

Title	Incorporating host-parasite biotic factors in species distribution models: Modelling the distribution of the Castor Bean tick, <i>Ixodes ricinus</i>
Authors	McDonough, Sinead;Holloway, Paul
Publication date	2020-11-01
Original Citation	McDonough, S. and Holloway, P. (2020) 'Incorporating host-parasite biotic factors in species distribution models: Modelling the distribution of the Castor Bean tick, <i>Ixodes ricinus</i> ', Irish Geography, 53(2), pp. 105-125. doi: 10.2014/igj.v53i2.1416
Type of publication	Article (peer-reviewed)
Link to publisher's version	<a href="http://irishgeography.ie/index.php/irishgeography/article/view/1416/1158">http://irishgeography.ie/index.php/irishgeography/article/view/1416/1158</a> - 10.2014/igj.v53i2.1416
Rights	© 2020, Geographical Society of Ireland. Published by Routledge – Taylor & Francis Group. All rights reserved. This is an Accepted Manuscript of an item published by Taylor & Francis in Irish Geography in November 2020, available online: <a href="https://dx.doi.org/10.2014/igj.v53i2.1416">https://dx.doi.org/10.2014/igj.v53i2.1416</a>
Download date	2025-05-16 05:19:58
Item downloaded from	<a href="https://hdl.handle.net/10468/11147">https://hdl.handle.net/10468/11147</a>



**UCC**

**University College Cork, Ireland**  
 Coláiste na hOllscoile Corcaigh

1 **Incorporating host-parasite biotic factors in species distribution models: Modelling the**  
2 **distribution of the castor bean tick, *Ixodes ricinus*.**

3 **Abstract**

4 Understanding where ticks are found, and the drivers of their geographic distributions is  
5 imperative for successful epidemiological precautions. Predictive models of tick distributions  
6 are often projected using solely abiotic (e.g., climate) variables, despite the strong biotic  
7 interaction that host species undoubtedly have with parasitic species. We used species  
8 distribution modelling to project the distribution of *Ixodes ricinus* in Ireland and the United  
9 Kingdom using different combinations of abiotic, biotic, and abiotic-biotic variables. We found  
10 that models parameterised solely on abiotic variables generally reported lower accuracy and  
11 ecological realism than models that incorporated biotic factors alongside climate. We also  
12 investigated representation of host distribution in models, testing four different methods  
13 (habitat suitability of individual hosts, presence-absence of individual hosts, ensembled habitat  
14 suitability, and ensembled presence-absence). Biotic representations of ensembled host  
15 distributions alongside abiotic variables reported the highest accuracy, with the variable  
16 representing host diversity (e.g., number of host species) the most important variable when  
17 measured using a jackknife test. Moreover, our results suggested how host distributions are  
18 represented (i.e., presence-absence, habitat suitability) greatly impacted results, with  
19 differences reported among habitat specialists and generalists. Results suggest that it is now  
20 imperative for projections of parasitic species to include a representation of biotic factors with  
21 host species. This research has improved our understanding of the drivers of tick distributions  
22 in a national context, and the investigation of biotic representation should foster discussion  
23 among researchers working in species distribution modelling and the wider biogeography  
24 discipline.

25 **Keywords:** biotic interactions; host species; parasitism; species distribution modelling; ticks

26

## 27 **Introduction**

28 Ticks are one of the most important global vectors in the transmission of disease (Kelly *et al.*  
29 2001), known to carry a wider variety of infectious agents compared to other arthropods  
30 (Soenshine 1991). Within Europe *Ixodes ricinus* is the most common arthropod vector (Parola  
31 & Raoult, 2001), with the ability to transmit various pathogens, including *Borrelia burgdorferi*  
32 (Lyme disease - Pietzsch *et al.* 2005). Global incidence rates of Lyme disease vary, but most  
33 reported cases are found in Europe and North America, with research indicating increasing  
34 incidence rates (Mead 2018). In Ireland, varied incidence and infection rates across both space  
35 and time have been reported (Gray *et al.* 1996; McKeown and Garvey 2009; Cullen 2010;  
36 Lambert *et al.* 2019), with the crude incidence rate of Lyme neuroborreliosis progressively  
37 increasing towards the south and west of the country (HSE 2019). Reports in the UK have  
38 identified an increase from 693 cases in 2005 to 1,310 cases in 2016 (Lorenc *et al.* 2017). Due  
39 to increasing incidence rates and the spatial variability in reports, understanding where ticks  
40 are found, and the drivers of their geographic distributions is imperative for successful  
41 epidemiological precautions.

42 In Ireland and the UK studies have predominantly focused on reporting the presence and  
43 density of tick species at a national scale using fieldwork and location maps (Martyn 1988;  
44 Kirstein *et al.* 1997; Kelly *et al.* 2001; Dobson *et al.* 2011a; Medlock and Leach, 2015; Zintl *et al.*  
45 2017). For example, both Kirstein *et al.* (1997) and Kelly *et al.* (2001) produced location  
46 maps indicating presence of tick species from field studies in Ireland, while in the UK, Martyn  
47 (1988) published the first atlas of tick distributions, with the tick surveillance scheme continued  
48 by Public Health England. Most recently, Zintl *et al.* (2017) reported 151 tick observations  
49 from 26 sites across Ireland. Such studies improve our understanding of the environmental  
50 drivers of tick abundance and densities, but are often restricted spatially to specific locations.

51 Species distribution models (SDMs) are one of the most important GIScience research areas in  
52 biogeography, and a powerful spatial ecological tool for studying the geographic distribution  
53 of plants, animals, and other taxa (Franklin 2009; Peterson *et al.* 2011). The modelling  
54 framework provides a methodology for researchers and practitioners to quantitatively assess  
55 the relationship between species distributions and environmental factors, having been widely  
56 used for various applications including predicting disease spread (Peterson *et al.* 2004; Crowl  
57 *et al.* 2008; Bhatt *et al.* 2013). In Europe, tick SDMs have primarily been undertaken at a  
58 continental scale. For example, Alkische *et al.* (2017) modelled the current and future (2050 and

59 2070) distribution of *I. ricinus* based on several climatic factors and identified that distributions  
60 are expected to increase in area under all future climate scenarios. Similarly, Williams *et al.*  
61 (2015) identified a shift to higher latitudes for eight European tick species for 2050 and 2098.  
62 Several other global studies have undertaken distribution projections for ticks at regional (De  
63 Clercq *et al.* 2015; Raghavan *et al.* 2016; St John *et al.* 2016), continental (Springer *et al.* 2015;  
64 Rubel *et al.* 2016; Hahn *et al.* 2016; Sage *et al.* 2017), and even global (Alkishe *et al.* 2020)  
65 scales; all identifying the importance of climatic factors, with many outlining the potential risk  
66 associated with increased human exposure to such a prevalent disease vector.

67 Climatic factors are an important determinant of species distributions; however, they are not  
68 the sole driver. The ‘BAM’ framework (Figure 1), developed by Soberón and Peterson (2005),  
69 illustrates the individual and joint effects of the three most important factors in determining a  
70 species distribution. Biotic (**B**) factors represent interactions with other species (e.g.,  
71 competition, parasitism), abiotic (**A**) factors represent the physiological tolerances of the  
72 species (e.g., temperature, precipitation), and movement (**M**) factors represent the ability of a  
73 species to access the habitats (e.g., dispersal, foraging). The importance of all three factors is  
74 well recognised, yet in the majority of the aforementioned studies projecting tick distributions,  
75 models were parameterised on only abiotic factors. The fact that ticks have a wide range of  
76 hosts, which often lack good-quality occurrence data, as well as uncertainty in how best to  
77 incorporate such data into model parameterisation has been cited as rationale for focusing on  
78 climatic factors when projecting tick distributions (Alkishe *et al.* 2020). Moreover, abiotic  
79 factors have often superseded biotic factors in SDMs due in part to the fact that abiotic factors  
80 often influence species at a broader spatial extent (Wisz *et al.* 2013; Miller and Holloway  
81 2015). Despite this, distribution studies that have included biotic factors have observed  
82 increases in model accuracy (Araújo and Luoto 2007; Heikkinen *et al.* 2007; Kissling *et al.*  
83 2008; Wisz *et al.* 2013; Raath *et al.* 2018), meaning distribution models could be improved by  
84 explicitly incorporating biotic factors.

85 [Figure 1 near here]

86 Studies have emerged that incorporate both abiotic and biotic factors when projecting tick  
87 distributions, often resulting in improved accuracy and ecological realism (Medlock *et al.*  
88 2013; Donaldson *et al.* 2016; Estrada-Peña and de la Fuente 2017). For example, in a study  
89 predicting *Ornithodoros turicata* distribution in the continental USA, Donaldson *et al.* (2016)  
90 overlaid the predicted tick distribution with the distributions of 58 host species, and identified

91 areas of significant overlap that could identify principal hosts. Similarly, Estrada-Peña and de  
92 la Fuente (2017) incorporated probabilities of occurrence of host species in their tick  
93 distribution models, finding significant tick-host relationships among the complex  
94 communities of host species. The importance of incorporating biotic factors as covariates in  
95 any statistical models projecting species distributions is well established; however, it is perhaps  
96 more pertinent for parasite species that have such a strong dependence on hosts (Wisz *et al.*  
97 2013). Recent reviews (Wisz *et al.* 2013; Dormann *et al.* 2018; Blanchet *et al.* 2020) have  
98 outlined several challenges associated with incorporating biotic factors within SDMs, one of  
99 which is the method through which the biotic factors are represented. Raath *et al.* (2018)  
100 recently compared four methods of incorporating biotic factors (host plant distribution) in  
101 SDMs projecting the distribution of the African silk moth in sub-Saharan Africa. They found  
102 that the method with which biotic factors were represented in the statistical model (e.g.,  
103 presence-absence of individual hosts, habitat suitability of individual hosts, combined  
104 presence-absence of all hosts, and combined habitat suitability of all hosts) greatly affected the  
105 predictive ability of the models, yet there was inconsistency among representations, suggesting  
106 a need for further research.

107 With the incorporation of biotic factors (i.e. host distributions) as environmental covariates in  
108 statistical models shown to improve projections of tick distributions, coupled with the fact that  
109 several tick SDM studies still neglect host distribution altogether, there persists a need to  
110 investigate the methods of representing biotic factors within SDM, and how their incorporation  
111 impacts the current geographic predictions of ticks. In this study, we investigated the role of  
112 biotic (**B**) and abiotic (**A**) factors in determining the distribution of *I. ricinus* in Ireland and the  
113 UK focusing on two specific research questions: 1) How do different ‘BAM’ scenarios (e.g.,  
114 **A**, **B**, **BA**) affect the accuracy and ecological realism of *I. ricinus* projections in Ireland and the  
115 UK? And 2) How do different representations of host species (e.g., **B**) affect the accuracy and  
116 ecological realism of the projections?

## 117 **Methodology**

### 118 ***Data Collection***

119 We acquired 172 primary occurrence records for *I. ricinus* from the Global Biodiversity  
120 Information Facility (GBIF 2020) across Ireland and the UK since 2000. Following GBIF data  
121 cleaning recommendations (Chapman 2005), incomplete points and duplicates were removed,  
122 resulting in 132 occurrence records for use in the statistical models. Data for 16 primary host

123 species were also obtained from GBIF (2019). Climate data was obtained from WorldClim  
124 (Fink and Hijmans *et al.*, 2017) and resampled using cubic convolution to 1000m. Corine land  
125 cover (EEA 2018) data was used to calculate the percentage of arable, forest, grass, pasture,  
126 urban, and water cover within a corresponding 1000m resolution.

### 127 ***Species Distribution Modelling***

128 SDMs generate a prediction of habitat suitability in both environmental and geographic space.  
129 The most common response variables used are data representing species presence-absence or  
130 presence-only, which are coupled with a set of corresponding and georeferenced environmental  
131 variables. Species-environment relationships are then estimated using a range of statistical  
132 approaches. These species-environment relationships can be extrapolated in space and time  
133 generating a habitat suitability map that identifies the probability that the environmental  
134 conditions at that location are favourable for the study species. To create a categorical  
135 representation of habitat suitability, a threshold that delineates between presence-absence can  
136 be selected, with habitat suitability values above and below this predicted as present and absent,  
137 respectively. The choice of response data, environmental data, and statistical models are all  
138 related to hypotheses derived from the underlying niche concept. As outlined above, the choice  
139 of environmental variables can have substantial implications on model output. Here, we  
140 investigated three different ‘BAM’ scenarios. Each scenario incorporates different  
141 methodologies to derive environmental variables based on the ‘BAM’ diagram (Figure 1).  
142 Figure 2 provides a methodological workflow of the SDM steps used in this study.

143 Maximum entropy (MaxEnt) was chosen as the statistical method to model the species-  
144 environment relationships for all scenarios (Phillips *et al.* 2006). MaxEnt selects suitable  
145 environmental variables by measuring how well they delineate between the recorded presence  
146 and pseudo-absence observations. The method chooses the split in environmental variables  
147 (e.g., temperature) that records the purest split between values in the binary response variable  
148 (e.g., presence and pseudo-absence). This method has been found to outperform other  
149 presence-only methods in an extensive comparison study (Elith and Graham 2009). Moreover,  
150 MaxEnt is robust to variable collinearity in model training and the removal of highly correlated  
151 variables prior to model parameterisation has little impact due to the fact the algorithm removes  
152 redundant variables (Feng *et al.* 2019), meaning MaxEnt is a robust method to evaluate  
153 different combinations of multiple environmental variables. However, the potential for  
154 spurious results and overfitting still exists and results need to be underpinned with a

155 mechanistic understanding related to their contribution in determining tick distributions  
156 (Holloway *et al.* 2018). To do this, we implemented the jackknife test to measure variable  
157 importance and explore the impact of variable representation on results. In the final MaxEnt  
158 model, the maximum number of iterations was set to 5,000 to allow model convergence; the  
159 number of pseudo-absences was set at 10,000 (following Barbet-Massin *et al.* 2012), and the  
160 model incorporated only linear and quadratic features to avoid over-fitting.

161 To account for unequal survey coverage in the species data, each SDM was fit with a bias grid  
162 to control for any violation in the assumption of independence in the response data that would  
163 increase the likelihood of type I errors (Fourcarde *et al.* 2014). A bias grid is the equivalent of  
164 a sampling probability surface, where the cell values reflect sampling effort and provides a  
165 weight to the pseudo-absence data used in modelling. Following Elith *et al.* (2010), a Gaussian  
166 kernel density map of the occurrence locations within 10km was derived and divided by the  
167 weighted number of terrestrial cells in the neighbourhood in order to avoid the edge effects at  
168 the coastline. This was then rescaled so that values varied between 1 and 20 in order to control  
169 for the impact of extreme variation. Bias grids are presented in Supplementary Information 1  
170 to provide an overview of the spatial bias in occurrence records.

171 [Figure 2 near here]

## 172 ***Scenario A***

173 Abiotic factors represent the physiological tolerances of species (i.e. temperature,  
174 precipitation) meaning climate variables best represent these factors. WorldClim v2.1 has  
175 interpolations of observed data from ~1970-2000, with 19 bioclimatic variables derived from  
176 the monthly temperature and rainfall values (Fick and Hijmans 2017). The 15 abiotic variables  
177 used by Alkische *et al.* (2017) were used within this study to allow for a direct comparison with  
178 recent European continental predictions, with variables 8-9 and 18-19 removed due to spatial  
179 artefacts. See Table 1 for variable code and description.

180 [Insert Table 1 near here]

## 181 ***Scenario B***

182 Biotic factors represent interactions with other species (i.e. parasitism, competition). Ticks are  
183 dependent on host species, meaning in this study, the biotic factors consisted of the  
184 distributions of sixteen mammalian hosts. Of the 27 host species listed by Kelly *et al.* (2001),  
185 nine were excluded due to their domestication which would prevent reliable habitat suitability

186 predictions. Humans were excluded as a host species, due to the obvious clustering in urban  
187 areas, and two other species (goat, mink) were excluded due to inadequate observations in  
188 GBIF (2019). The final list of mammalian species consisted of; hedgehog, pygmy shrew,  
189 rabbit, grey squirrel, red squirrel, bank vole, wood mouse, brown rat, fox, pine marten, stoat,  
190 badger, fallow deer, red deer, sika deer, and hare. The distribution of these sixteen species was  
191 predicted using the observed locations from GBIF (2019) since 2000, the fifteen abiotic  
192 variables (Fink and Hijmans 2017), the land cover data (EEA 2018), and projected using  
193 MaxEnt with the same settings and bias grids outlined above.

194 Recent research has indicated different predictive accuracies in using presence-absence of host  
195 species compared to habitat suitability (Raath *et al.* 2018). Therefore, four methods of  
196 incorporating biotic factors were investigated (Table 2). Both habitat suitability (hs.ind) and  
197 presence-absence (pa.ind) of all sixteen species were incorporated as individual explanatory  
198 variables, as well as an ensemble model (i.e., average or summation of multiple projections) of  
199 both habitat suitability (hs.ens) and presence-absence (pa.ens). Thresholds for generating  
200 binary presence-absence models from habitat suitability were calculated using maximum  
201 sensitivity plus specificity.

202 [Insert Table 2 near here]

### 203 ***Scenario BA***

204 This scenario projects the potential distribution in the BAM framework. Both abiotic (i.e., the  
205 fifteen bioclimatic variables) and biotic (i.e., distributions of the sixteen host species) variables  
206 were used as the input variables in the statistical model. Again, all four methods of  
207 incorporating biotic factors were included in different iterations of the statistical model.

### 208 ***Model Evaluation***

209 Model evaluation in SDM focuses on quantifying prediction accuracy to ascertain model  
210 validity. To test our models, we withheld 20% of the 132 tick occurrence records from training  
211 the statistical model for testing as is common practice in SDM research when independent test  
212 data are lacking (Franklin 2009). This testing dataset was then used to evaluate the performance  
213 of our models. We used three accuracy metrics that are best suited for presence-only data;  
214 lowest possible threshold (LPT), minimum predicted area (MPA) and the Boyce Index (BI).  
215 LPT is the value that results in zero omission errors, with higher values indicating a better  
216 model. MPA represents the area encompassing a predefined proportion of observed species



217 occurrences (in this case 90 percent), with a lower value (measured in geographic area)  
218 representing a more parsimonious model (Engler et al. 2004). The BI implements a moving  
219 window analysis across the predicted values, using Spearman rank correlation coefficient to  
220 measure the monotonic increase in the predicted-to-expected frequency ratio with increasing  
221 habitat suitability (Boyce *et al.* 2002). This method was undertaken using the *ecospat* package  
222 (Broenniman *et al.* 2014) in R 3.6.1 (R Development Core Team 2008) using default settings,  
223 with values ranging from -1 to 1, with higher values representing more parsimonious models,  
224 and 0 indicating predictions indifferent from a random model.

## 225 **Results**

226 The projected tick distributions and subsequent model accuracies varied substantially across  
227 the different implementations of **A**, **B**, and **BA** (Figure 3, Table 3). The model parameterised  
228 only on **A** (Figure 3a) reported high probability (>0.5) for the south and west of Ireland, as well  
229 as Wales, the south, southwest and northwest of England, and eastern Scotland, with low  
230 probability (<0.5) of occurrence for the remainder of Ireland and eastern England. Models  
231 parameterised including representations of host species (**B**, **BA**) consistently projected the  
232 southwest of the UK as suitable, with much wider spatial variation elsewhere. When the  
233 percentage contribution and permutation importance of the host species was explored in models  
234 that incorporated biotic factors (Table 4), the distribution of sika deer and bank voles were  
235 continually high across all models (10-40%), as well as contributing to large gains in the final  
236 models (Supplementary Information 2), suggesting the distribution of these host species may  
237 be an important driver in determining tick distributions and supporting the high probability  
238 within the southwest regions of the study area. For **B** (pa.ens) and **B** (hs.ens) where only the  
239 ensembled representation of biotic factors were used in model parameterisation, clear  
240 overfitting of the data occurred, with large swaths of the study area projected as suitable (Figure  
241 3d-e). The **BA** models appear to have much more localised areas of high suitability (>0.5)  
242 across all counties in Ireland and the UK, which most likely reflects a more ecologically  
243 realistic distribution.

244 [Table 3 near here]

245 [Figure 3 near here]

246 Certain species were reported as contributing heavily to the projections when measured using  
247 one representation, but not the other (Table 4, Supplementary Information 2). For example,  
248 rabbit and hare distributions had higher gains for models when measured using habitat

249 suitability compared to presence-absence (Supplementary Information 2), and contributed  
250 more to final models when represented as habitat suitability, with rabbits contributing 0% to  
251 the final models when represented as presence-absence (Table 4). Alternatively, red deer  
252 distributions contributed more when measured using presence-absence compared to habitat  
253 suitability (Table 4).

254 There was discrepancy in model validity when measured using the different accuracy metrics,  
255 with LPT, MPA, and BI indicating the most accurate models as B (hs.ens), BA (hs.ens), and  
256 BA (pa.ens) respectively (Table 3). LPT rewards correct prediction of presences, often  
257 resulting in a high commission error (i.e. high false positive rate), which aligns with the  
258 observed overfitting for models B (pa.ens) and B (hs.ens) (Figure 3d-e). MPA and BI both  
259 attempt to control omission errors (i.e. false negatives), with the BI considered the more robust  
260 metric due to its ability to measure the monotonic increase in the predicted-to-expected  
261 frequency ratio with increasing habitat suitability (Hirzel *et al.* 2006). Models BA (hs.ens) and  
262 BA (pa.ens) both reported BI values above 0.9 and low MPA values suggesting a combination  
263 of abiotic variables with only one variable representing host diversity reports the most valid  
264 projection of tick distributions. Jackknife testing was used to evaluate variable importance to  
265 the final models, with the ensemble representations of host species reporting the highest gain  
266 in model performance when used in isolation for the BA (pa.ens) parameterisations and the  
267 largest decrease in gain in model performance when removed from both the BA (pa.ens) and  
268 BA (hs.ens) parameterisations (Figure 4).

269 [Figure 4 near here]

270 [Table 4 near here]

## 271 **Discussion**

272 The aim of this study was to investigate the role of abiotic and biotic factors in determining  
273 tick distributions in Ireland and the UK, comparing methods of incorporating biotic variables  
274 into SDMs. From examining the models (Figure 3), it is evident that there are certain regions  
275 that are more suitable for *I. ricinus* when the roles of **B**, **A**, and **BA** are compared. The most  
276 suitable areas in Ireland appear to be in the south, west, southeast, and north, while in the UK  
277 the most suitable areas are the south-west region of England, Wales, and western Scotland,  
278 although we did identify high probabilities for suitable habitat in every county of Ireland and  
279 the UK (Figure 3). However, these areas do vary depending on the choice of environmental  
280 variables incorporated in the model. Previous continental scale projections of *I. ricinus* have

281 identified the importance of abiotic variables in determining both current and future  
282 distributions (e.g., Alkische *et al.* 2017), with our result identifying the importance of mean  
283 annual temperature and precipitation of the driest month as important variables (Figure 4, Table  
284 4, Supplementary Information 2); however, the results of this study also highlight the need to  
285 incorporate factors related to the distribution of host species.

286 The relatively lower accuracy of *I. ricinus* projections for models parameterised solely on  
287 abiotic (**A**) factors compared to both biotic and abiotic (**BA**) factors (Table 3) suggests that  
288 abiotic factors alone do not adequately capture the drivers of tick distributions. Due to the  
289 strong dependence of ticks on host species, we expected that tick distributions would exhibit  
290 positive relationships with host distributions. The presence of large mammals (e.g., sika deer,  
291 red deer) were often important in the final entropy models, while smaller mammals (e.g., bank  
292 voles) were also shown to be important to the relative percentage contribution and regularised  
293 training gain in the final models (Table 4, Supplementary Information 2). Improvements in  
294 model accuracy (Table 3) and ecological realism (Figure 3) corroborate other studies that have  
295 included explicit variables measuring biotic factors (e.g., Kissling *et al.* 2008; Donaldson *et al.*  
296 2016), and our results suggest that to accurately and realistically ascertain the distribution of  
297 parasitic species, variables that account for such relationships must be included.

298 The models parameterised on the different representations of **B** varied greatly, highlighting the  
299 need for continued research into how such conceptualisations can affect results. We found that  
300 BA (pa.ens) was the most accurate when the BI was used as the accuracy metric (Table 3),  
301 corroborating the findings of Raath *et al.* (2018) who also found that a combined presence-  
302 absence map was the most accurate representation of **B** in their study of African silk moths.  
303 This variable represents the total number of projected host species in that grid (i.e., equivalent  
304 of host species richness). Subsequently, this model may outperform others due to its ability to  
305 incorporate information on all host species, rather than assuming dependence on only one or a  
306 handful of species. The areas predicted as highly suitable by this model (Figure 3e) correspond  
307 with areas that were projected to support 12-16 host species. Subsequently, a variable  
308 representing the host diversity in the model potentially captures the parasite-host relationships  
309 more realistically than individual habitat suitability maps of all host species or an average  
310 habitat suitability of all hosts. Moreover, when each variable was tested for its importance to  
311 the final model through the jackknife test, the ensemble representations of host species reported  
312 the highest gain in model performance when used in isolation for BA (pa.ens)  
313 parameterisations, and the largest decrease in gain in model performance when removed from

314 both the BA (pa.ens) and BA (hs.ens) parameterisations (Figure 4). Tick interactions with  
315 mammals such as sika deer (Kimura *et al.* 1995; Gray *et al.* 1999; Braticikov *et al.* 2019), red  
316 deer (Gray *et al.* 1999; Zintl *et al.* 2011; Razanske *et al.* 2019), foxes (Lappin 2016; D'Amico  
317 *et al.* 2017; Sándor *et al.* 2017) and small rodents (Cayol *et al.* 2017; Cull *et al.* 2018) are well  
318 established, meaning the dependence of *I. ricinus* on multiple host species is best represented  
319 through a measure of diversity and distribution rather than as a series of individual host  
320 distributions.

321 Unsurprisingly many host distributions were continuously important in final models, both  
322 through the combined representation of host diversity, but also when included as individual  
323 species (Table 4, Supplementary Information 2). Sika deer and bank voles were consistently  
324 reported as important variables in final models, with these species exhibiting a strong positive  
325 relationship with forest cover (results not shown). The importance of vegetation cover for tick  
326 distributions and abundance is well established (Dobson *et al.* 2011a, 2011b), meaning there is  
327 the potential that our models confounded the importance of host distribution with land cover.  
328 To investigate this, we parameterised a model using the abiotic variables and land cover  
329 (Supplementary Information 3). This model performed poorly (LPT 0.299, MPA 150787, BI  
330 0.727) compared to models parameterised with host distributions, highlighting the importance  
331 of considering the full ecological niche requirements of host species when projecting tick  
332 distributions.

333 Moreover, the relative contribution and subsequent regularised training gain of biotic variables  
334 altered for individual species depending on whether the variable was represented as continuous  
335 habitat suitability or categorical presence-absence (Table 4, Supplementary Information 2).  
336 The preference for binary representations appeared to occur for habitat specialist species (e.g.,  
337 red deer, wood mouse, bank vole, stoat), suggesting that delineating the most suitable habitat  
338 from lesser habitat into presence-absence improves projections of parasite-host relationships.  
339 A continuous representation favoured more habitat generalists (e.g., grey squirrel, rat, rabbit),  
340 in part due to their ubiquitous distributions, meaning binary representations predicted them as  
341 present throughout the study area. Therefore, for generalist species a continuous variable that  
342 captures nuances in the habitat suitability improves models when compared to a thresholded  
343 presence-absence map, although generalist species continually reported low contributions and  
344 gains to final models (Table 4, Supplementary Information 2). This should have important  
345 considerations for future SDM research when study species may only have one or two primary  
346 biotic interactions determining their potential distributions.

347 Our results highlight the substantial impact of user decisions in how biotic factors are  
348 represented (Tables 3-4, Figures 3-4), and subsequently we have refrained from making  
349 grandiose statements considering our projections to be the definitive distribution of *I. ricinus*  
350 in Ireland and the UK. Research using such approaches for epidemiologically important vector  
351 species must be aware of such uncertainties in model parameterisations and strive to  
352 incorporate or quantify these in projections (see Alkishe *et al.* 2020 for a comprehensive  
353 assessment of model uncertainty projecting global tick distributions). Similarly, while spatial  
354 bias in response data was controlled for by using a bias grid (Supplementary Information 1),  
355 the tick data (GBIF 2020) does vary spatially with a wider coverage of the UK (particularly  
356 England and Scotland) and a lower number of observations in Ireland and Wales, with a higher  
357 reporting of occurrence in coastal regions (i.e. four observations on the Aran Islands alone) .  
358 Future research collecting verified tick locations that are made available through open data  
359 repositories is needed to improve model calibration and tick predictions, and supports recent  
360 calls for more comprehensive surveys, particularly in Ireland (Zintl *et al.* 2017). Despite these  
361 caveats, we do report high accuracy values (Table 3), allowing confidence in the discussion of  
362 the important environmental drivers of *I. ricinus* distributions.

363 The importance of considering environmental variables, particularly in the context of the  
364 'BAM' framework (Figure 1) is therefore imperative for distribution studies; however, we did  
365 not incorporate movement factors (**M**) in our models. **M** refers to the area that has been or will  
366 be accessible to a species within a certain timeframe. The majority of SDM research only  
367 incorporates movement when investigating response to climate change (Franklin 2010;  
368 Holloway *et al.* 2016; Holloway and Miller, 2017). Recent research on tick distributions  
369 incorporating movement has identified the role of temporal variation in mammal movements  
370 resulting from land cover, and that such movements can have implications for tick distributions  
371 (Martin *et al.* 2018). Similarly, Halsey and Miller (2018) developed a spatial agent-based  
372 model to explore host-tick interactions, highlighting the possibility to combine such approaches  
373 and better inform on host-tick movements and interactions. Such approaches combining SDM  
374 and agent-based models are beginning to emerge that could create finer scale distribution  
375 models that can account for all three 'BAM' factors (e.g., Holloway 2018), meaning future  
376 research should continue to explore the inter-linked relationships between biotic, abiotic, and  
377 movement variables.

378 Finally, there persists a need to explore habitat suitability of *I. ricinus* in conjunction with other  
379 measures of biodiversity and epidemiology. It is important to note that ground feeding species

380 of birds (e.g., blackbirds, redwings) have been found to be important hosts for *Ixodes* ticks,  
381 with some rural populations reportedly having infestation rates as high as 74% (Gregoire *et al.*  
382 2002; Singh and Girshick 2003). We opted only to investigate the role of mammalian host  
383 species; however, future research should look to incorporate information on avian distributions  
384 when projecting the geographic ranges of ticks. Similarly, we should also note that high habitat  
385 suitability does not necessarily indicate presence of breeding populations of ticks or Lyme  
386 disease. For example, Talleklint and Jaenson (1996) found a significant relationship between  
387 tick density and infection rates (of Lyme disease) at lower tick densities, while Gray *et al.*  
388 (1999) found that tick infection rates were inversely related to tick abundance. Dobson *et al.*  
389 (2011a) also found spatial variability in the abundance of ticks across recreational areas in the  
390 UK, recording uniform presence across vegetation types, but higher densities in plots which  
391 contained trees. Healy and Bourke (2004) note that clusters of ticks and larvae should be used  
392 to identify breeding populations, which contradicts the predominant method in SDM of using  
393 binary representation of presence-absence. New methods, such as Poisson point process  
394 models, are emerging as suitable tools to predict the distribution and abundance of species from  
395 presence-only data (Warton and Shepherd, 2010; Renner and Warton 2013; Schank *et al.*  
396 2017), and given the uncertain relationships identified, future studies should focus on  
397 developing predictions of both the distribution and density of ticks to inform epidemiological  
398 research.

### 399 **Conclusion**

400 Understanding the drivers of tick distributions is imperative for successful epidemiological  
401 precautions. Previous research has tended to focus only on the role of abiotic variables in  
402 determining tick distributions, and while research is beginning to emerge that highlights the  
403 importance of biotic variables (e.g., host distributions), there persists a need for the combined  
404 influence of **A** and **B** to be investigated together, as well as exploration into how best to  
405 represent biotic factors within SDM. Large differences in accuracy (Table 3) and area (Figures  
406 3) were identified depending on the combination of ‘BAM’ variables used during model  
407 parameterisation. Representations of **B** varied substantially with results suggesting that a  
408 variable of host diversity captures the parasite-host relationships most accurately when coupled  
409 with abiotic factors (**BA**). The representation of host distribution also appeared to differ for  
410 habitat specialists and generalists (Table 4, Supplementary Information 2), suggesting use of a  
411 binary presence-absence conceptualisation and a continuous habitat suitability  
412 conceptualisation for specialist and generalist host species, respectively. There is now a

413 minimum requirement for any research into tick distributions to explicitly incorporate biotic  
414 factors, as excluding the host-parasite relationship masks the true distribution of species and  
415 does not capture an ecological realistic distribution of a species' habitat suitability.

#### 416 **Abstract**

417 We would like to thank the anonymous reviewers and Ronan Foley for their comments and  
418 suggestions.

#### 419 **References**

420 Alkische, A., Townsend Peterson, A., & Samy, A. (2017). Climate change influences on the  
421 potential geographic distribution of the disease vector tick *Ixodes ricinus*. *PLoS ONE*, 12 (12)

422 Alkische, A., Cobos, M.E., Townsend Peterson, A., & Samy, A.M. (2020). Recognising  
423 sources of uncertainty in disease vector ecological niche models: An example with the tick  
424 *Rhipicephalus sanguineus sensu lato*. *Perspectives in Ecology and Conservation*.

425 Araújo, M.B. and Luoto, M., (2007). The importance of biotic interactions for modelling  
426 species distributions under climate change. *Global Ecology and Biogeography*, 16(6),  
427 pp.743-753.

428 Barbet-Massin, M., F. Jiguet., C.H. Albert., and W. Thuiller. (2012). Selecting pseudo-  
429 absences for species distribution models: how, where and how many?. *Methods in Ecology*  
430 *and Evolution*, 3(2): 327-338.

431 Bhatt, S., Gething, P.W., Brady, O.J., Messina, J.P., Farlow, A.W., Moyes, C.L., Drake, J.M.,  
432 Brownstein, J.S., Hoen, A.G., Sankoh, O. and Myers, M.F., (2013). The global distribution and  
433 burden of dengue. *Nature*, 496(7446), p.504.

434 Blanchet, F.G., Cazelles, K., Gravel, D. (2020). Co-occurrence is not evidence of ecological  
435 interactions. *Ecology Letters*, 23: 1050-1063.

436 Boyce, M.S., P.R. Vernier., S.E. Nielsen., and F.K. Schmiegelow. (2002). Evaluating  
437 resource selection functions. *Ecological modelling*, 157(2): 281-300.

438 Bratickov, M., Paulauskas, A., Rosef, O., Ražanskė, I., Gričiuvienė, L. and Radzijeuskaja, J.,  
439 (2019). Prevalence and co-infection with tick-borne *Anaplasma phagocytophilum* and *Babesia*

440 spp. in red deer (*Cervus elaphus*) and roe deer (*Capreolus capreolus*) in Southern  
441 Norway. *International journal for parasitology: parasites and wildlife*. Amsterdam: Elsevier  
442 Ltd, 2019, Vol. 8.

443 Broenniman, O., B. Petitpierre., C. Randin., R. Engler., F. Breiner., M. D'Amen., L. Pellissier.,  
444 J. Pottier., D. Pio., R. Garcia Mateo, *et al.* (2014). ecospat: Spatial ecology miscellaneous  
445 methods. R package version 1.0.

446 Cayol, C., Koskela, E., Mappes, T., Siukkola, A. and Kallio, E.R., (2017). Temporal dynamics  
447 of the tick *Ixodes ricinus* in northern Europe: epidemiological implications. *Parasites &*  
448 *vectors*, 10(1), p.166.

449 Chapman A. D. (2005) Principles and Methods of Data Cleaning - Primary Species and  
450 Species-Occurrence Data. version 1.0. Report for the Global Biodiversity Information Facility,  
451 Copenhagen. Available online at [http://www.gbif.org/orc/?doc\\_id=1262](http://www.gbif.org/orc/?doc_id=1262)

452 Cowl, T.A., Crist, T.O., Parmenter, R.R., Belovsky, G. and Lugo, A.E., (2008). The spread of  
453 invasive species and infectious disease as drivers of ecosystem change. *Frontiers in Ecology*  
454 *and the Environment*, 6(5), pp.238-246.

455 Cull, B., Pietzsch, M.E., Hansford, K.M., Gillingham, E.L. and Medlock, J.M., (2018).  
456 Surveillance of British ticks: An overview of species records, host associations, and new  
457 records of *Ixodes ricinus* distribution. *Ticks and tick-borne diseases*, 9(3), pp.605-614.

458 Cullen, E. (2010). Lyme disease and climate change. *Irish Medical Journal*. 103:1, 1-2.

459 D'Amico, G., Juránková, J., Tăbăran, F.A., Frgelecová, L., Forejtek, P., Matei, I.A., Ionică,  
460 A.M., Hodžić, A., Modrý, D. and Mihalca, A.D., (2017). Occurrence of ticks in the  
461 subcutaneous tissue of red foxes, *Vulpes vulpes* in Czech Republic and Romania. *Ticks and*  
462 *tick-borne diseases*, 8(2), pp.309-312.

463 De Clercq, E.M., Leta, S., Estrada-Peña, A., Madder, M., Adehan, S. and Vanwambeke, S.O.,  
464 (2015). Species distribution modelling for *Rhipicephalus microplus* (Acari: Ixodidae) in Benin,  
465 West Africa: Comparing datasets and modelling algorithms. *Preventive veterinary*  
466 *medicine*, 118(1), pp.8-21.

467 Dobson., A.D.M., Taylor, J.L., & Randolph, S.E. (2011a). Ticks (*Ixodes ricinus*) abundance  
468 and seasonality at recreational sites in the UK: Hazards in relation to fine-scale habitat types  
469 revealed by complimentary sampling methods. *Ticks and Tick-borne Diseases*, 2: 67-74.



470 Dobson, A.D.M., Finnie, T.J.R., & Randolph, S.E. (2011b). A modified matrix model to  
471 describe the seasonal population ecology of the European tick *Ixodes ricinus*. *Journal of*  
472 *Applied Ecology*, 48: 1017-1028.

473 Donaldson, T.G., de Leon, A.A.P., Li, A.I., Castro-Arellano, I., Wozniak, E., Boyle, W.K.,  
474 Hargrove, R., Wilder, H.K., Kim, H.J., Teel, P.D. and Lopez, J.E., (2016). Assessment of the  
475 geographic distribution of *Ornithodoros turicata* (Argasidae): climate variation and host  
476 diversity. *PLoS neglected tropical diseases*, 10(2), p.e0004383.

477 Dormann, C.F., Bobrowski, M., Dehling, D.M., Harris, D.J., Hartig, F., Lischke, H., Moretti,  
478 M.D., Pagel, J., Pinkert, S., Schleuning, M. and Schmidt, S.I., (2018). Biotic interactions in  
479 species distribution modelling: 10 questions to guide interpretation and avoid false  
480 conclusions. *Global ecology and biogeography*, 27(9), pp.1004-1016.

481 EEA. (2018). "Corine Land Cover (CLC) 2018, Version 20b2." *Release Date: 21-12-*  
482 *2018*. European Environment Agency. [https://land.copernicus.eu/pan-european/corine-land-](https://land.copernicus.eu/pan-european/corine-land-cover/clc2018)  
483 [cover/clc2018](https://land.copernicus.eu/pan-european/corine-land-cover/clc2018)

484 Elith, J., and C.H. Graham. (2009). Do they? How do they? WHY do they differ? On finding  
485 reasons for differing performances of species distribution models. *Ecography*, 32(1): 66-77.

486 Elith, J., M. Kearney., and S. Phillips. (2010). The art of modelling range-shifting species.  
487 *Methods in Ecology and Evolution*, 1: 330–342

488 Engler, R., Guisan, A. and Rechsteiner, L., 2004. An improved approach for predicting the  
489 distribution of rare and endangered species from occurrence and pseudo-absence data. *Journal*  
490 *of applied ecology*, 41: 263-274.

491 Estrada-Peña, A. and de la Fuente, J., (2017). Host distribution does not limit the range of the  
492 tick *Ixodes ricinus* but impacts the circulation of transmitted pathogens. *Frontiers in Cellular*  
493 *and Infection Microbiology*, 7, p.405.

494 Feng, X., Park, D.S., Liang, Y., Pandey, R. and Papeş, M., 2019. Collinearity in ecological  
495 niche modeling: Confusions and challenges. *Ecology and evolution*, 9: 10365-10376.

496 Fourcade, Y., J.O. Engler., D. Rödder., and J. Secondi. (2014) Mapping Species Distributions  
497 with MAXENT Using a Geographically Biased Sample of Presence Data: A Performance  
498 Assessment of Methods for Correcting Sampling Bias, *PLoS One*, 9(5): e97122.

499 Franklin, J. (2009). *Mapping Species Distributions: Spatial Inference Prediction*. Cambridge:  
500 Cambridge University Press.

501 Franklin, J. (2010). Moving beyond static species distribution models in support of  
502 conservation biogeography. *Diversity and Distributions*, 16 (3): 321-330.

503 GBIF (20120) GBIF Occurrence Download <https://doi.org/10.15468/dl.zkqf26>

504 GBIF (2019) GBIF Occurrence Download <https://doi.org/10.15468/dl.wxxwh9>

505 Gray, J.S., Cutler, S., Robertson, J. & O'Connell, S. (1996). Lyme disease in the Republic of  
506 Ireland, p.6. In 'Who Workshop on Lyme Borreliosis diagnosis and surveillance'. Warsaw,  
507 Poland 20<sup>th</sup> – 22<sup>nd</sup> June 1995.

508 Gray, J.S., Kirstein, F., Robertson, J.N., Stein, J. and Kahl, O., (1999). *Borrelia burgdorferi*  
509 *sensu lato* in *Ixodes ricinus* ticks and rodents in a recreational park in south-western  
510 Ireland. *Experimental & applied acarology*, 23(9), pp.717-729.

511 Gregoire, A., Faivre, B., Heeb, P. and Cezilly, F., (2002). A comparison of infestation patterns  
512 by *Ixodes* ticks in urban and rural populations of the Common Blackbird *Turdus*  
513 *merula*. *Ibis*, 144(4), pp.640-645.

514 Hahn, M.B., Jarnevich, C.S., Monaghan, A.J. and Eisen, R.J., (2016). Modeling the geographic  
515 distribution of *Ixodes scapularis* and *Ixodes pacificus* (Acari: Ixodidae) in the contiguous  
516 United States. *Journal of medical entomology*, 53(5), pp.1176-1191.

517 Halsey, S.J. and Miller, J.R., (2018). A spatial agent-based model of the disease vector *Ixodes*  
518 *scapularis* to explore host-tick associations. *Ecological modelling*, 387, pp.96-106.

519 Healy, J.A.E. and Bourke, P., (2004). Field evidence for aggregating behaviour in the tick  
520 *Ixodes ricinus* L. *Acarologia*, 44(1-2), pp.3-13.

521 Heikkinen, R.K., Luoto, M., Virkkala, R., Pearson, R.G. and Körber, J.H., (2007). Biotic  
522 interactions improve prediction of boreal bird distributions at macro-scales. *Global Ecology*  
523 *and Biogeography*, 16(6), pp.754-763.

524 Hirzel, A.H., G. Le Lay., V. Helfer., C. Randin., and A. Guisan. 2006. Evaluating the ability  
525 of habitat suitability models to predict species presences. *Ecological Modelling*, 199(2): 142-  
526 152.

527 Holloway, P., (2018). Simulating movement-related resource dynamics to improve species  
528 distribution models: a case study with oilbirds in Northern South America. *The Professional*  
529 *Geographer*, 70(4), pp.528-540.

530 Holloway, P., and Miller. J.A., (2017). A quantitative synthesis of the movement concepts  
531 used in species distribution modelling. *Ecological Modelling*, 356: 91-103.

532 Holloway, P., Kudenko, D., and Bell, J. (2018). Dynamic selection of environmental  
533 variables to improve the prediction of aphid phenology: A machine learning approach.  
534 *Ecological Indicators*, 88; 512-521.

535 Holloway, P., Miller, J.A. and Gillings, S., (2016). Incorporating movement in species  
536 distribution models: how do simulations of dispersal affect the accuracy and uncertainty of  
537 projections?. *International Journal of Geographical Information Science*, 30(10), pp.2050-  
538 2074.

539 HSE (2019). Primary Prevention and Surveillance of Lyme Borreliosis in Ireland. Report of  
540 the Lyme Borreliosis Subcommittee of the Scientific Advisory Committee of the Health  
541 Protection Surveillance Centre. Available at: [https://www.hpsc.ie/a-](https://www.hpsc.ie/a-z/vectorborne/lymedisease/informationforhealthcareprofessionals/)  
542 [z/vectorborne/lymedisease/informationforhealthcareprofessionals/](https://www.hpsc.ie/a-z/vectorborne/lymedisease/informationforhealthcareprofessionals/). Accessed on 29<sup>th</sup> June  
543 2020.

544 Kelly, T. C., Sleeman, D.P., Fennessy, G.J., Dillon, A., Walton, G.A. & Gray, J. (2001). The  
545 mammal ticks (Acari: Ixodoidea) of Ireland: Their distribution and hosts. *Irish National*  
546 *Journal*. 26:3, 63 – 70.

547 Kimura, K., Isogai, E., Isogai, H., Kamewaka, Y., Nishikawa, T., Ishii, N. and Fujii, N., (1995).  
548 Detection of Lyme disease spirochetes in the skin of naturally infected wild sika deer (*Cervus*  
549 *nippon yesoensis*) by PCR. *Appl. Environ. Microbiol.*, 61(4), pp.1641-1642.

550 Kirsten, F., Rijpkema, S., Molkenboer, M. & Gray, J.S., (1997). The distribution and  
551 prevalence of *B. burgdorferi* genomospecies in *Ixodes ricinus* ticks in Ireland. *European*  
552 *Journal of Epidemiology*. 13, 67.

553 Kissling, W.D., R. Field., and K. Böhning-Gaese. (2008). Spatial patterns of woody plant and  
554 bird diversity: functional relationships or environmental effects?. *Global Ecology and*  
555 *Biogeography*, 17(3): 327-339.

556 Lambert, J.S., Cook, M.J., Healy, J.E., Murtagh, R., Avramovic, G. and Lee, S.H., (2019).  
557 Metagenomic 16S rRNA gene sequencing survey of *Borrelia* species in Irish samples of *Ixodes*  
558 *ricinus* ticks. *PloS one*, 14(4).

559 Lappin, J.W.E., (2016). *Ixodid ticks: vectors for emerging diseases in Northern*  
560 *Ireland* (Doctoral dissertation, Queen's University Belfast).

561 Lorenc, T., Jones-Diette, J., Blanchard, L., Sutcliffe, K., Stansfield, C., Harden, M., Sowden,  
562 A. and Thomas, J., (2017). *Incidence and surveillance of Lyme disease*. London: EPPI-Centre,  
563 Social Science Research Unit, UCL Institute of Education, University College London.

564 Martin, J., Vourc'h, G., Bonnot, N., Cargnelutti, B., Chaval, Y., Lourtet, B., Goulard, M.,  
565 Hoch, T., Plantard, O., Hewison, A.M. and Morellet, N., (2018). Temporal shifts in landscape  
566 connectivity for an ecosystem engineer, the roe deer, across a multiple-use  
567 landscape. *Landscape ecology*, 33(6), pp.937-954.

568 Martyn, K.P. (1988) *Provisional Atlas of The Ticks (Ixodoidea) of the British Isles*. Biological  
569 Records Centre Institute of Terrestrial Ecology, Abbots Ripton, Huntingdon.

570 McKeown, P. & Garvey, P. (2009). Lyme disease often under diagnosed says HPSC. *Epi-*  
571 *Insight*. 10, 1 – 5.

572 Mead, P.S., (2018). Global Epidemiology of *Borrelia burgdorferi* Infections. *Lyme Disease:*  
573 *an Evidence-based Approach*, p.76.

574 Medlock, J.M. and Leach, S.A., (2015). Effect of climate change on vector-borne disease risk  
575 in the UK. *The Lancet Infectious Diseases*, 15(6), pp.721-730.

576 Medlock, J.M., Hansford, K.M., Bormane, A., Derdakova, M., Estrada-Peña, A., George, J.C.,  
577 Golovljova, I., Jaenson, T.G., Jensen, J.K., Jensen, P.M. and Kazimirova, M., 2013. Driving  
578 forces for changes in geographical distribution of *Ixodes ricinus* ticks in Europe. *Parasites &*  
579 *vectors*, 6:1.

580 Miller, J. A., and P. Holloway. (2015). Incorporating movement in species distribution models.  
581 *Progress in Physical Geography*, 39(6): 837-849.

582 Parola, P. & Raoult, D. (2001). Tick-borne bacterial diseases emerging in Europe. *Clinical*  
583 *Microbiology and Infection*. 7, 80-83.

584 Peterson, A.T., Bauer, J.T. and Mills, J.N., (2004). Ecologic and geographic distribution of  
585 filovirus disease. *Emerging infectious diseases*, 10(1), p.40.

586 Peterson, A. T., J. Soberón., R.G. Pearson., R.P. Anderson., E. Martínez-Meyer., M.  
587 Nakamura., and M.B. Araújo. (2011). Ecological niches and geographic distributions (MPB-  
588 49). Princeton University Press.

589 Phillips, S. J., Anderson, R.P., & Schapire, R. E. (2006) Maximum entropy modelling of  
590 species geographic distributions. *Ecological Modelling*, 190, 231-259

591 Pietzsch, M.E., Medlock, J.M, Jones, L., Avenell, D., Abbott, J., Harding, P. & Leach, S.  
592 (2005). Distribution of *Ixodes ricinus* in the British Isles: Investigation of historical  
593 records. *Medical and Veterinary Entomology*. 19, 306-314.

594 R Development Core Team (2008). R: A language and environment for statistical computing.  
595 Vienna, Austria: R Foundation for Statistical Computing.

596 Raath, M.J., Le Roux, P.C., Veldtman, R. and Greve, M., (2018). Incorporating biotic  
597 interactions in the distribution models of African wild silk moths (Gonometa species,  
598 Lasiocampidae) using different representations of modelled host tree distributions. *Austral*  
599 *ecology*, 43(3), pp.316-327.

600 Razanske, I., Rosef, O., Radzijeuskaja, J., Bratchikov, M., Gričiuviene, L. and Paulauskas, A.,  
601 (2019). Prevalence and co-infection with tick-borne *Anaplasma phagocytophilum* and *Babesia*  
602 spp. in red deer (*Cervus elaphus*) and roe deer (*Capreolus capreolus*) in Southern  
603 Norway. *International Journal for Parasitology: Parasites and Wildlife*, 8, pp.127-134.

604 Renner, I.W. and Warton, D.I., (2013). Equivalence of MAXENT and Poisson point process  
605 models for species distribution modeling in ecology. *Biometrics*, 69(1), pp.274-281.

606 Raghavan, R.K., Goodin, D.G., Hanzlicek, G.A., Zolnerowich, G., Dryden, M.W., Anderson,  
607 G.A. and Ganta, R.R., (2016). maximum entropy-based ecological niche model and bio-  
608 climatic determinants of lone star tick (*Amblyomma americanum*) niche. *Vector-Borne and*  
609 *Zoonotic Diseases*, 16(3), pp.205-211.

610 Rubel, F., Brugger, K., Pfeffer, M., Chitimia-Dobler, L., Didyk, Y.M., Leverenz, S., Dautel,  
611 H. and Kahl, O., (2016). Geographical distribution of *Dermacentor marginatus* and  
612 *Dermacentor reticulatus* in Europe. *Ticks and tick-borne diseases*, 7(1), pp.224-233.

613 Sage, K.M., Johnson, T.L., Teglas, M.B., Nieto, N.C. and Schwan, T.G., (2017). Ecological  
614 niche modeling and distribution of *Ornithodoros hermsi* associated with tick-borne relapsing  
615 fever in western North America. *PLoS neglected tropical diseases*, 11(10), p.e0006047.

616 Sándor, A.D., D'Amico, G., Gherman, C.M., Dumitrache, M.O., Domşa, C. and Mihalca, A.D.,  
617 (2017). Mesocarnivores and macroparasites: altitude and land use predict the ticks occurring  
618 on red foxes (*Vulpes vulpes*). *Parasites & vectors*, 10(1), p.173.

619 Schank, C.J., Cove, M.V., Kelly, M.J., Mendoza, E., O'Farrill, G., Reyna-Hurtado, R., Meyer,  
620 N., Jordan, C.A., González-Maya, J.F., Lizcano, D.J. and Moreno, R., (2017). Using a novel  
621 model approach to assess the distribution and conservation status of the endangered Baird's  
622 tapir. *Diversity and Distributions*, 23(12), pp.1459-1471.

623 Singh, S.K. and Girschick, H.J., (2003). Tick-host interactions and their immunological  
624 implications in tick-borne diseases. *Current science*, 85(9), pp.1284-1298.

625 Soberón, J. & Peterson, A. (2005). Interpretation of models of fundamental ecological niches  
626 and species' distributional areas. *Biodiversity Informatics*. 2:1-10

627 Soberón, J. (2007). Grinnellian and Eltonian niches and geographic distributions of species.  
628 *Ecology letters*, 10 (12): 1115-1123.

629 Soenshine, D.E. (1991) *Biology of ticks*. Oxford University Press, New York

630 Springer, Y.P., Jarnevich, C.S., Barnett, D.T., Monaghan, A.J. and Eisen, R.J., (2015).  
631 Modeling the present and future geographic distribution of the Lone Star Tick, *Amblyomma*  
632 *americanum* (Ixodida: Ixodidae), in the continental United States. *The American journal of*  
633 *tropical medicine and hygiene*, 93(4), pp.875-890.

634 St. John, H.K., Adams, M.L., Masuoka, P.M., Flyer-Adams, J.G., Jiang, J., Rozmajzl, P.J.,  
635 Stromdahl, E.Y. and Richards, A.L., (2016). Prevalence, distribution, and development of an  
636 ecological niche model of *Dermacentor variabilis* ticks positive for *Rickettsia*  
637 *montanensis*. *Vector-Borne and Zoonotic Diseases*, 16(4), pp.253-263.

638 Tälleklint, L. and Jaenson, T.G.T. (1996). Relationship between *Ixodes ricinus* density and  
639 prevalence of infection with *Borrelia*-like spirochetes and density of infected ticks. *Journal of*  
640 *Medical Entomology*. 33, 805–81

641 Warton, D.I. and Shepherd, L.C., (2010). Poisson point process models solve the “pseudo-  
642 absence problem” for presence-only data in ecology. *The Annals of Applied Statistics*, 4(3),  
643 pp.1383-1402.

644 Williams, H. W., Cross, D. E., Crump, H. L., Drost, C. J., & Thomas, C. J. (2015). Climate  
645 suitability for European ticks: assessing species distribution models against null models and  
646 projection under AR5 climate. *Parasites & vectors*, 8(1), 440.

647 Wisz, M. S., J. Pottier., W.D. Kissling., L. Pellissier., J. Lenoir., C.F. Damgaard., C.F.  
648 Dormann., M.C. Forchhammer., J.A. Grythnes., A. Guisan., *et al.* (2013). The role of biotic  
649 interactions in shaping distributions and realised assemblages of species: implications for  
650 species distribution modelling. *Biological Reviews*, 88(1): 15-30.

651 Zintl, A., Finnerty, E.J., Murphy, T.M., de Waal, T. and Gray, J.S., (2011). Babesias of red  
652 deer (*Cervus elaphus*) in Ireland. *Veterinary research*, 42(1), p.7.

653 Zintl, A., Moutailler, S., Stuart, P., Paredis, L., Dutraive, J., Gonzalez, E., O’Connor, J.,  
654 Devillers, E., Good, B., O’Muireagain, C. and De Waal, T., (2017). Ticks and Tick-borne  
655 diseases in Ireland. *Irish veterinary journal*, 70(1), p.4.

656

657

658 **Tables**

659 Table 1. Variable information for the 19 bioclimatic variables acquired from Fick and Hijmans  
660 (2017).

Code	Variable
Bio1	Annual mean temperature
Bio2	Mean diurnal range
Bio3	Isothermality
Bio4	Temperature seasonality
Bio5	Maximum temperature of warmest month
Bio6	Minimum temperature of coldest month
Bio7	Temperature annual range
Bio8	Mean temperature of wettest quarter
Bio9	Mean temperature of driest quarter
Bio10	Mean temperature of warmest quarter
Bio11	Mean temperature of coldest quarter
Bio12	Annual precipitation
Bio13	Precipitation of wettest month
Bio14	Precipitation of driest month
Bio15	Precipitation seasonality
Bio16	Precipitation of wettest quarter
Bio17	Precipitation of driest quarter
Bio18	Precipitation of warmest quarter
Bio19	Precipitation of coldest quarter

661



662 Table 2. Description of the biotic representations used within parameterised models

Biotic Representation	Explanation
hs.ind	Individual maps of habitat suitability for all 16 host species
pa.ind	Individual maps of presence-absence for all 16 host species
hs.ens	Ensemble map of habitat suitability projections. The mean habitat suitability of all 16 host species
pa.ens	Ensemble map of presence-absence projections. Sum of all presence-absence maps for the 16 host species. Equivalent of projected host diversity

663

664 Table 3. The three accuracy metrics used to evaluate model performance; lowest possible  
665 threshold (LPT), minimum predicted area (MPA) and the Boyce Index (BI). Three ‘BAM’  
666 scenarios were explored; biotic (B), abiotic (A) and biotic-abiotic (BA) scenarios, and four  
667 representations of biotic interactions. The four representations are individual habitat suitability  
668 maps of all hosts (hs.ind), individual presence-absence maps of all hosts (pa.ind), ensemble  
669 map of mean habitat suitability for all hosts (hs.ens), and ensemble map representing  
670 summation of all binary presence-absence maps (pa.ens).

	<b>LPT</b>	<b>MPA</b>	<b>BI</b>
<b>A</b>	0.250	214066	0.896
<b>B (HS.IND)</b>	0.238	147349	0.769
<b>B (PA.IND)</b>	0.252	145331	0.727
<b>B (HS.ENS)</b>	<b>0.540</b>	241599	0.842
<b>B (PA.ENS)</b>	0.327	220647	0.879
<b>BA (HS.IND)</b>	0.278	112787	0.664
<b>BA (PA.IND)</b>	0.247	171757	0.873
<b>BA (HS.ENS)</b>	0.151	<b>92740</b>	0.925
<b>BA (PA.ENS)</b>	0.262	110689	<b>0.934</b>

671

672

673 Table 4: Percent Contribution (PC) and Permutation Importance (PI) of the different environmental variables used in the final models. The four  
674 representations are individual habitat suitability maps of all hosts (hs.ind), individual presence-absence maps of all hosts (pa.ind), ensemble map  
675 of average habitat suitability for all hosts (hs.ens), and ensemble map representing summation of all binary presence-absence maps (pa.ens). Biotic  
676 representations using ensembles B (hs.ens) and B (pa.ens) are not included in table as they only contain one variable, which contributes 100% to  
677 the final models. See Table 1 for a description of the abiotic (A) codes.

678

	Abiotic (A)		Biotic (hs.ind)		Biotic (pa.ind)		Biotic Abiotic (hs.ind)		Biotic Abiotic (pa.ind)		Biotic Abiotic (hs.ens)		Biotic Abiotic (pa.ens)	
	PC	PI	PC	PI	PC	PI	PC	PI	PC	PI	PC	PI	PC	PI
<b>A (bio1)</b>	27.3	19.4					0.2	0.0	0.5	11.3	20.6	18.4	11.1	17.5
<b>A (bio2)</b>	1.5	9.5					2.8	1.1	0.7	9.1	1.2	13.6	0.7	10.0
<b>A (bio3)</b>	9.9	0.4					5.3	4.6	15.8	0.0	15.2	0.0	21.8	0.3
<b>A (bio4)</b>	3.7	7.9					0.0	0.0	1.3	10.3	6.6	5.1	6.5	7.8
<b>A (bio5)</b>	1.0	10.5					0.0	0.0	0.3	5.2	3.3	1.4	2.1	7.5
<b>A (bio6)</b>	4.2	8.1					0.2	2.2	0.4	5.8	7.3	12.8	1.2	9.9
<b>A (bio7)</b>	0.0	0.0					0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.0
<b>A (bio10)</b>	0.4	0.0					0.9	0.0	2.0	0.2	0.0	0.0	0.0	0.0
<b>A (bio11)</b>	0.0	0.0					0.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0
<b>A (bio12)</b>	0.2	2.6					3.1	0.0	0.4	3.9	0.7	7.6	0.5	3.6
<b>A (bio13)</b>	0.3	0.0					0.6	6.1	0.4	2.4	0.4	0.0	0.0	0.0

<b>A (bio14)</b>	19.2	22.8					3.5	17.2	6.7	21.9	14.5	18.7	11.6	18.9
<b>A (bio15)</b>	14.8	1.3					0.4	1.8	1.3	0.4	7.6	0.0	5.1	0.0
<b>A (bio16)</b>	0.0	0.0					0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0
<b>A (bio17)</b>	17.7	17.6					1.0	16.1	4.8	18.5	8.5	14.8	9.2	17.2
<b>B (ensemble)</b>											13.8	7.5	30.2	7.4
<b>B (fox)</b>			0.2	0.0	0.0	0.0	1.9	0.8	0.0	0.0				
<b>B (sika deer)</b>			37.0	17.9	21.4	16.5	32.6	0.5	13.5	1.0				
<b>B (red deer)</b>			2.7	8.8	11.8	11.4	0.9	6.8	7.6	1.0				
<b>B (fallow deer)</b>			1.1	0.7	2.2	0.0	1.6	0.0	1.3	0.0				
<b>B (hare)</b>			7.1	15.4	4.4	5.5	1.9	2.3	1.9	0.0				
<b>B (wood mouse)</b>			0.0	0.0	5.9	5.9	0.3	0.0	4.4	0.8				
<b>B (pygmy shrew)</b>			2.0	4.2	1.2	2.3	0.1	0.7	0.2	0.0				
<b>B (grey squirrel)</b>			1.6	8.9	1.6	4.1	3.5	8.8	3.6	1.6				
<b>B (red squirrel)</b>			3.9	18.1	7.3	8.8	1.8	13.8	5.8	3.8				
<b>B (bank vole)</b>			15.0	16.2	33.3	27.6	12.8	14.5	21.2	1.2				
<b>B (pine marten)</b>			0.1	1.8	7.1	9.8	0.9	0.0	2.9	0.0				
<b>B (rat)</b>			3.0	6.0	1.3	1.8	2.5	0.0	0.1	0.0				
<b>B (rabbit)</b>			10.0	0.2	0.0	0.0	6.2	0.0	0.0	0.0				
<b>B (stoat)</b>			0.9	1.0	2.5	6.4	0.5	2.5	2.8	1.5				
<b>B (badger)</b>			0.1	0.7	0.0	0.0	0.4	0.0	0.0	0.0				
<b>B (hedgehog)</b>			15.2	0.0	0.0	0.0	13.4	0.0	0.0	0.0				



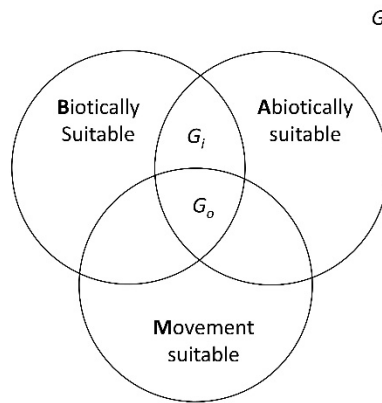
680 **Figure Captions**

681 Figure 1. The BAM diagram, which depicts the interaction between biotic (**B**), abiotic (**A**),  
682 and movement (**M**) factors. Three areas are depicted: **G** the geographical space within which  
683 analyses are developed, **G<sub>o</sub>** = the occupied distributional area, and **G<sub>i</sub>** = the invadable or  
684 potential distribution or **BA**. Modified from Soberon (2007).

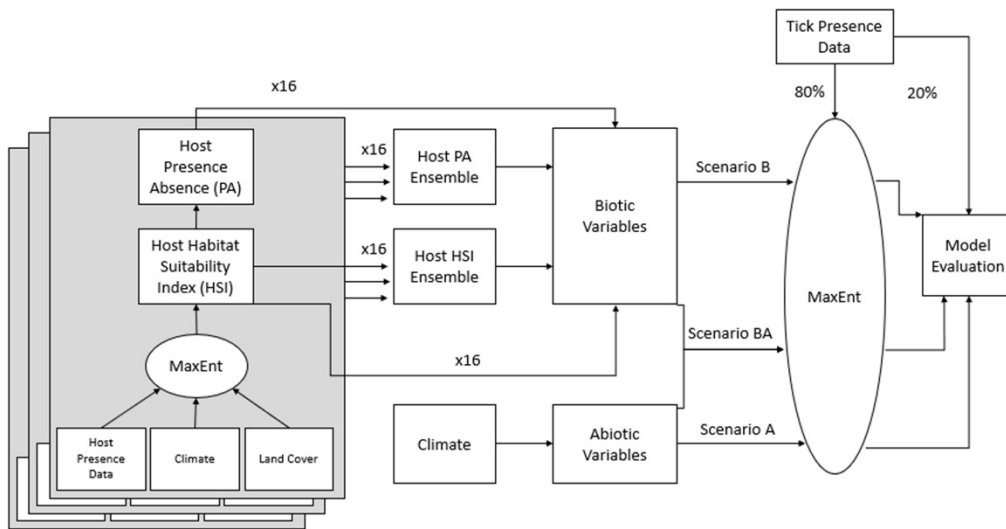
685 Figure 2. Conceptual workflow of the methodological steps

686 Figure 3. Distribution maps for *Ixodes Ricinus* in Ireland and the UK. Models parameterized  
687 on a) abiotic variables, b) biotic (hs.ind), c) biotic (pa.ind), d) biotic (hs.ens), e) biotic (pa.ens),  
688 f) biotic-abiotic (hs.ind), g) biotic-abiotic (pa.ind), h) biotic-abiotic (hs.ens), and i) biotic-  
689 abiotic (pa.ens).

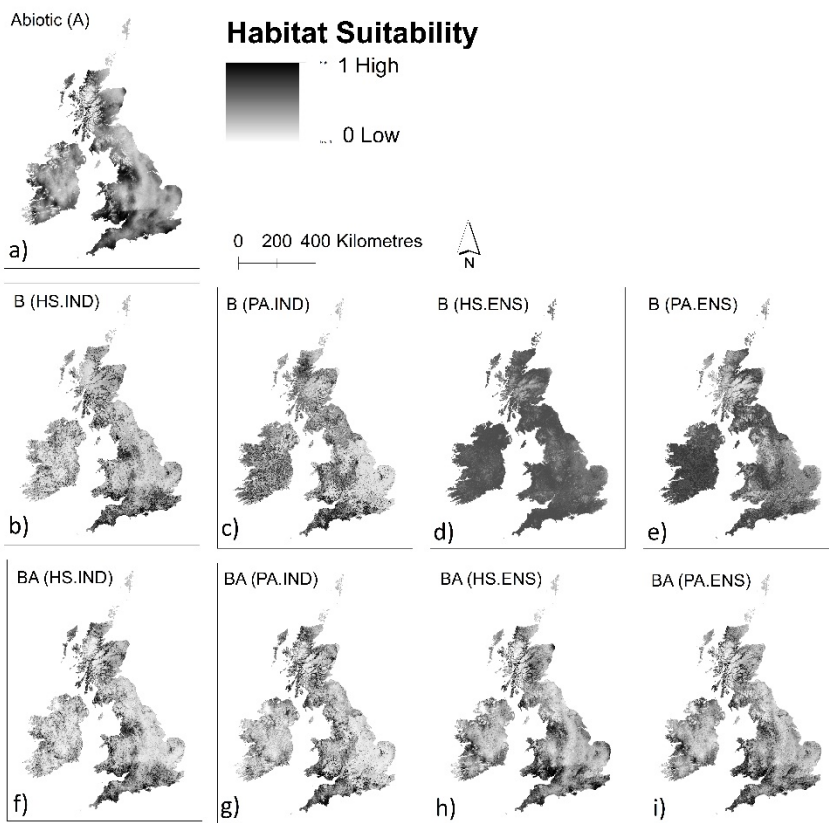
690 Figure 4. Regularised training gain from the jackknife test of variable importance. The  
691 environmental variable with the highest gain has the most useful information when models are  
692 parameterised solely on that variable. The environmental variable with the largest decrease in gain when  
693 omitted from model parameterisation has the most i



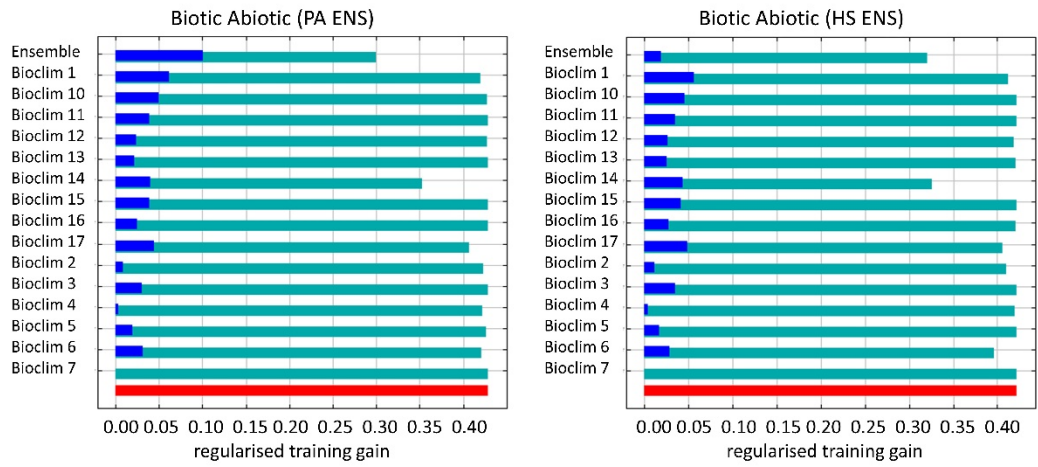
694



695



696



Without variable ■  
 With only variable ■  
 With all variables ■

697

698 information that is not present in the other variables.