from WorldClim at a 2.5° cell resolution
164 (<http://www.worldclim.org/bioclim>). A total of 8 variables were used, namely annual mean
165 temperature (Bio1), temperature seasonality (Bio4), maximum temperature of warmest month
166 (Bio5), minimum temperature of coldest month (Bio6), annual precipitation (Bio12),
167 precipitation seasonality (Bio15), mean precipitation of wettest quarter (Bio16) and
168 precipitation of driest quarter (Bio17). These variables were selected based on their relevance
169 to the ecology of freshwater plants. All other WorldClim variables were excluded in order to
170 reduce multi-collinearity and subsequent model over-fitting (Jiménez-Valverde *et al.*, 2011).
171 Terrestrial climate variables will not exactly reflect the conditions of freshwater systems;
172 however, terrestrial climate data have been shown to be an effective proxy for shallow waters
173 (e.g. Reshetnikov & Ficetola, 2011).

174 Forecast projections for future climatic scenarios were from the International Panel on
175 Climate Change (IPCC) 4th assessment report (IPCC, 2007), in particular, the Special Report
176 Emissions Scenarios (SRES) ‘A2’ (“high energy requirements and emissions”) and ‘B2’
177 (“low energy requirements and emissions”). Spatial climate projections were the CSIRO
178 MK2 global climate model for each scenario for the 2020s (2010 to 2039), 2050s (2050 to
179 2069) and 2080s (2070 to 2099) and were downloaded from <http://www.ccafsc->
180 [climate.org/data/](http://www.ccafsc-climate.org/data/).

181 Plant records were split into a 75% model ‘training set’ and a 25% model ‘test set’.
182 Duplicate records were excluded. A combination of linear, quadratic and product features
183 was used for estimating the response of each species to each predictor variable. Extrapolation
184 was not used so climate suitability was not projected where conditions were outside the range
185 of the training data (i.e. background and presence data).

186 Presence-only modelling techniques are particularly susceptible to recorder-bias (Phillips
187 *et al.*, 2009; Yackulic *et al.*, 2013), due to the fact that they rely on random background data
188 points rather than recorded absences. Sample bias can be minimised in MaxEnt by restricting
189 the selection of background points to represent the environmental conditions that were
190 sampled (Elith *et al.*, 2011). Here, we used GBIF records of two major aquatic plant families
191 (Hydrocharitaceae and Menyanthaceae) to account for the distribution of recording effort of
192 aquatic botanists. These plant families have a cosmopolitan distribution, and contain only
193 aquatic and semi-aquatic species. Background points ($n = 10,000$) were chosen exclusively
194 from locations in which these aquatic plant families or the invasive species studied were
195 recorded, rather than assuming uniform recording effort.

196

197 **(B) Regional Environmental Niche Models (ENMs)**

198 Data on species distributions in Ireland were collated from local government agencies, data
199 providers and botanical societies (see Appendix S1 in Supporting Information). Regional
200 ENMs were built using Irish data only and within the spatial extent of the island (for those 8
201 species already established in Ireland). Records with a resolution of >100 m were excluded in
202 order to match the cell size of the environmental data. As with the global dataset, duplicates
203 and records pre-1950 were excluded. Species had between 13 and 1,646 records (mean =
204 262) (data from within Ireland only).

205 Environmental variables were chosen based on ecological relevance to study species, with
206 both “direct” and “indirect” associations with the distributions of study species (Austin,
207 2007). Variables included were human impact i.e. propagule transport and disturbance
208 facilitating establishment (Buchan & Padilla, 2000; Aznar *et al.*, 2003; Leung, 2006; Capers
209 *et al.*, 2009); land use, including forestry (Buchan & Padilla, 2000) and agriculture (Heegaard
210 *et al.*, 2001) which affect aquatic species indirectly due to their impact on surface water run-
211 off and water chemistry; and physicochemical parameters directly influencing plant growth
212 including pH, nutrient availability, substratum types and amount of standing freshwater
213 (Heegaard *et al.*, 2001; Capers *et al.*, 2009; Alahuhta *et al.*, 2011).

214 Landscape composition was derived from CORINE Land Cover maps (EEA, 2002).
215 Specifically, the coverage of arable land, broad-leaved woodland, coniferous plantations,
216 pasture, urban and standing freshwater was quantified at seven candidate spatial scales (0.5
217 km, 1.5 km, 2.5 km, 4.5 km, 6.5 km, 10.5 km and 20.5 km) and attributed to the central 0.5
218 km grid cell around which they had been extracted. These spatial scales were chosen to
219 represent the scales on which land use and human influence could plausibly influence the
220 establishment of aquatic plants. The total length of riparian corridor and an index of human
221 influence, representing a proxy for human-mediated vectors of dispersal, were also quantified
222 at the same scales. The Human Influence Index was a composite of population density, access
223 infrastructure (major roads, railways, navigable rivers and coastline) and night-time light
224 emissions downloaded from the Socio-Economic Data and Applications Centre (Wildlife
225 Conservation Society & Center for International Earth Science Information Network, 2005).

226 Nutrient concentrations were obtained from the Environmental Protection Agency of
227 Ireland (EPA) and the Northern Ireland Environment Agency (NIEA), expressed as
228 ‘orthophosphate’ (mg/l) measured at river monitoring sites. Data for missing areas were
229 interpolated between empirical data points using the Kriging function in the Spatial Analyst

230 toolbar for ArcGIS 10.1 (ESRI, California, USA). Soil pH was taken as a proxy of water pH
231 and downloaded from <http://eussoils.jrc.ec.europa.eu>. Soil type (texture) was expressed as a
232 categorical variable representing levels of mineral coarseness (Gardi *et al.*, 2008).
233 Orthophosphate, soil pH and soil texture were expressed at the 0.5 km scale. The projected
234 output raster from the global CNM for each species, representing the climatic suitability of
235 each cell for the target species, was also incorporated into the regional ENMs.

236 As with global records, Irish records were biased due to sampling effort. However, the
237 nature of this bias differed from that of the GBIF records. Aquatic plants have been
238 systematically recorded throughout Ireland in the vast majority of lakes and rivers (as part of
239 European Water Framework Directive monitoring). Hence, it was not necessary to use the
240 recording of other major aquatic plant species as a proxy for aquatic plant recording.
241 However, it was evident that the volume of recording per unit area for Northern Ireland was
242 substantially greater than that in the Republic of Ireland. To account for the
243 disproportionately high density of records in Northern Ireland, a 'bias' layer was constructed
244 with values of 1 for the Republic of Ireland and 10 for Northern Ireland based on relative
245 prevalence of records of the most commonly recorded species (*Elodea canadensis*). This bias
246 layer was included in MaxEnt models (following Elith *et al.*, 2011). Models of invasive
247 species may also be biased when background data are included from outside the area where
248 the target species has had the opportunity to colonise (Elith *et al.*, 2010; Jiménez-Valverde *et*
249 *al.*, 2011). Therefore, background points ($n=10,000$) were randomly selected within the
250 'invaded range' of each species (based on the minimum convex polygons (MCP) that
251 included all records).

252 Prior to final model construction, separate models were constructed for each land use
253 variable and the Human Influence Index to ascertain the optimal spatial scale for each
254 variable for each species. One variable (e.g. area of arable land) was fitted at all seven

255 candidate spatial scales. The optimal spatial scale for each variable was chosen based on the
256 maximum Area Under the Curve (AUC) value when only one spatial scale was included
257 (based on Jackknife AUC results for test data; see Appendix S2). Response features were
258 limited to linear and quadratic excluding product features as we were not interested in
259 interactions between scales within variables.

260 The final regional ENMs used linear, quadratic and product features. A 10-fold replicated
261 cross-validation routine was used for all species. The averaged model across replicates was
262 used for analysis of variable permutation importance and estimated suitable range. As with
263 global CNMs, all models were projected up to 2080 under A2 and B2 scenarios. All non-
264 climate variables were held constant in future scenarios.

265

266 **(B) Statistical analyses**

267 The projected 'estimated suitable range' was calculated for each species, model and scenario
268 based on the percentage of grid cells with an occurrence probability of greater than the 10th
269 percentile training presence threshold (providing a conservative estimate allowing records in
270 suboptimal habitats). Changes in the estimated suitable range were expressed as the increase
271 or decrease compared to the total land area of Ireland, rather than percentage increase
272 compared with the current range (e.g. an expansion of estimated suitable range from 20% to
273 40% of the island was expressed as an increase of 20%, not 100%).

274 For species currently invasive in Ireland ($n=8$) differences in the estimated suitable range
275 between species, scenarios and years were examined using a General Linear Model (GLM).
276 Species, model type (i.e. global CNM or regional ENM) and scenario were fitted as fixed
277 factors, year as an ordinal factor and all two-way interactions were included. GLM residuals
278 were checked to ensure they conformed to normality.

279 Similarly, for species currently invasive in Ireland the performance of both model types
280 (global CNMs and regional ENMs) was evaluated using AUC values (as a threshold
281 independent metric), sensitivity (proportion of presences which are predicted correctly),
282 specificity (proportion of absences which are predicted correctly) and True Skill Statistic or
283 TSS (using the 10th percentile training presence as the threshold value). TSS is equal to
284 sensitivity plus specificity minus one and is a prevalence independent model evaluation
285 metric (Allouche *et al.*, 2006). Therefore, TSS is the only metric in this study which can be
286 compared between species. TSS ranges from -1 to +1, where 1 indicates perfect agreement
287 and values of less than 0 indicate models which perform worse than random. We include
288 sensitivity and specificity separately from AUC and TSS to clarify the relative importance of
289 omission and commission errors in model performance (Jiménez-Valverde *et al.*, 2008; Lobo
290 *et al.*, 2008).

291 While specificity conveys important information about commission errors, it should be
292 interpreted with caution in studies such as this one, where species are not at equilibrium with
293 the environment. In particular, in the case of invasive species misleading commission errors
294 are likely to arise in areas which are potentially suitable for species, but which the species
295 have not yet colonised. Hence, low specificity (i.e. high commission error) may arise when
296 models are actually accurately predicting areas of suitability which have not yet been
297 colonised. This is most likely in the early stages of invasion (Václavík & Meentemeyer,
298 2012). This will affect AUC values, specificity and TSS which all include commission error.
299 Here, we have minimised this source of error by selecting pseudo-absences from within a
300 minimum convex polygon containing currently invaded sites.

301 These evaluation statistics were calculated for global and regional models using the same
302 testing of sub-sets of Irish data and background points within the invaded range in Ireland.
303 The presence records (test datasets) and numbers of background points were the same as

304 those used in the 10-fold cross-validation of regional ENMs (such that no model was tested
305 against the same data set that was used to build the model). In our analyses, we focus on
306 differences in evaluation metrics between global CNMs and regional ENMs within species
307 where evaluation metrics were calculated on same datasets (hence equal prevalence). Model
308 performance of global CNMs and regional ENMs was then compared using a Generalised
309 Linear Model (GLM), with model type and species as fixed factors. A gamma error
310 distribution with a log link function was used to account for non-normality of the response
311 variables.

312 The importance of each predictor variable in global CNMs and regional ENMs was
313 described using the permutation importance (%). For regional ENMs the permutation
314 importance values were grouped into ‘Land use and nutrient concentration’ (arable farming,
315 broad-leaved woodland, coniferous plantations, farming type and water nutrient
316 concentration), ‘Human Influence Index’, ‘Natural landscapes’ (riparian length, soil pH, soil
317 type and amount of standing freshwater) and ‘Climate’ (output from global CNMs).

318 The estimated suitability of projections for multiple species was calculated for each grid
319 cell by summing the number of species for which the projected suitability was greater than
320 the 10th percentile training presence threshold under the high emissions scenario. Suitability
321 for multiple invasive species was mapped and coloured to identify putative hot- and coldspots
322 of invasion. Mapping and spatial analysis were conducted in ArcGIS 10. Model evaluation
323 statistics, GLM and GLMM analyses were conducted in R 2.14.1, using the add-on packages
324 “PresenceAbsence” (Freeman & Moisen, 2008) and “glmmADMB” (Fournier *et al.*, 2012).
325 Graphs were generated in Sigma Plot 10 (Systat, 2010).

326

327 **(A) Results**

328

329 **(B) Comparison between global CNMs and regional ENMs**

330

331 Regional ENMs had significantly higher AUC, TSS and specificity values than global CNMs,
332 whilst the opposite was true for sensitivity values (Table 1). Species which showed an
333 increase in estimated suitable range over time in global CNMs also increased in regional
334 ENMs, similarly species which showed a decrease in estimated suitable range over time in
335 global CNMs also decreased in regional ENMs. Estimates of suitable range were
336 significantly larger in global CNMs than in regional ENMs (Table 1 and Fig. 1).

337

338 **(B) Factors influencing invasive aquatic plant distributions**

339

340 The permutation importance of variables varied considerably between species for both global
341 climatic and regional landscape variables. In the global CNMs, annual mean temperature
342 (Bio1) and minimum temperature of coldest month (Bio6) were the most important variables
343 (Fig. 2 and Appendix S3). In regional ENMs, land use and nutrient concentration variables
344 had the greatest overall permutation importance. Climate was generally least important in
345 regional ENMs (Fig. 3 and Appendix S4). However, climate was the single most important
346 variable associated with the range of *Azolla filiculoides* (permutation importance = 32%),
347 although for other species it was less informative (permutation importance 0-7%).

348

349 **(B) Case study: Invasion risk in Ireland**

350 The estimated suitable range for each invasive species in Ireland varied between species ($F =$
351 59.31 , $p < 0.001$). Changes in the estimated suitable range differed between species, with
352 some species showing range expansions under projected climate scenarios and some
353 declining ($F = 5.97$, $p < 0.001$; Fig. 1 and Appendix S5). There was an increase in the
354 estimated suitable range for aquatic invasive plants under both climate scenarios between
355 current conditions and projections up to 2080 ($F = 3.54$; $p = 0.018$). Typically, such changes
356 were greater under scenario A2 or high emissions than B2 or low emissions.

357 Global CNMs projected that under the high emissions scenario the estimated suitable
358 range would increase moderately for three species already well-established in Ireland (*A.*
359 *filiculoides* +7%, *Lemna minuta* +7% and *Myriophyllum aquaticum* +12%). The ranges of a
360 further three species recorded from only one or two sites in Ireland were projected to increase
361 dramatically under the same scenario (*Egeria densa* +91%, *Hydrilla verticillata* +40% and
362 *Ludwigia grandiflora* +33%). In contrast, the ranges of two well-established species (*Elodea*
363 *canadensis* -9%, and *E. nuttallii* -32%), were projected to decrease under this scenario.
364 Concurrently, there was projected to be little change in three well-established species
365 (*Crassula helmsii*, *Lemna major* and *Nymphoides peltata*). Current and future climates were
366 projected to be unsuitable for three ‘high risk’ invaders (*Cabomba caroliniana*, *Eichhornia*
367 *crassipes* and *Salvinia molesta*) and indicated that the climate may remain unsuitable for
368 these species until at least 2080. *Hydrocotyle ranunculoides* is present at only a few sites but
369 projections indicated that climate is currently, and will remain, very suitable.

370 Regional ENMs projected that the estimated suitable range under high emissions would
371 increase for two of the three species which were also projected to increase using the global
372 CNM models (*A. filiculoides* +22% and *M. aquaticum* +44%). In addition, the range
373 decreased for one well-established species (*E. canadensis* -31%) for which declines were

374 projected by global CNMs. Concurrently, regional ENMs projected little change in the
375 remaining five species (*C. helmsii*, *E. nuttallii*, *L. major*, *L. minuta* and *N. peltata*; Fig. 1).

376

377 **(B) Temporal changes in hot- and coldspots of invasion**

378 Projected hotspots of invasion suitability exist currently around major cities and river systems
379 (Fig. 4). Future projections based on regional ENMs under the high emissions scenario
380 indicated that the major temporal change will be the increasing suitability of currently
381 unsuitable locations (Fig. 4 a-d). Approximately 6% of cells were projected to be unsuitable
382 for any of the 15 invasive species under current conditions, however, this decreased to <1%
383 by 2080. There was also a decrease in the number of cells projected to have just one species
384 (from *ca.* 20% to *ca.* 15% by 2080). Predicted changes appear to make areas currently
385 suboptimal (principally upland and northern areas) increasingly suitable over time. To a
386 lesser extent, it was projected that the number of cells with 3-4 species may increase by 2080.
387 Changes in hot- and coldspots of invasion were mainly due to the projected expansion of *A.*
388 *filiculoides* and *M. aquaticum* concomitant with a decline in *E. canadensis*. Projections only
389 took account of current invasive species and did not include potential invaders.

390

391 **(A) Discussion**

392 Our results illustrate how global climate niche models can be integrated with regional
393 environmental niche models to improve the precision of invasive species risk by accounting
394 for landscape factors which limit species establishment at local scales. Model AUC values
395 increased by 0.18 and TSS values by 0.36. The estimated suitable range of invasive aquatic
396 plants in Ireland was reduced by 50% of the total land area when global CNMs were
397 combined with regional ENMs accounting for land use, nutrient concentration, natural
398 landscape and human influence. However, global CNMs did show higher sensitivity (average
399 difference = 0.18) than regional ENMs. Conversely, the specificity of global CNMs was
400 lower than regional ENMs (average difference = 0.54). The ability of models to correctly
401 predict presence is of particular concern for invasive species, as failure to identify invaded
402 locations can cause problems for management and eradication. However, for management
403 purposes 'estimated suitability maps' with higher sensitivity could easily be achieved using
404 regional ENMs by changing the presence-absence threshold to achieve the desired sensitivity
405 level or using the minimum training presence threshold rather than the 10th percentile training
406 presence threshold. These options would increase the estimated suitable area and decrease
407 model specificity, reducing the risk of under-prediction but increasing the cost of survey and
408 intervention.

409 True evaluations of model performance are difficult during the invasion process as species
410 are not at equilibrium, and false-presences may be misleading if they represent areas that
411 have yet to be colonised (Peterson *et al.*, 2008). In addition, models based on data from
412 invaded ranges may underestimate suitable range of species, due to the fact that suitable areas
413 which have not yet been colonised are treated as pseudo-absences in model building. This
414 effect has been shown to be most evident in the early stages of invasion (Václavík &
415 Meentemeyer, 2012). Species in this study represent a range of invasion stages, with the

416 earliest species introduced in 1836 and the most recent in 1993. If stage of invasion were a
417 key factor causing the difference between estimated suitable ranges by global CNMs and
418 regional ENMs we would expect to see a smaller difference between range estimates of
419 species which had been introduced earlier than for those introduced more recently. No such
420 pattern was evident in this study and the two earliest introductions (*Elodea canadensis* and
421 *Nymphoides peltata*, both introduced in 1836) differed in range estimates between global
422 CNMs and regional ENMs by 58% and 33% respectively.

423 Alternatively, presences may occur in areas of marginal habitat quality where species
424 would be unable to reproduce, reach high abundance or survive in the long term. This may
425 lead to an overestimation of the potential impact of invasive species (Bradley, 2013).

426 Future projections have increased uncertainty due to the unpredictability of species
427 responses to novel combinations of landscape and climate, and uncertainties associated with
428 meteorological forecasting. Hence, uncertainties increase with distance into the future.
429 Moreover, as there are no test datasets available for future time periods it is difficult to
430 evaluate the accuracy of future projections (Elith & Leathwick, 2009). We recommend that
431 our models are re-evaluated and revised throughout the lifetime of their projections, for
432 example by using iterative survey methods (e.g. Stohlgren & Schnase, 2006) or by
433 incorporating novel data from physiological studies (Kearney & Porter, 2009).

434 Climate was the least important factor associated with the regional distribution of seven
435 out of eight aquatic invasive species in Ireland. This may be due to the low climatic variation
436 at regional scales characteristic of temperate maritime islands such as Ireland. The global
437 ranges of the majority of species examined were most strongly associated with climatic
438 tolerances set by mean annual temperature and minimum temperature of coldest month. It is
439 possible that a stronger association with climatic variables would emerge if variables specific

440 to freshwater habitats were used (e.g. max, min and mean water surface temperatures).
441 However, we are not aware of any such global databases.

442 Climate acts as an initial coarse filter which allows for the establishment of invasive
443 species regionally, and is a useful tool for assessing risk for species not currently present.
444 Here, global CNMs allowed us to distinguish between species for which climate suitability
445 was likely to increase considerably by 2080 (*E. densa*, *H. verticillata* and *L. grandiflora*) and
446 those for which there was a low risk of widespread establishment (*C. caroliniana*, *E.*
447 *crassipes* and *S. molesta*). Hence, global CNMs can be used to prioritise species for
448 preventative action such as legislation on sale and importation. However, in the case of these
449 latter three species, it remains possible that they will establish in Ireland as a result of
450 phenotypic adaptation to colder climates or through the introduction or evolution of novel
451 genotypes (Clements & Ditommaso, 2011; Shaw & Etterson, 2012).

452 At the regional scale, land use, nutrient concentration and natural landscape were the most
453 important factors associated with species ranges. However, the response of individual species
454 varied considerably. Similar species-specific responses to landscape have been noted
455 previously for many invasive terrestrial plants (e.g. Ibáñez *et al.*, 2009; Syphard & Franklin,
456 2009).

457 Human Influence Index was positively associated with the presence of all species except
458 *L. major*. Association with human activity has been reported before for invasive species and
459 aquatic plants (e.g. Buchan & Padilla, 2000; Mack *et al.*, 2000; Aznar *et al.*, 2003; Thuiller *et*
460 *al.*, 2005; Capers *et al.*, 2009) and probably reflects an increase in propagule pressure and
461 disturbance facilitating establishment. Risk projections could be improved by including
462 further data on known vectors e.g. boating and angling (Leung *et al.*, 2006)

463 Present ranges are likely to be restricted by invasion histories. In Ireland, all study species
464 occupied only a small fraction of the range that was projected to be at risk and substantial
465 expansion is expected over the next century, even in the absence of climate change.

466 Notably, the range of a high-latitude species (*E. canadensis*) was projected to contract
467 throughout the 21st century. Retreat in range and vigour of some species may provide
468 opportunities for the restoration of native plant communities (Bradley *et al.*, 2009). Ireland is
469 predicted to become generally more suitable for colonisation, resulting in fewer coldspots in
470 upland and northern parts of the island, which is of particular concern for upland oligotrophic
471 lakes as they currently represent an important refuge for rare native aquatic species
472 (Heegaard *et al.*, 2001).

473 Despite the potential limitations of ecological niche modelling and the uncertainties
474 inherent in future projections, the amalgamation of global CNMs and regional ENMs
475 provides a useful framework for the estimation of invasion risk. In the case of invasive
476 aquatic plants, this approach improved model performance in terms of AUC, TSS and
477 specificity and greatly reduced estimates of suitable ranges. Hence, integrated models may
478 provide more precise information about invasion risk at sub-regional scales and provide
479 useful information for species management and the prioritisation of resources.

480

481 **(A) Acknowledgements**

482 This research was funded by the Natural Heritage Research Partnership (NHRP) between the
483 Northern Ireland Environment Agency (NIEA) and *Quercus*, Queen's University Belfast
484 (QUB) under a PhD studentship (QU08-05). We thank the client officers, John Early and
485 Tony Waterman for their support. Thanks also to two anonymous reviewers and Dr Matthew
486 Fitzpatrick who provided useful feedback which substantially improved the manuscript.

487

488 (A) References

489

490 Alahuhta, J., Heino, J. & Luoto, M. (2011) Climate change and the future distributions of
491 aquatic macrophytes across boreal catchments. *Journal of Biogeography*, **38**, 383-393.

492

493 Allouche, O., Tsoar, A. & Ronen, K. (2006) Assessing the accuracy of species distribution
494 models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, **43**,
495 1223-1232.

496

497 Araújo, M. & Guisan, A. (2006) Five (or so) challenges for species distribution modelling.
498 *Journal of Biogeography*, **33**, 1677-1688.

499

500 Austin, M. (2007) Species distribution models and ecological theory: A critical assessment
501 and some possible new approaches. *Ecological Modelling*, **200**, 1-19.

502

503 Austin, M. (2011) Improving species distribution models for climate change studies: variable
504 selection and scale. *Journal of Biogeography*, **38**, 1-8.

505

506 Aznar, J-C., Dervieux, A. & Grillas, P. (2003) Association between aquatic vegetation and
507 landscape indicators of human pressure. *Wetlands*, **23**, 149-160.

508

509 Beaumont, L.J., Gallagher, R.V., Thuiller, W., O'Downey, P., Leishman, M.R. & Hughes, L.
510 (2009) Different climatic envelopes among invasive populations may lead to
511 underestimations of current and future biological invasions. *Diversity and Distributions*, **15**,
512 409-420.

513 Bossenbroek, J.M., Wagner, H.H. & Wiens, J.A. (2004) Taxon-dependent scaling: beetles,
514 birds, and vegetation at four North American grassland sites. *Landscape Ecology*, **20**, 675–
515 688.

516

517 Buchan, L.A. & Padilla, D.K. (2000) Predicting the likelihood of Eurasian Watermilfoil
518 presence in lakes, a macrophyte monitoring tool. *Ecological Applications*, **10**, 1442–1455.

519

520 Bradley, B.A., Openheimer, M., & Wilcove, D.S. (2009) Climate change and plant invasions:
521 restoration opportunities ahead? *Global Change Biology*, **15**, 1511-1521.

522

523 Bradley, B.A. (2013) Distribution models of invasive plants over-estimate potential impact.
524 *Biological Invasions*, **15**, 1417-1429

525

526 Capers, R.S., Selsky, R. Bugbee, G.J. & White, J.C. (2009) Species richness of both native
527 and invasive aquatic plants influenced by environmental conditions and human activity.
528 *Botany*, **87**, 306-314.

529

530 Capinha, C. & Anastácio, P. (2011) Assessing the environmental requirements of invaders
531 using ensembles of distribution models. *Diversity and Distributions*, **17**, 13–24.

532

533 Clements, D.R. & Ditommaso, A. (2011) Climate change and weed adaptation: can evolution
534 of invasive plants lead to greater range expansion than forecasted? *Weed Research*, **51**, 227-
535 240.

536

537 Compton, T.J., De Winton, M., Leathwick, J.R. & Wadhwa, S. (2012) Predicting spread of
538 invasive macrophytes in New Zealand lakes using indirect measures of human accessibility.
539 *Freshwater Biology*, **57**, 938-948.
540

541 EEA (2002) *CORINE land cover update, I&CLC2000 project, technical guidelines*.
542 European Environment Agency, Copenhagen.
543

544 Elith, J. & Leathwick, J.R. (2009) Species distribution models: ecological explanation and
545 prediction across space and time. *Annual Review of Ecology Evolution and Systematics*, **40**,
546 677-697.
547

548 Elith, J., Kearney, M. & Phillips, S.J (2010) The art of modelling range-shifting species.
549 *Methods in Ecology and Evolution*, **1**, 330-342.
550

551 Elith, J., Phillips, S.J., Hastie, T., Dudík, M., Chee, Y.E. & Yates, C. (2011) A statistical
552 explanation of MaxEnt for ecologists. *Diversity and Distributions*, **17**, 43-57.
553

554 Evangelista, P.H., Kumar, S., Stohlgren, T.J., Jarnevich, C.S., Crall, A.W., Norman III, J.B.
555 & Barnett, D.T. (2008) Modelling invasion for a habitat generalist and a specialist plant
556 species. *Diversity and Distributions*, **14**, 808-817.
557

558 Fournier, D.A., Skaug, H.J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M., Nielsen, A.
559 & Sibert, J. (2012). AD Model Builder: using automatic differentiation for statistical
560 inference of highly parameterized complex nonlinear models. *Optimization Methods and*
561 *Software*, **27**, 233-249.
562

563 Freeman, E.A. & Moisen, G. (2008). PresenceAbsence: An R package for presence-absence
564 model analysis. *Journal of Statistical Software*, **23**, 1-31.
565

566 Gardi, C., Panagos, P., Hiederer, R., Montanarella, L. & Micale, F. (2008) *Report on the*
567 *activities realized within the Service Level Agreement between JRC and EFSA as a support*
568 *of the FATE and ECOREGION Working Groups of EFSA PPR (SLA/EFSA-JRC/2008/01)*.
569 Luxembourg, Publications Office of the European Union.
570

571 Gaston, K.J. (2003) *The structure and dynamics of geographic ranges*. Oxford University
572 Press, Oxford.
573

574 Gillingham, P.K., Palmer, S.C.F., Huntley, B, Kunin, W.E., Chipperfield, J.D. & Thomas,
575 C.D. (2012) The relative importance of climate and habitat in determining distributions of
576 species at different spatial scales: a case study with ground beetles in Great Britain.
577 *Ecography*, **35**, 831-838.
578

579 Heegaard, E., Birks, H.H., Gibson, C.E., Smith, S.J. & Wolfe-Murphy, S. (2001) Species-
580 environmental relationships of aquatic macrophytes in Northern Ireland. *Aquatic Botany*,
581 **70**, 175-223.
582

583 Ibáñez, I., Silander, J.A, Allen, J., Treanor, S. & Wilson, A. (2009) Identifying hotspots for
584 plant invasions and forecasting focal points of further spread. *Journal of Applied Ecology*,
585 **46**, 1219-1228.
586

- 587 IPCC (2007) *Climate change 2007: the physical science basis. Contribution of Working*
588 *Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate*
589 *Change* (ed. by S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M.
590 Tignor and H.L. Miller). Cambridge University Press, Cambridge.
591
- 592 Jarnevich, C.S., Evangelista, P, Stohlgren, T.J. & Morisette, J. (2011) Improving national-
593 scale invasion maps: Tamarisk in the Western United States. *Western North American*
594 *Naturalist*, **71**, 164–175.
595
- 596 Jiménez-Valverde, A., Lobo, J.M. & Hortal, J. (2008) Not as good as they seem: the
597 importance of concepts in species distribution modelling. *Diversity and Distributions*, **14**,
598 885-890.
599
- 600 Jiménez-Valverde, A., Peterson, A.T., Soberón, J. Overton, J.M., Aragón, P. & Lobo, J.M.
601 (2011) Use of niche models in invasive species risk assessments. *Biological Invasions*, **13**,
602 2785-2797.
603
- 604 Kearney, M. & Porter, W. (2009) Mechanistic niche modelling: combining physiological and
605 spatial data to predict species' ranges. *Ecology letters*, **12**, 1-17.
606
- 607 Leung, B., Bossenbroek, J. M. & Lodge, D. M. (2006) Boats, pathways, and aquatic
608 biological invasions: estimating dispersal potential with gravity models. *Biological*
609 *Invasions*, **8**, 241-254.
610
- 611 Liu, X., Guo, Z., Ke, Z., Wang, S. & Li, Y (2011) Increasing potential risk of a global aquatic
612 invader in Europe in contrast to other continents under future climate change. *Plos One*, **6**,
613 e18429.
614
- 615 Lobo, J.M., Jiménez-Valverde, A. & Real, R. (2008) AUC: a misleading measure of the
616 performance of predictive distribution models. *Global Ecology and Biogeography*, **17**, 145-
617 151.
618
- 619 Mack, R.N., Simberloff, D., Lonsdale, W.M., Evans, H., Clout, M. & Bazzaz, F.A. (2000)
620 Biotic invasions: Causes, epidemiology, global consequences and control. *Ecological*
621 *Applications*, **10**, 689-710.
622
- 623 Milbau, A, Stout, J.C., Graae, B.J. & Nijs, I. (2009) A hierarchical framework for integrating
624 invasibility experiments incorporating different factors and spatial scales. *Biological*
625 *Invasions*, **11**, 941-950.
626
- 627 Minchin, D. (2007) A checklist of alien and cryptogenic aquatic species in Ireland. *Aquatic*
628 *Invasions*, **2**, 341-366.
629
- 630 Nielsen, C., Hartvig, P. & Kollmann, J. (2008) Predicting the distribution of the invasive
631 alien *Heracleum mantegazzianum* at two different spatial scales. *Diversity and*
632 *Distributions*, **14**, 307–317.
633
- 634 O'Donnell, J., Gallagher, R.V., Wilson, P.D., O'Downey, P., Hughes, L. & Leishman, M.R.
635 (2012) Invasion hotspots for non-native plants in Australia under current and future
636 climates. *Global Change Biology*, **18**, 617-629.

637
638 Peterson, A.T., Papes, M. & Soberon, J. (2008) Rethinking receiver operating characteristic
639 analysis applications in ecological niche modelling. *Ecological Modelling*, **213**, 63-72.
640
641 Phillips, S.J., Dudík, M., Elith, J., Graham, C.H., Lehmann, A., Leathwick, J. & Ferrier, S.
642 (2009) Sample selection bias and presence-only distribution models: implications for back-
643 ground and pseudo-absence data. *Ecological Applications*, **19**, 181–197.
644
645 Phillips, S., Dudík, M. & Schapire, R. (2010) “Maxent Software, ver. 3.3.3k”
646
647 Reshetnikov, A.N. & Ficetola G.F. (2011) Potential range of the invasive fish rotan
648 (*Perccottus glenii*) in the Holarctic. *Biological Invasions*, **13**, 2967-2980.
649
650 Shaw, R.G. & Etterson, J.R. (2012) Rapid climate change and the rate of adaptation: insight
651 from experimental quantitative genetics. *New Phytologist*, **195**, 752-65.
652
653 Stohlgren, T.J. & Schnase, J.L. (2006) Risk analysis for biological hazards: what we need to
654 know about invasive species. *Risk Analysis*, **26**, 163-173.
655
656 Syphard, A.D. & Franklin, J. (2009) Differences in spatial predictions among species
657 distribution modeling methods vary with species traits and environmental predictors.
658 *Ecography*, **32**, 907-918
659
660 Theoharides, K.A. & Dukes, J.S. (2007) Plant invasion across space and time: factors
661 affecting nonindigenous species success during four stages of invasion. *New Phytologist*,
662 **176**, 256-273.
663
664 Thuiller, W., Richardson, D.M., Pyšek, P., Midgley, G.F. Hughes, G.O. & Rouget, M. (2005)
665 Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global
666 scale. *Global Change Biology*, **11**, 2234-2250.
667
668 Thuiller, W., Richardson, D.M., Rouget, M., Proches, S. & Wilson, J.R.U. (2006) Interactions
669 between environment species traits and human uses describe patterns of plant invasions.
670 *Ecology*, **87**, 1755-1769.
671
672 Václavík, T. & Meentemeyer, R.K. (2012) Equilibrium or not? Modelling potential
673 distribution of invasive species in different stages of invasion. *Diversity and Distributions*,
674 **18**, 73-83.
675
676 Walther, G.R., Roques, A., Hulme, P.E., Sykes, M.T., Pysek, P., Kühn, I., Zobel, M., Bacher,
677 S., Botta-Dukát, Z., Bugmann, H., Czúcz, B., Dauber, J., Hickler, T., Jarošík, V., Kenis, M.,
678 Klotz, S., Minchin, D., Moora, M., Nentwig, W., Ott, J., Panov, V.E., Reineking, B.,
679 Robinet, C., Semchenko, V., Solarz, W., Thuiller, W., Vilà, M., Vohland, K. & Settele, J.
680 (2009) Alien species in a warmer world: risks and opportunities. *Trends in Ecology &*
681 *Evolution*, **24**, 686–693.
682
683 Wildlife Conservation Society (WCS); Center for International Earth Science Information
684 Network (CIESIN), Columbia University. (2005) Last of the Wild Project, Version 2, 2005
685 (LWP-2): Global Human Influence Index (HII) Dataset (Geographic). NASA
686 Socioeconomic Data and Applications Center (SEDAC), New York. Available at:

687 <http://sedac.ciesin.columbia.edu/data/set/wildareas-v2-human-influence-index-geographic>
688 (accessed 12 December 2011)

689

690 Yackulic, C.B., Chandler, R., Zipkin, E. F., Royle, J.A., Nichols, J.D., Campbell Grant, E.H.
691 & Veran, S. (2013) Presence-only modelling using MAXENT: when can we trust the
692 inferences? *Methods in Ecology and Evolution*, **4**, 236–243.

693

694 **(A) Biosketch**

695

696 Research group: *Quercus* is Northern Ireland's Centre for Biodiversity and Conservation
697 Science, which uses species distribution modelling as a tool for identifying areas of high
698 conservation value (www.quercus.ac.uk).

699

700 Author contributions: All authors contributed to the conceptualisation and development of
701 this work. Data collation and analysis was conducted by R.K. Advice on analysis was
702 provided by A.C. and N.R. The manuscript was drafted by R.K. and edited by K.L., C.A.M.
703 and N.R.

704

705 **Table 1.** Model evaluation metrics and percentage estimated suitable area for
 706 global Climate Niche Models (CMNs) and regional Environmental Niche Models
 707 (ENMs) of 8 invasive freshwater plants in Ireland. Generalised linear model
 708 results show differences between global CNMs and regional ENMs for each
 709 model evaluation metric.

710

Metric	Mean values (standard error.)		GLM result	
	Global CMNs	Regional ENMs	Wald	p
AUC	0.64 (0.01)	0.82 (0.01)	117.12	<0.001
Sensitivity	0.96 (<0.01)	0.78 (0.30)	9.97	0.002
Specificity	0.07 (<0.01)	0.61 (0.02)	725.75	<0.001
TSS	0.03 (<0.01)	0.39 (0.03)	180.12	<0.001
% estimated suitable area	86.5 (3.0)	35.5 (3.0)	1060.56	<0.001

711

712

713 **Figure legends**

714

715 **Fig. 1** Suitable climate range from global climate niche models (A2 = brown line, B2 = red
716 line) and suitable environmental niche (from regional environmental niche models including
717 local landscape data) (A2 scenario = dark blue line, B2 scenario = light blue line) for current
718 conditions, 2020, 2050 and 2080).

719

720 **Fig. 2** Permutation importance of predictor variables in global climate niche models for 15
721 invasive species. Boxplot shows 10th, 25th, 50th, 75th, 90th percentiles. Dotted lines indicate
722 mean values.

723

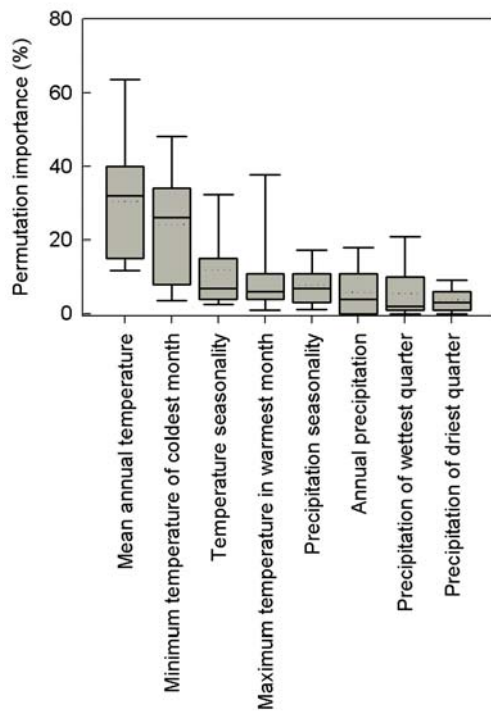
724 **Fig. 3** Permutation importance of predictor variables in environmental niche models for 8
725 invasive species grouped into descriptive categories. Boxplot shows 10th, 25th, 50th, 75th,
726 90th percentiles. Dotted lines indicate mean values.

727

728 **Fig. 4** Temporal changes in ‘hot- and coldspots’ of invasions under **a)** current conditions, **b)**
729 2020, **c)** 2050 and **d)** 2080. Future projections are shown for the A2 or high CO₂ scenario.
730 Colours indicate the number of invasive species. Histograms show the distributions of
731 invasive species richness.

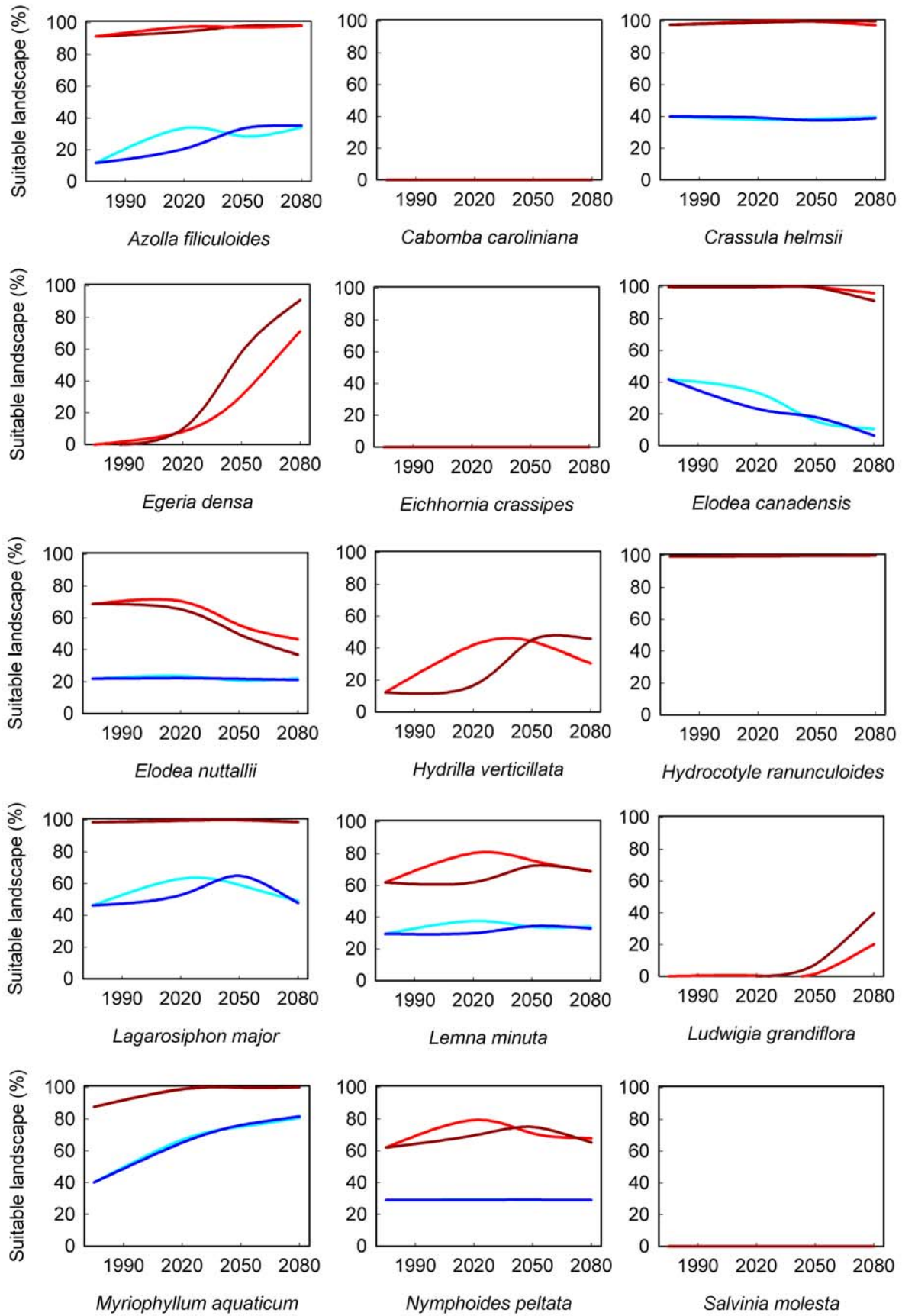
732

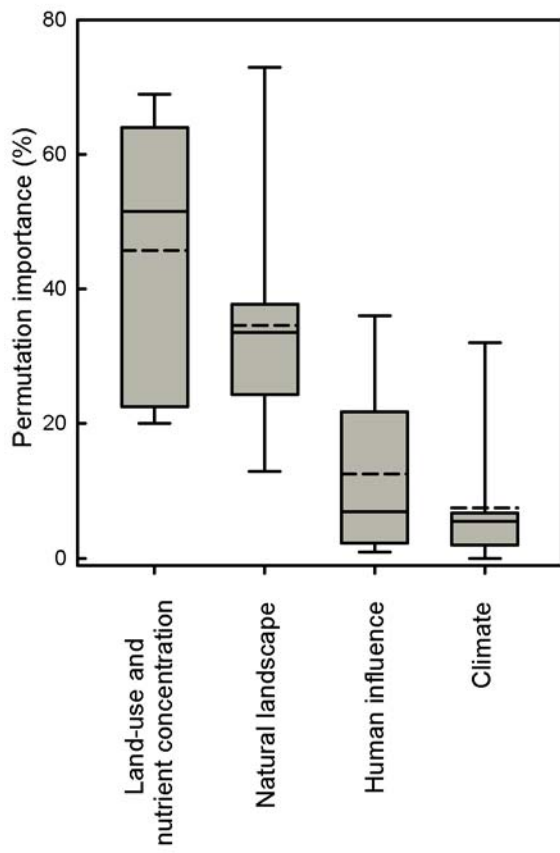
733 Fig 1.



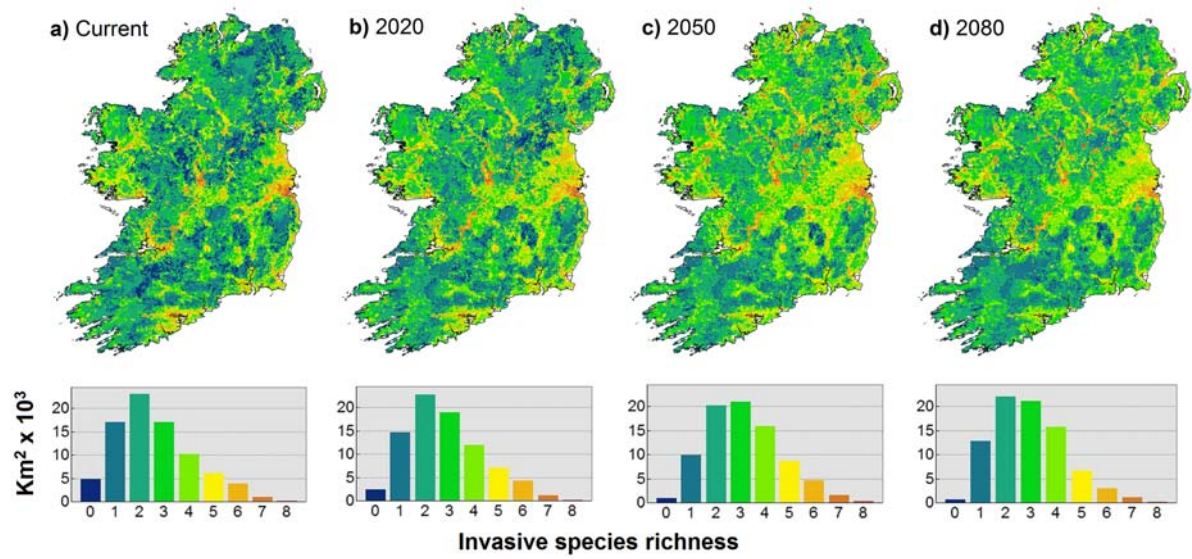
734

735





741 Fig. 4



742

Supplementary Information – supplied as separate word document

- Appendix S1** Data agencies by whom data on species distributions in the Republic of Ireland and Northern Ireland was contributed.
- Appendix S2** Optimal spatial scales of land use and human influence variables, showing Jackknife AUC value of the model with only that variable at that spatial scale.
- Appendix S3** Permutation importance of climatic variables in global climate niche models.
- Appendix S4** Permutation importance of climate, landscape, human influence index and natural landscapes in Irish species distribution models for 8 invasive aquatic plants already established in Ireland.
- Appendix S5** Projected ‘estimated suitable range’ of each species within Ireland based on regional environmental niche models.
- Appendix S6** Projected ‘estimated suitable range’ of each species within Ireland based on global climate niche models.

Appendix S1

Data agencies by whom data on species distributions in the Republic of Ireland and Northern Ireland was contributed.

- Botanical Society of the British Isles (BSBI)
- Centre for Environmental Data and Recording (CEDaR)
- Invasive Species Ireland (ISI)
- National Biodiversity Data Centre (NBDC)
- National Biodiversity Network (NBN)
- Northern Ireland Environment Agency (NIEA)

Appendix S2

Optimal spatial scales of land use and human influence variables; showing Jackknife AUC value of the model with only that variable at that spatial scale.

Species	Variable	Spatial scale (km ²)	Jackknife AUC
<i>Azolla filiculoides</i>	Arable land	10.5	0.550
	Broad leaved woodland	10.5	0.633
	Coniferous woodland	6.5	0.740
	Human impact	4.5	0.727
	Pastural farming	1.5	0.634
	Riparian length	20.5	0.647
	Standing freshwater	1.5	0.682
<i>Crassula helmsii</i>	Arable land	1.5	0.614
	Broad leaved woodland	1.5	0.517
	Coniferous woodland	6.5	0.679
	Human impact	20.5	0.552
	Pastural farming	1.5	0.737
	Riparian length	0.5	0.569
	Standing freshwater	4.5	0.562
<i>Elodea canadensis</i>	Arable land	10.5	0.623
	Broad leaved woodland	20.5	0.663
	Coniferous woodland	10.5	0.571
	Human impact	20.5	0.581
	Pastural farming	4.5	0.558
	Riparian length	1.5	0.714
	Standing freshwater	2.5	0.707
<i>Elodea nuttallii</i>	Arable land	10.5	0.649
	Broad leaved woodland	10.5	0.625
	Coniferous woodland	6.5	0.656
	Human impact	0.5	0.846
	Pastural farming	0.5	0.664
	Riparian length	0.5	0.787
	Standing freshwater	1.5	0.803
<i>Lagarosiphon major</i>	Arable land	0.5	0.557
	Broad leaved woodland	10.5	0.624
	Coniferous woodland	4.5	0.627
	Human impact	20.5	0.624
	Pastural farming	1.5	0.778
	Riparian length	1.5	0.867
	Standing freshwater	1.5	0.819
<i>Lemna minuta</i>	Arable land	10.5	0.527
	Broad leaved woodland	10.5	0.687
	Coniferous woodland	6.5	0.589
	Human impact	4.5	0.692
	Pastural farming	6.5	0.645
	Riparian length	1.5	0.660
	Standing freshwater	4.5	0.673
<i>M. aquaticum</i>	Arable land	4.5	0.636
	Broad leaved woodland	10.5	0.661
	Coniferous woodland	4.5	0.632
	Human impact	1.5	0.770
	Pastural farming	10.5	0.713
	Riparian length	20.5	0.701
	Standing freshwater	20.5	0.560
<i>Nymphoides peltata</i>	Arable land	20.5	0.679
	Broad leaved woodland	10.5	0.703
	Coniferous woodland	20.5	0.624
	Human impact	10.5	0.813
	Pastural farming	1.5	0.711
	Riparian length	0.5	0.707
	Standing freshwater	4.5	0.656

Appendix S3

Permutation importance of climatic variables in global climate niche models.

Species	N	Mean annual temperature	Temperature seasonality	Maximum temperature (in hottest month)	Minimum temperature (in coldest month)	Annual precipitation	Precipitation seasonality	Maximum precipitation (in wettest month)	Minimum precipitation (in driest month)
<i>A. filiculoides</i>	2,035	15	4	1	56	7	11	0	6
<i>C. caroliniana</i>	260	40	5	10	15	0	10	14	6
<i>C. helmsii</i>	879	21	27	8	14	18	2	5	6
<i>E. densa</i>	570	16	7	57	7	1	4	1	8
<i>E. crassipes</i>	1,261	34	7	19	3	3	21	3	11
<i>E. canadensis</i>	4474	58	4	2	26	6	4	1	0
<i>E. nuttallii</i>	3,070	13	13	11	32	11	0	20	0
<i>H. verticillata</i>	1,935	33	11	4	4	15	11	22	0
<i>H. ranunculoides</i>	449	14	39	1	27	0	15	1	3
<i>L. major</i>	572	10	28	5	24	18	4	10	1
<i>L. minuta</i>	1,839	32	3	4	43	0	12	0	6
<i>L. grandiflora</i>	508	18	6	25	34	4	11	1	2
<i>M. aquaticum</i>	1,139	46	8	7	30	4	2	1	3
<i>N. peltata</i>	1,810	34	15	6	40	0	3	2	1
<i>S. molesta</i>	122	72	2	5	8	0	7	2	4
Mean	1,395	30	12	11	24	6	8	6	4

Appendix S4

Permutation importance of climate, landscape, human influence index and natural landscapes in Irish species distribution models for 8 invasive aquatic plants already established in Ireland.

Species	Climate	Land use and eutrophication					Human influence	Natural landscape			
	Predicted probability from global models	Arable	Broad-leaved woodland	Coniferous plantations	Pastoral farming	Eutrophication Orthophosphate	Human Influence Index	Riparian length	Standing freshwater	Soil Type	Soil pH
<i>A. filiculoides</i>	32	3	2	11	7	4	6	10	20	4	0
<i>C. helmsii</i>	2	24	0	12	11	22	6	4	7	12	1
<i>E. canadensis</i>	7	0	12	3	4	1	1	53	20	0	0
<i>E. nuttallii</i>	5	4	6	3	8	0	36	14	23	0	1
<i>L. major</i>	6	13	1	10	20	17	1	12	4	12	5
<i>L. minuta</i>	2	14	10	25	3	1	8	16	8	13	0
<i>M. aquaticum</i>	6	6	5	29	7	3	21	7	4	12	2
<i>N. peltata</i>	0	22	17	4	1	21	22	5	2	5	1
Mean	8	11	7	12	8	9	13	15	11	7	1

Appendix S5

Predicted 'suitable range' within Ireland based on regional environmental niche models.

Species	N	Current recorded range (No. of 500m cells (%))	Current suitable range (%)	B2 'low emmissions'			A2 'high emmissions'		
				Difference in suitable range			Difference in suitable range		
				2020	2050	2080	2020	2050	2080
<i>A. filiculoides</i>	110	116 (0.03)	12	+9	+22	+23	+22	+17	+22
<i>C. helmsii</i>	13	20 (0.01)	40	-1	-3	-1	-2	-2	0
<i>E. canadensis</i>	1,646	1659 (0.50)	42	-19	-24	-35	-8	-26	-31
<i>E. nuttallii</i>	104	112 (0.03)	22	+0	0	-1	+2	-1	0
<i>L. major</i>	77	84 (0.03)	46	+7	+19	+2	+17	+13	+3
<i>L. minuta</i>	98	116 (0.03)	29	+0	+5	+3	+8	+4	+4
<i>M. aquaticum</i>	18	28 (<0.01)	40	+25	+36	+42	+27	+35	+41
<i>N. peltata</i>	29	31 (<0.01)	29	0	0	0	0	0	0
Mean	262	271 (0.08)	32	+3	+7	+4	+8	+5	+5