

1 **Bayesian species distribution models integrate presence-only and presence-absence data**
2 **to predict deer distribution and relative abundance.**

3 Virginia Morera-Pujol¹, Philip S. Mostert², Kilian Murphy¹, Tim Burkitt³, Barry Coad⁴, Barry

4 J. McMahon⁵, Maarten Nieuwenhuis⁶, Kevin Morelle^{7,8}, Alastair Ward⁹, Simone Ciuti¹

5 *1 - Laboratory of Wildlife Ecology and Behaviour, School of Biology and Environmental Science,*
6 *University College Dublin, Dublin, Ireland*

7 *2 - Department of Mathematical Sciences, NTNU Norwegian University of Science and Technology,*
8 *Gløshaugen, Norway*

9 *3 - Deer Biologist and Management Consultant, DEER MANAGEMENT SOLUTIONS, Coolies,*
10 *Muckcross, Killarney, Co. Kerry V93R380*

11 *4 - Coillte Forest, Coillte, Dublin Road, Newtownmountkennedy, Co Wicklow, Ireland A63 DN25*

12 *5 - UCD School of Agriculture & Food Science, University College Dublin, Belfield, Dublin, Ireland*

13 *6 - UCD Forestry, School of Agriculture and Food Science, University College Dublin, Belfield,*
14 *Dublin, Ireland*

15 *7 - Department of Migration, Max Planck Institute of Animal Behaviour, Radolfzell, Germany*

16 *8 - Department of Game Management and Wildlife Biology, Czech University of Life Science, Prague,*
17 *Czech Republic*

18 *9 - School of Biology, University of Leeds, Leeds, LS2 9JT.*

19

20 **Abstract**

21 The use of georeferenced information on the presence of a species to predict its distribution
22 across a geographic area is one of the most common tools in management and conservation.
23 The collection of high-quality presence-absence data through structured surveys is, however,
24 expensive, and managers usually have more abundant low-quality presence-only data
25 collected by citizen scientists, opportunistic observations, and culling returns for game
26 species. Integrated Species Distribution Models (ISDMs) have been developed to make the
27 most of the data available by combining the higher-quality, but usually less abundant and
28 more spatially restricted presence-absence data, with the lower quality, unstructured, but
29 usually more extensive and abundant presence-only data. Joint-likelihood ISDMs can be run
30 in a Bayesian context using INLA (Integrated Nested Laplace Approximation) methods that
31 allow the addition of a spatially structured random effect to account for data spatial
32 autocorrelation. These models, however, have only been applied to simulated data so far.
33 Here, for the first time, we apply this approach to empirical data, using presence-absence and
34 presence-only data for the three main deer species in Ireland: red, fallow and sika deer. We
35 collated all deer data available for the past 15 years and fitted models predicting distribution
36 and relative abundance at a 25 km² resolution across the island. Models' predictions were
37 associated to spatial estimate of uncertainty, allowing us to assess the quality of the model
38 and the effect that data scarcity has on the certainty of predictions. Furthermore, we validated
39 the three species-specific models using independent deer hunting returns. Our work clearly
40 demonstrates the applicability of spatially-explicit ISDMs to empirical data in a Bayesian
41 context, providing a blueprint for managers to exploit unused and seemingly unusable data
42 that can, when modelled with the proper tools, serve to inform management and conservation
43 policies.

44 **1 - Introduction**

45 Methods to accurately predict species distributions have been central to wildlife
46 management, conservation of endangered species, control of invasive species, and
47 improvement of human-wildlife coexistence (Nyhus 2016, Frans et al. 2021). Species
48 distribution models (SDMs) correlate species occurrence to variables reflecting climatic and
49 environmental conditions, allowing us to understand spatiotemporal drivers of species
50 occurrence in different areas or under different climatic conditions (Guisan and Zimmermann
51 2000). SDMs have increased in complexity since their origin, aiming to improve the
52 predictions based on environmental variables, to account for spatial autocorrelation, and to
53 include different data types, such as presence-only, occurrence, or presence-absence (Guisan
54 and Thuiller 2005, Elith and Leathwick 2009, Guillera-Aroita et al. 2015).

55 SDMs have been developed to deal with systematically collected data with strict
56 control for effort, methodology and spatial coverage, although these are typically expensive
57 to collect and are thus scarce and with low spatial coverage (Hortal and Lobo 2005, Miller et
58 al. 2019). Unstructured data, where collection effort, protocol, and exact location may not be
59 specified, offer an alternative, more abundant even though less accurate source of information
60 with the potential to give relevant insights about species ecology. Unstructured data may
61 range between museum records and opportunistic citizen science observations, sometimes
62 collected using recent advances in technology such as smartphone applications (Boyce and
63 Corrigan 2017, Pacifici et al. 2017); in game species, unstructured data can be originated
64 from culling returns (Nagy-Reis et al. 2021). Although unstructured datasets may be more
65 abundant and have wider spatial and temporal coverage than structured data, their use in
66 SDMs raises issues such as the need to carefully consider observation bias and the
67 underestimation of local occurrence rates due to the lack of information on the observational
68 process (Yackulic et al. 2013, Pacifici et al. 2017).

69 Differently from structured data, unstructured and opportunistic datasets do not
70 include species absences and, to be used in a species distribution model, pseudo-absences
71 need to be randomly generated in locations where the species could have been present but
72 were not observed (Lobo et al. 2010). Although different SDM techniques have been
73 developed to work specifically with one (e.g. presence-absences) or another (e.g. presence-
74 only) type of data (Elith et al. 2006, Aarts et al. 2012, Isaac et al. 2019), both data types are
75 often available for a single species, area, and time period, introducing the possibility of
76 combining them. Two approaches have been developed to cope with this analytical
77 challenge: data pooling and model-based data integration (or Integrated Species Distribution
78 Models, ISDMs).

79 The data pooling approach combines datasets prior to entering a model, by degrading
80 the higher quality dataset until it has a common observation process with the lower quality
81 dataset (e.g. converting a presence-absence dataset to presence-only observations, Ahmad
82 Suhaimi et al. 2021). Alternatively, ISDMs avoid losing data quality in the most accurate
83 dataset by considering the two datasets as different representations of the same distribution,
84 and thus modelling them together combining the two likelihoods (joint-likelihood approach,
85 Pacifici et al. 2017). Additional advantages have become obvious in ISDMs: on the one hand,
86 including an unbiased structured dataset (i.e. a presence-absence dataset) helps compensate
87 for potential biases in presence-only datasets (Simmonds et al. 2020); on the other hand,
88 ISDMs improve the ability to predict over a wider geographic area by combining a spatially
89 restricted presence-absence dataset with an overlapping, but more extensive, presence-only
90 dataset (Simmonds et al. 2020).

91 As datasets become increasingly complex, the challenge for SDMs is to find
92 appropriate ways to account for the spatial structure of the observations and their intrinsic
93 autocorrelation. Hierarchical Bayesian models allow for the inclusion of a spatially structured

94 random effect (i.e. spatial field) that captures all the spatially explicit structures that might
95 influence the distribution of observations (Paradinas et al. 2017, Lezama-Ochoa et al. 2020).
96 In addition, Integrated Nested Laplace Approximation (INLA) methods have recently been
97 implemented within the *R-INLA* package (Rue et al. 2009, Bakka et al. 2018) and the more
98 recent development of the *inlabru* package (which provides easy access to most *R-INLA*
99 functionality for spatially structured data, Bachl et al. 2019). INLA provides a
100 computationally fast modelling environment for hierarchical Bayesian models where
101 complex spatially structured random effects can be added to models for a wide variety of
102 response variables (e.g. binomial models for presence-absence data or Poisson models for
103 presence-only and count data, Bakka et al. 2018).

104 The above mentioned methods can help to model species distributions and understand
105 their drivers, turning them into a great tool for wildlife conservation and management
106 (Linnell and Zachos 2011). Since the latter half of the 20th century, ungulate populations
107 across Europe have shown similar expansive trends and increased local densities (Apollonio
108 et al. 2010, Putman et al. 2011), placing them at the heart of human-wildlife coexistence
109 research. Human-ungulate coexistence has permeated a wide variety of land-uses, among
110 them the damage to commercial forestry plantations (Chadwick et al. 1996, Spake et al.
111 2020) and crops (Linnell et al. 2020); the transmission of diseases to livestock and eventually
112 humans (Gortázar et al. 2012); and collisions with vehicles (Langbein et al. 2011). Most
113 management plans depend on regulating the populations through hunting quotas, which
114 requires a good assessment of population densities, locally and globally (Putman et al. 2011,
115 Krausman and Bleich 2013, Richardson et al. 2020). However, despite the importance of
116 having accurate estimates of population densities and distributions to inform management,
117 survey methods are rarely coordinated or standardised, and most information comes from

118 private stakeholders' efforts to survey local populations (Liu and Nieuwenhuis 2014) or, at
119 most, population estimates based on hunting returns (Apollonio et al. 2010).

120 Ireland provides a representative study case to apply recent advances with ISDMs to
121 ungulate management, being home to expanding populations of native red deer (*Cervus*
122 *elaphus*), and non-native fallow (*Dama dama*) and sika deer (*Cervus nippon*, Carden et al.
123 2011). Despite the recent population expansion of the three species (Purser et al. 2010, Liu
124 and Nieuwenhuis 2018), Ireland lacks a national management plan for any of its deer species
125 and, currently, management is limited to hunting permits that do not limit hunters on where
126 (e.g. high-density hotspots), how many, and which deer (e.g. species, age and sex classes) to
127 hunt. This is due to the lack of an empirical basis on deer distribution and relative abundance
128 needed to set harvest quotas, maintain healthy populations and improve human-wildlife
129 coexistence (Millspaugh et al. 2009, Williams 2011, Nagy-Reis et al. 2021). Up until now
130 ISDMs within an INLA context had only been applied to simulated data (Simmonds et al.
131 2020, Ahmad Suhaimi et al. 2021); here, for the first time, we demonstrate how this approach
132 can be applied to empirical data.

133 Specifically, we collated all data available on deer distribution in Ireland previously
134 collected by several stakeholders at different spatio-temporal scales. We also collected
135 original data using *ad hoc* web tools we created and made accessible to deer stakeholders.
136 Our goal is to demonstrate how ISDMs can integrate structured and unstructured data to
137 produce and validate predicted distributions for each species of deer present in Ireland,
138 fundamental to inform science-based management practices. This study aims at
139 demonstrating the applicability of an approach that can be adapted more broadly, and
140 ultimately produce more accurate distributions of species that can be used for science-
141 informed wildlife conservation and for the management of human-wildlife conflicts.

142 **2 – Methods**

143 2.1 Studied species

144 There are three species of deer well distributed through Ireland, red deer, sika deer,
145 and fallow deer. Red deer are native to Ireland (but see Carden et al. 2012), whereas fallow
146 deer were introduced by the Anglo-Normans in the 12th century (Beglane et al. 2018) and
147 sika deer were initially introduced for ornamental purposes in 1860s in the Wicklow
148 mountains not far from Dublin (Powerscourt 1884).

149 To gather all data available on deer in Ireland and Northern Ireland (NI, UK), we
150 contacted (1) Coillte (<https://www.coillte.ie/>), which provided the results of the systematic
151 deer presence-absence surveys in part of the 440,000 ha of forests they manage in Ireland,
152 and (2) the British Deer Society (<https://bds.org.uk/>), which provided survey data on the
153 presence-absence of deer in NI. These first two datasets were the only presence-absence (PA)
154 data available for the entire island. We collated presence-only (PO) data from (1) the British
155 Agri-Food and Biosciences Institute (<https://www.afbini.gov.uk/>) which provided geotagged
156 data on culling returns from NI. We also downloaded all observations from (2) Ireland's
157 National Biodiversity Database (<https://biodiversityireland.ie/>), a citizen science platform
158 where users can submit deer observations, (3) iNaturalist (<https://www.inaturalist.org/>), an
159 international platform with the same goal; (4) and the platform CEDaR
160 ([https://www.nmni.com/CEDaR/CEDaR-Centre-for-Environmental-Data-and-](https://www.nmni.com/CEDaR/CEDaR-Centre-for-Environmental-Data-and-Recording.aspx)
161 [Recording.aspx](https://www.nmni.com/CEDaR/CEDaR-Centre-for-Environmental-Data-and-Recording.aspx)) which curates all data for NI obtained from citizen science platforms and
162 other surveys; and (5) the web survey (<https://smartdeer.ie/>) we developed *ad hoc* to collect
163 PO data from Irish deer stakeholders

164 We obtained a total of 29,140 PA observations and 4,185 PO observations, spanning
165 between 2007 and 2022 (the vast majority being collected in the last decade, see Table 1 for

166 full details on the temporal resolution of data). From these, we generated three separate
167 datasets, one for each species (red, sika, and fallow deer), to run one model for each. In
168 addition to the PO and PA data introduced above, we gathered hunting culling returns from
169 the National Park & Wildlife Service (NPWS, <https://www.npws.ie/>), responsible for issuing
170 hunting licences. Culling returns are an alternative source of data (Milner et al. 2006, Forsyth
171 et al. 2022), and we retained this dataset to validate the ISDMs we built by integrating PO
172 and PA data.

173 2.2 Data collection and pre-processing

174 *2.2.1 Presence absence (PA) data*

175 PA data for each species were obtained from Coillte based on surveys performed in a
176 fraction of the 6,000 properties they manage (Table 1), by asking property managers (who
177 visit the forests they manage on a regular basis) whether deer were present and, if so, what
178 species. Properties range in size from less than one to around 2,900 ha, and to assign the PA
179 value to a specific location, we calculated the centroid of each property using the function
180 *st_centroid()* from the package *sf* in R (Pebesma 2018). The survey was mainly performed in
181 2010 and 2013, in addition to further data collected between 2014 and 2016. Some properties
182 were surveyed only once in the period 2010-2016, but for those that were surveyed more than
183 once, the value for that location was considered “absence” if deer had never been detected in
184 the property in any of the surveys, and “presence” in all other cases. In addition to these
185 surveys, Coillte commissioned density surveys based on faecal pellet sampling in a subset of
186 their properties between the years 2007 and 2020. Any non-zero densities in these data were
187 considered “presences”, and all zeros were considered “absences”. These data were also
188 summarised across years when a property had been repeatedly sampled, and counted as
189 presence if deer had been detected in any of the samples (Table 1).

190 PA data for NI were obtained from a survey carried out by the British Deer Society in
191 2016. The survey divided the British territory in 100 km² grid cells and deer presence was
192 assigned based on public contributions, which were then reviewed and collated by experts.
193 Since 100 km² grid cells are quite large, we did not, as with the Coillte properties, calculate
194 the centroid of each cell and assign the PA value of the cell to it. Instead, we randomly
195 simulated positions within each cell and assigned the presence or absence value of the cell to
196 each of them. We performed a sensitivity analysis to calculate an optimal number of positions
197 that would capture the environmental variability within each cell (Suppl material S1), which
198 was set to 5 random positions per grid cell. After processing, we obtained a total of 920 PA
199 data across NI (Fig 1a).

200 *2.2.2 Presence only (PO) data*

201 PO data were collected from various sources, mainly (but not only) from citizen
202 science initiatives. The National Biodiversity Data Centre (NBDC) is an Irish initiative that
203 collates biodiversity data coming from different sources, from published studies to citizen
204 contributions. From them, we obtained all contributions on the three species, a total of 1,430
205 records. To this, we added the 164 records of deer in Ireland downloaded from the iNaturalist
206 site, another citizen contributed database that collects the same type of data. From the
207 resulting dataset, we (1) removed all observations with a spatial resolution lower than 1 km²;
208 (2) did a visual inspection of the data and comments, and removed all observations that were
209 obviously incorrect (i.e. at sea or that the comment specified it was a different species); (3)
210 filtered out all the fallow deer reported in Dublin's enclosed city park (Phoenix Park) to avoid
211 biases caused by the large amount of people reporting deer from the capital; and (4) filtered
212 duplicate observations by retaining only one observation per user, location, and day. The
213 Centre for Environmental Data and Recording (CEDaR) plays a similar role to the NBDC in
214 Northern Ireland. They provided 872 records of deer in NI, coming from different survey,

215 scientific, and citizen science initiatives, from which we removed all records provided with a
 216 spatial resolution lower than 1 km². The location and species of 469 deer culled between
 217 2019 and 2021 in NI were obtained from the British Agri-Food and Biosciences Institute. For
 218 the observations that did not have specific coordinates, we derived them from the location
 219 name or postcode if provided.

220 As part of a nationally funded initiative to improve deer monitoring in Ireland
 221 (SMARTDEER), we developed a bespoke online tool to facilitate the reporting of deer
 222 observations by the general public and all relevant stakeholders e.g. hunters, farmers, or
 223 foresters. Observations were reported in 2021 and 2022 by clicking on a map to indicate a 1
 224 km² area where deer have been observed. For each user and session, we calculated the area of
 225 the surface covered in squares, and simulated a number of positions proportional to the size
 226 of the polygon and distributed them within it to generate a number of exact positions
 227 equivalent to the area were the user had indicated an observation (details in Supp material
 228 S1). In total, the SMARTDEER tool allowed us to collect 4,078 presences across Ireland and
 229 NI (Table 1, Fig. 1b).

Table 1. Summary of the presence-absence (PA, structured data) and presence-only (PO, unstructured data) datasets gathered for Ireland and Northern-Ireland (NI, UK).					
Data type	Source	Years	Red deer	Sika deer	Fallow deer
PA	BDS survey	2016	920	920	920
	Coillte density surveys	2007 - 2020	417	417	417
	Coillte desk surveys	2010 - 2016	4 936	4 936	4 936
	TOTAL		6 273	6 273	6 273
PO	Citizen Science	2005 - 2021	408	573	394
	AFBI culling returns	2017 - 2021	7	169	259
	Smartdeer web survey	2021 - 2022	507	460	528
	Others	2001 - 2018	51	35	69
	TOTAL		973	1 237	1 250

230

231

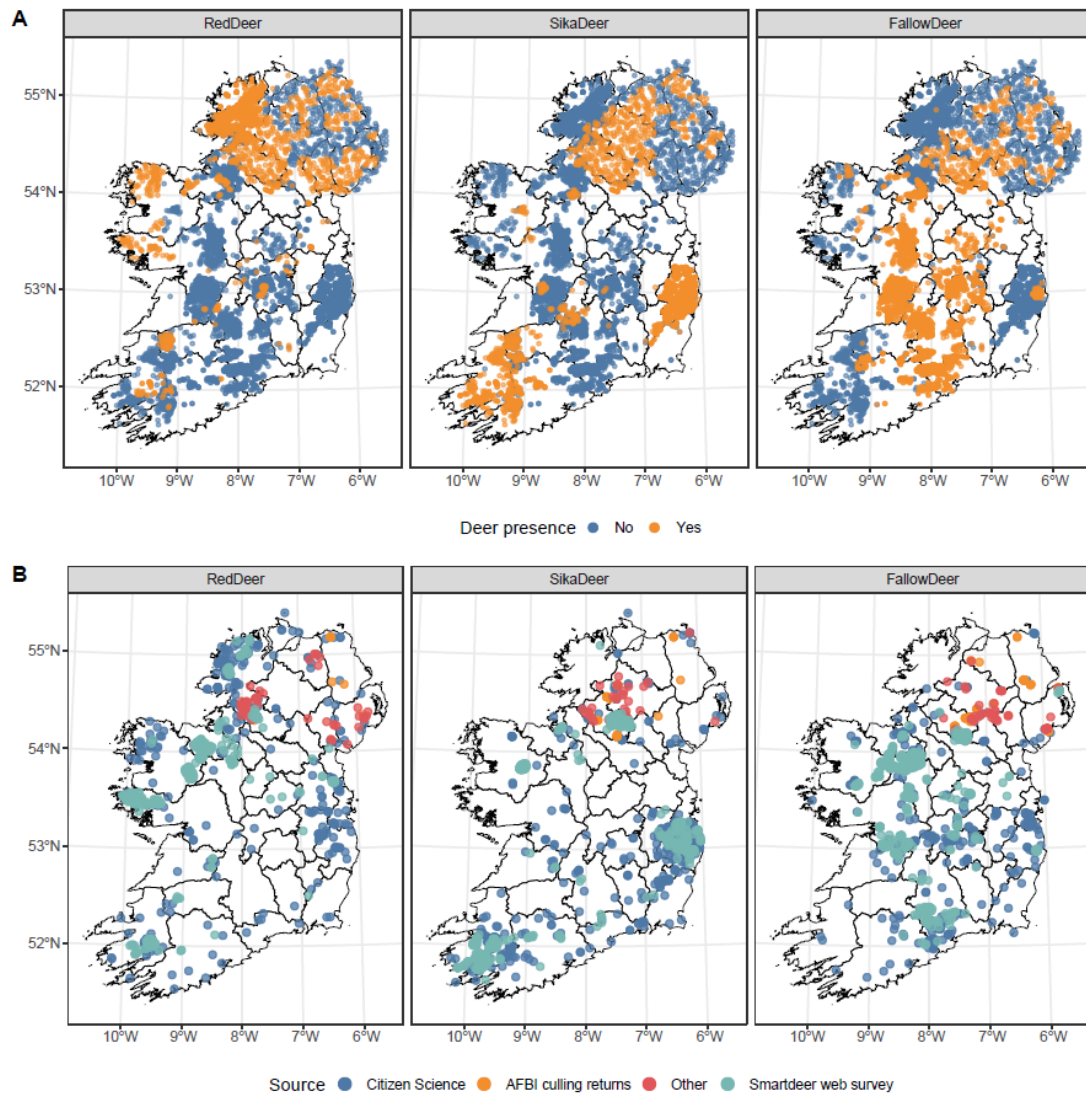


Figure 1: presence-absence (A) and presence-only (B) data for each deer species (see Table 1 for sample sizes and temporal resolution). Presence-absence data were provided by Coillte in Ireland and by the British Deer Society in Northern Ireland (NI, UK), while presence-only data were collated from a wide variety of sources including citizen science data, location of culled animals, and our own web tools specifically designed for deer stakeholders (<https://smartdeer.ie/>).

232 The data used in the models were collected between 2007 and 2022. Deer populations
233 expanded in Ireland until 2008 (Carden et al. 2011), and according to culling return data have
234 somewhat stabilised since then (NPWS official data). Although the range expansion of deer
235 species would merit further investigation, here we provide for the first time an accurate
236 modelled distribution of the three main species of deer in Ireland, and since the data are
237 scarce, we have made use of all available data without considering the temporal trends. A

238 continued data collection scheme will provide enough data to study population size and range
239 changes, but this is beyond the scope of this manuscript.

240 2.3 Statistical model

241 To integrate the two datasets into one model for each species, we used functions from
242 the *PointedSDMs* package (<https://github.com/PhilipMostert/PointedSDMs>). This package is
243 designed to construct ISDMs from different data sources by employing the functions from the
244 *inlabru* package (Bachl et al. 2019), within the R-INLA framework (Rue et al. 2009).
245 Through the functions provided in the package, we constructed a joint likelihood model, with
246 the PA data modelled as a Bernouilli distribution (Isaac et al. 2019), modelling the
247 probability of observing an individual at each location s

$$248 \quad Y_s \sim \text{Bernouilli}(p_s)$$

$$249 \quad \text{cloglog}(p_s) = \alpha_1 + \beta_1 x_1 + \dots + \beta_n x_n + \xi_s$$

250 where Y_s is the binary response variable (PA) and p_s is the probability of presence. This is
251 linked to the linear predictor by a complementary log-log link function (cloglog, Ahmad
252 Suhaimi et al. 2021). The linear predictor is composed of a dataset-specific intercept (α_1), a
253 set of covariates (x_1 to x_n) and their coefficients (β_1 to β_n), and a dataset-specific random
254 spatial effect (spatial field) to account for the spatial structure of the data (ξ_s).

255 In turn, the PO data are modelled as a log-Gaussian Cox process with intensity
256 function

$$257 \quad N(A) \sim \text{Poisson} \int_A \lambda(s) d(s)$$

$$258 \quad \log(\lambda(s)) = \alpha_2 + \beta_1 x_1 + \dots + \beta_n x_n + \omega_s$$

259 where N is the expected number of presences in the study area (A), λ is the intensity function,
260 α_2 is the dataset-specific intercept, the same vector of covariates with their effect sizes as in

261 the PA model, and another data-specific spatial field (ω_s). The use of the cloglog link in the
262 binomial model allows its response to be interpreted on the same scale as the response of the
263 Poisson model, which allows the sharing of parameters between likelihoods (Bowler et al.
264 2019). Spatial fields from both processes (PO, PA) are modelled as Gaussian random fields
265 with Matérn covariance functions, which are approximated using a triangulation of the study
266 area (called a mesh) through stochastic partial differential equations fitted in a Bayesian
267 context through integrated nested Laplace approximations (Lindgren et al. 2011).

268 *2.3.1 Prior specification*

269 The spatial fields are controlled by two hyperparameters –range and marginal variance.
270 The range controls the smoothness of the spatial field (i.e. the distance between peaks and
271 throughs), and the variance controls the magnitude of these peaks and throughs. In the
272 Bayesian context in which we are fitting this model, we need to set prior values to these two
273 hyperparameters. To do so, we use Penalised Complexity (PC) priors, a newly developed
274 framework that allows easily interpretable and controllable priors (Simpson et al. 2014). PC
275 priors are weakly informative (allowing the posterior of each hyperparameter to be mainly
276 controlled by the data) and penalise model complexity by “pulling” the model towards its
277 simplest realisation (the “base” model), which has infinite range and zero variance (i.e. a
278 completely flat spatial field, absence of spatial structure). To set the priors, we inform the
279 model of “how far it is allowed to deviate” from those base models using the following
280 specifications:

- 281 - The prior on the range (ρ) is set providing the lower tail quantile ρ_0 and the
282 probability $P(\rho)$ so that

$$283 \quad \text{Prob}(\rho < \rho_0) = P(\rho)$$

284 or “the probability that the true range (ρ) is smaller than ρ_0 is $P(\rho)$ ”. For example, if we set
285 ρ_0 to be 50 and $P(\rho)$ to be 0.05, we are telling the model that the probability of the true range
286 of the spatial field being smaller than 50 km is 5%. In this way we are limiting the range to
287 values between infinite (the base model) and 50, i.e. we are saying that the smallest that the
288 range could possibly be is 50 km with a probability of 95%.

289 - The prior on the variance is set on the standard deviation, providing the upper tail
290 quantile σ_0 and the probability $P(\sigma)$ so that

$$291 \quad \text{Prob}(\sigma > \sigma_0) = P(\sigma)$$

292 or “the probability that the true standard deviation σ is larger than σ_0 is $P(\sigma)$ ”. For example,
293 if we set σ_0 to be 0.5 and $P(\sigma)$ to be 0.05, the probability of the true standard deviation being
294 larger than 0.5 is 5%, so effectively the standard deviation value is limited between 0 (the
295 base model) and 0.5 with a 95% probability.

296 Priors have to be carefully specified, but there is no absolute rule for it, so the
297 decisions that go into the prior choice are an essential part of the modelling process. In this
298 case, we started off with a prior for the range (for both the PO and PA spatial fields, and for
299 the three species) that was equal to the size of the triangles of the mesh (40 km). In this way,
300 we are providing the minimum amount of information to the model, as we are setting the
301 lower limit for the range as the limit of the resolution of the model. For the standard
302 deviation, we started with a prior of 1 for all spatial fields, a value large enough to serve as an
303 appropriate upper limit.

304 *2.3.2 Covariate selection*

305 Raster environmental covariates used in the models were obtained from the
306 Copernicus Land Monitoring Service (© European Union, Copernicus Land Monitoring
307 Service 2018, European Environment Agency (EEA)), whereas the vector layers (roads,

308 paths) were obtained from the Open Street Map service (OpenStreetMap contributors, 2017.
309 Planet dump [Data file from January 2022]. <https://planet.openstreetmap.org>). Vector layers
310 were transformed into distance layers (distance to roads, distance to paths) using the
311 *distance()* function from the package *raster*, and into density layers (density of roads, paths)
312 using the *rasterize()* function of the same package (Hijmans 2021). All raster layers were
313 resampled to the lowest resolution available in the used covariates, resulting in a 1 km²
314 resolution. A full description of the process of covariate selection (including screening for
315 collinearity) can be found in the supplementary material (Supp mat. S1). The covariates
316 eventually used in the model were elevation (m), slope (degrees), tree cover (%), small
317 woody feature density (%), distances to forest edge (m, positive distances indicate a location
318 outside a forest, negative distances indicate a location within a forest), and human footprint
319 index (Venter et al. 2016, 2018). All covariates were scaled by subtracting the mean and
320 dividing by the standard deviation before entering the model (function *scale()* from the *raster*
321 package).

322 2.3.3 Spatial predictions

323 From the fitted models, we used the *predict()* function from the *inlabru* package to
324 obtain predicted deer densities in a 25 km² grid. Since models were fitted in a Bayesian
325 context, the prediction obtained at each location is not a point value but a distribution, from
326 which we can produce the mean and the standard deviation, thus obtaining a spatial estimate
327 of the uncertainty of the prediction. We used the same function to obtain the prediction of the
328 spatial effects, which can provide an indication of the spatial autocorrelation structure of each
329 of the datasets. The model is designed on the assumption that not all individuals have been
330 observed and although in theory the total abundance can be predicted integrating the intensity
331 of the process over all the study area, an imperfect detection will affect the predicted total
332 abundance. In all our models the total predicted abundances were grossly underestimated, so

333 we decided to use the predictions in the linear scale and, rescaled from 0 to 1, use them as
334 relative abundances instead of total abundances or densities.

335 2.4 Model validation

336 To validate the results of our models, we obtained culling return data from the NPWS,
337 aggregated by county between 2008 and 2018. The data consist of the number of harvested
338 deer of each species by county (ranging from 826 to 7,500 km²) and year, and the number of
339 hunting licences issued. To consider the increase in hunting pressure affecting the number of
340 deer harvested, we corrected each year and species data by the number of licences issued, and
341 then aggregated the data of the past 10 years by calculating the mean. Thus, we obtained
342 average deer harvested (corrected by number of licences) for each county. From our ISDMs,
343 we obtained the predictions this time in the response scale, to obtain aggregated abundances
344 by county, and then used a linear model to investigate how well our models predicted county-
345 level culling returns, using the R² score to evaluate the performance of the models.

346 **3 – Results**

347 We developed one model for each species, including effects for six covariates (tree
348 cover, density of small woody features, distance to the forest edge, slope, elevation, and
349 human footprint index), and two spatial fields, one for the PO data and one for the PA data
350 (Fig. S6). For red and sika deer the priors specified above for the spatial fields provided good
351 enough posterior estimates, and we did not modify them to allow the data distribution to
352 inform the model output. For fallow deer, a standard deviation of one proved insufficient to
353 capture the variability of the PO spatial field, so we ran the model again with a prior value of
354 two (Table 2).

355

Table 2. Priors' specification and posterior distribution estimated for the spatial fields in the three species-specific models (red, sika, fallow deer) for both presence-absence (PA) and present-only (PO) data. Spatial fields are defined in our model by their range and standard deviation (St. dev). Priors are set on these parameters as point values, and posterior distributions are obtained which we have summarised here as "mean (standard deviation)".

		PA data		PO data	
Species		Range	St. dev	Range	St. dev.
Red deer	Prior	40	1	40	1
	Posterior distr.	269.1 (80.0)	0.6 (0.1)	206.5 (21.5)	1.3 (0.1)
Sika deer	Prior	40	1	40	1
	Posterior distr.	266.5 (70.6)	0.6 (0.1)	207.4 (33.4)	1.5 (0.1)
Fallow deer	Prior	40	1	40	2
	Posterior distr.	193.8 (47.7)	0.7 (0.1)	171.5 (26.8)	1.46 (0.1)

356 The posterior range and SD of the spatial fields showed larger ranges and smaller SD
357 for the PA data than for the PO data, reflecting the differences in spatial structure of each
358 dataset. PA data points are more evenly distributed throughout, while PO data points display
359 more clustering.

360 The covariate effects for the three models (Fig. 2) showed that the three species had,
361 in general, similar ecology in terms of environmental preferences, i.e. sika, red, and fallow
362 deer were more likely to be observed within forests (negative values of distance to forest
363 edge) with high tree cover densities. Elevation had a small but significantly negative effect on
364 the distribution of the three species, and while slope did not have a clear effect in red and
365 fallow deer distribution (CIs overlap zero), sika deer seemed to prefer areas with steeper
366 slopes. The three species distributions seemed to match areas with greater human footprint, in
367 line with the expectation that bare and unpopulated lands are less attractive to deer.

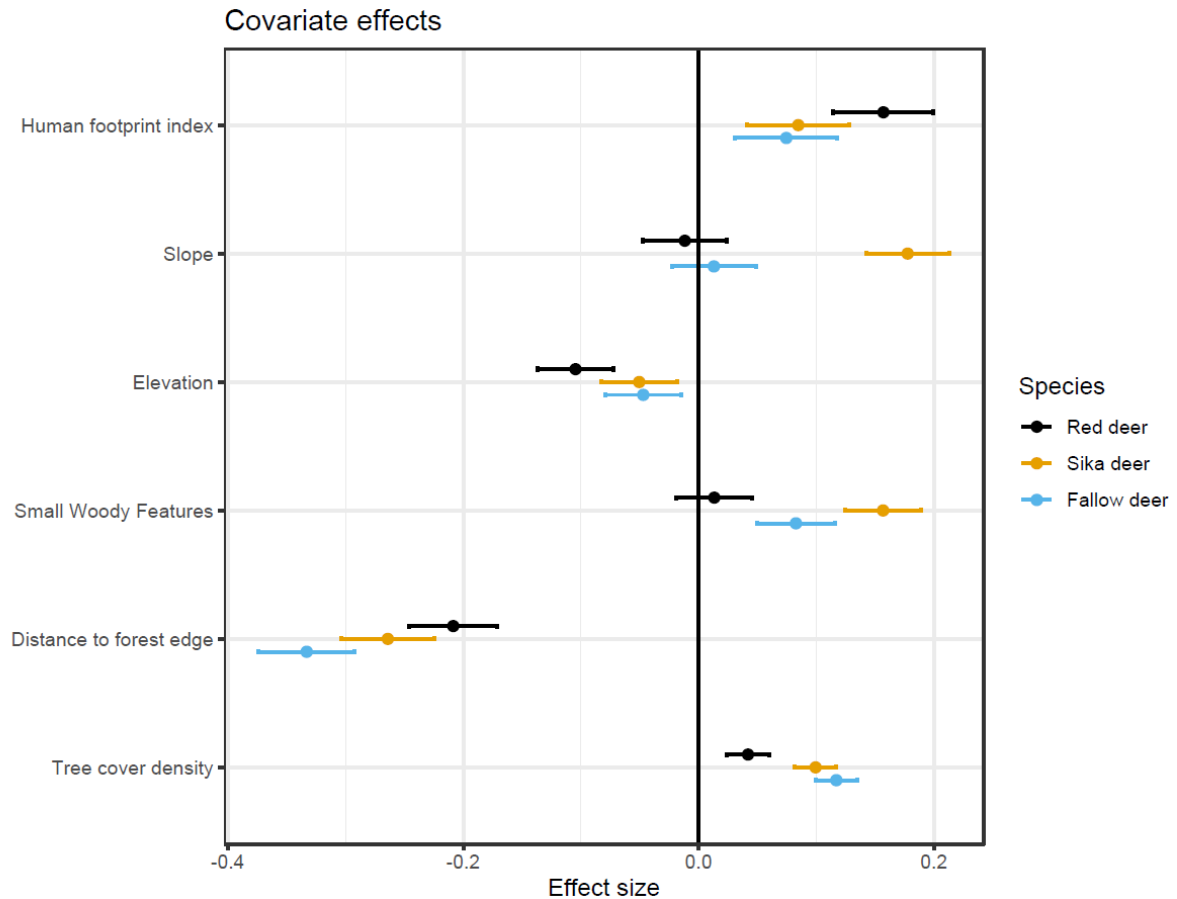


Figure 2. Covariate effects for each of the models for red (top, black), sika (middle, orange) and fallow (bottom, blue) deer. Circles represent the median value of the effect, while the bars represent the 95% credible intervals (CIs).

368 From each of the models we obtained a spatial prediction that allowed us to plot a
369 mean prediction and its standard deviation (Fig. 3). Red deer hotspots were detected in the
370 NW and SW of Ireland. Sika deer were present at higher relative abundances in a hotspot at
371 the east coast, and more diffusely in the SW, overlapping with a red deer hotspot. Lastly,
372 fallow deer are mainly distributed in the midlands. For all species, the standard deviation was
373 larger in NI, reflecting the scarcity of PO data in that region.

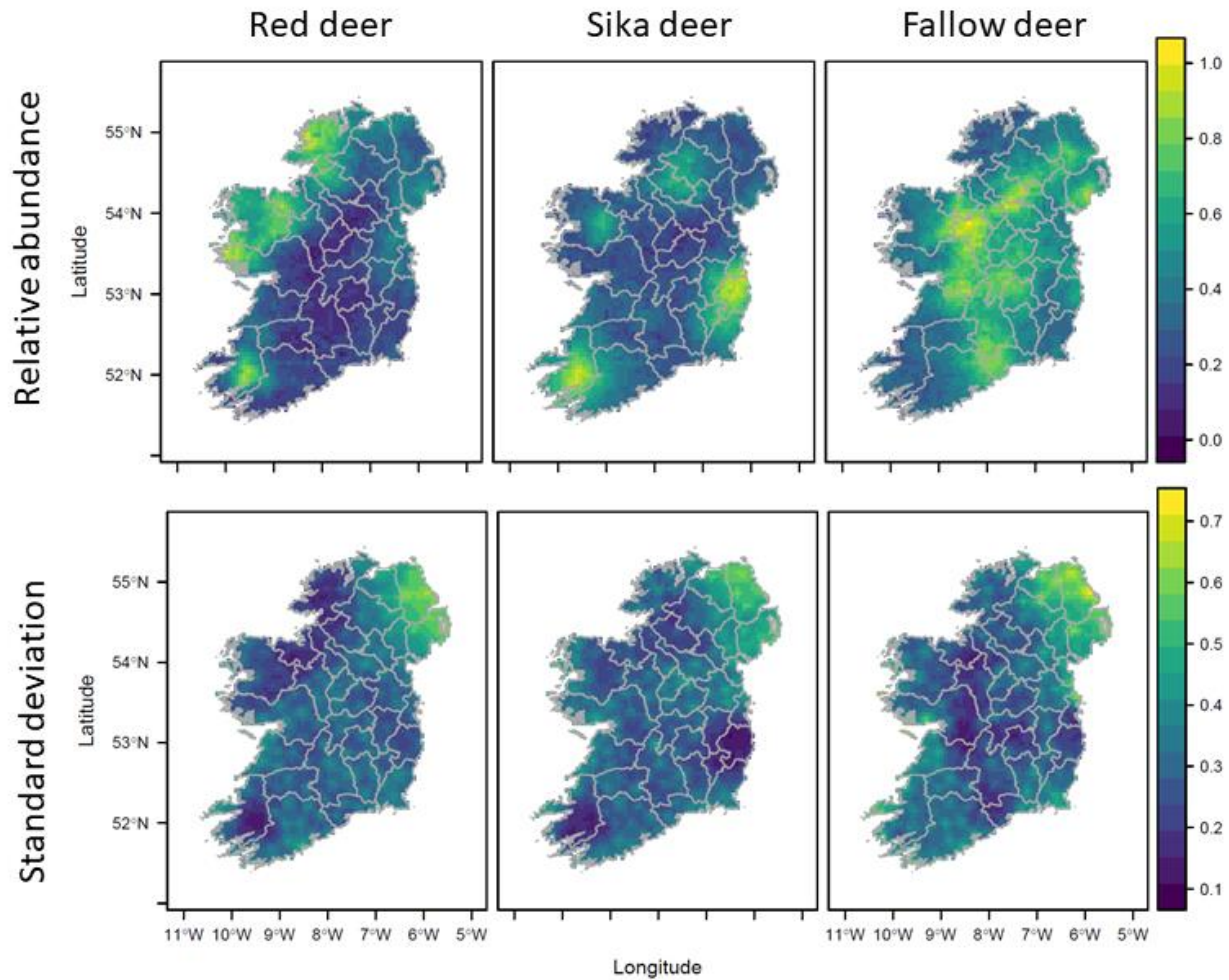


Figure 3. Mean (top) and standard deviation (bottom) of the spatial predictions for red, sika and fallow deer. The values indicate relative abundances, with 0 reflecting absence of the species and values closer to 1 representing the areas where the species is more abundant.

374 For the three models, ranges were larger and marginal variances smaller for the spatial
375 fields of the PA datasets (Table 2, Fig. S1) than for those of the PO datasets, reflecting the
376 more regular structure and thus lesser spatial autocorrelation of the dataset.

377 Our ISDMs predicted distribution and relative abundance across Ireland, and, when
378 aggregated by county, these predictions were in high agreement with the independent dataset
379 of culling returns corrected by hunters' licences. The validation analysis showed that our
380 models were particularly good in predicting distribution and relative abundance for sika deer
381 ($R^2 = 0.69$), followed by red ($R^2 = 0.52$) and fallow deer ($R^2 = 0.44$).

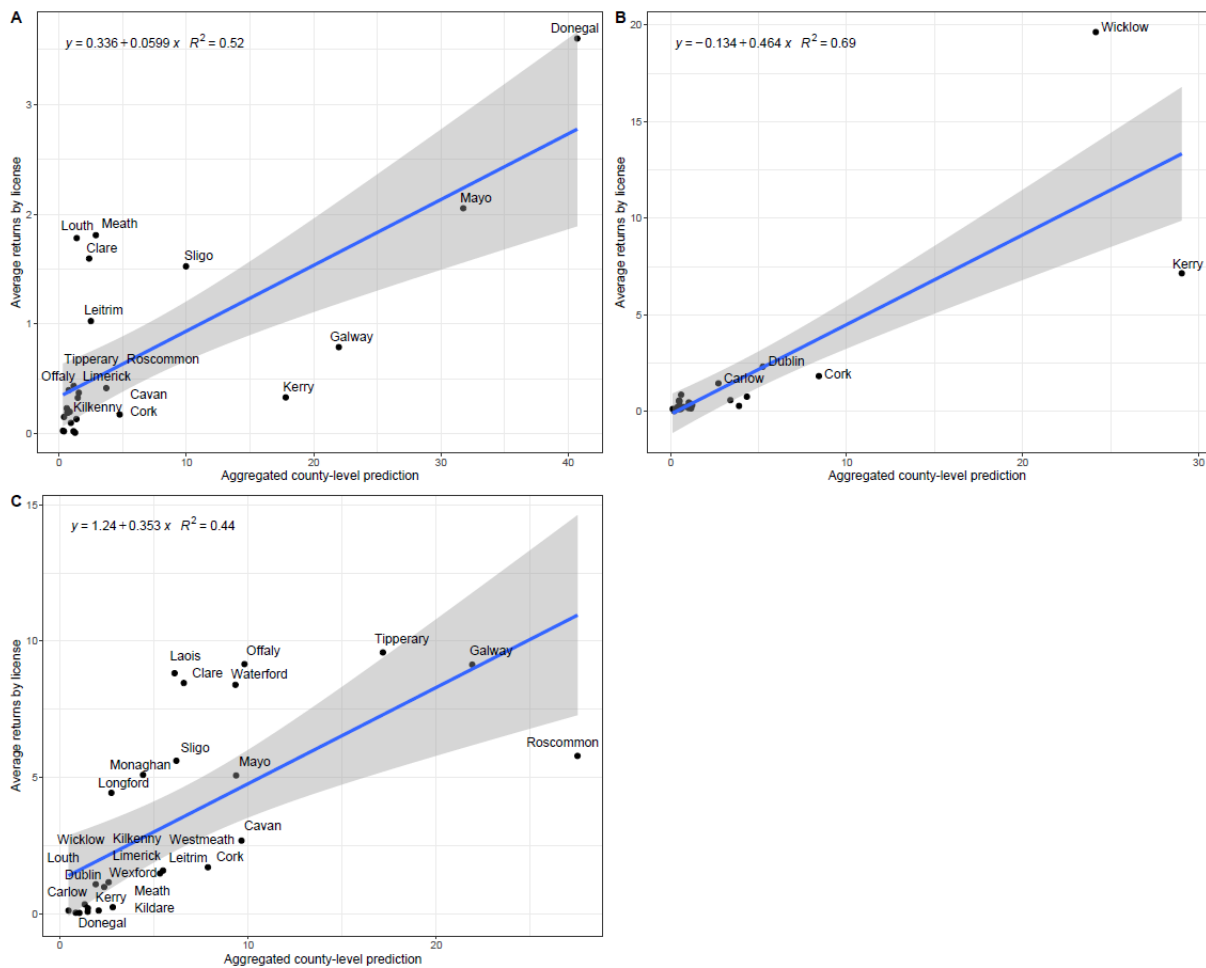


Figure 4 Validation plots for the ISDMs predicting red (A), sika (B) and fallow deer (C) distribution and relative abundance. Predictions of the ISDMs (x-axis, aggregated county-level abundance) are regressed against average culling returns (corrected by hunter licences at the county level, y-axis)

382

383 4 – Discussion

384 4.1 Applicability of joint likelihood models in an INLA context to real data

385 Our results demonstrated the practical applications of ISDM in the INLA Bayesian
386 context for the first time with real data, a method that so far had only been applied to
387 simulated datasets (Simmonds et al. 2020, Ahmad Suhaimi et al. 2021). Despite the scarcity
388 and low quality of the data, our models managed to successfully produce not only a
389 prediction of the distribution for each species, but also to map the uncertainty. The predicted
390 distributions displayed small standard deviations across most of the island, efficiently

391 reflecting the regions where data are less abundant, demonstrating how fewer data relate to
392 less certain predictions. Furthermore, we validated the predictions with an external dataset to
393 ensure their accuracy, finding that our models performed well in predicting county-level
394 culling returns. Thus, we provide accurate science-based relative abundance maps that
395 integrate all previous knowledge about deer distribution in Ireland, setting a path for future
396 data gathering initiatives with conservation and management in sight.

397 The separate spatial random fields for each dataset allowed us to capture the different
398 observational processes. Although usually PA data come from organised surveys designed to
399 avoid exhibiting any spatial structure, the PA data in our model might have exhibited some
400 spatial structure, which would have been absorbed by the PA-specific spatial field. In the
401 same way, PO data came from many different sources, including citizen science initiatives
402 that would have a clear observational bias towards more populated areas or those used for
403 recreation, but also other opportunistic observations that would have a less clearly defined
404 observational bias. Thus, the use of a PO specific spatial field was more suited for capturing
405 the spatial structure in that dataset than the addition of a covariate that could represent the
406 bias, such as the human footprint index or the distance to roads (Dorazio 2014).

407 4.2 Deer distributions and relative abundances in Ireland

408 Our model predicted two main hotspots for red deer. The hotspot in the SW was
409 centred around the Killarney National Park, a herd under conservation measures such as a
410 hunting ban in the area (Carden et al. 2012). This ban is reflected in our validation plots,
411 where our red deer model seemed to predict a larger abundance than what is reflected in the
412 culling returns, since the culling returns of red deer in that county would be disproportionate
413 small compared to the reality as much of their range is protected from hunting. The other
414 hotspots to the NW coincided with areas where modern introductions of red deer have taken

415 place in the past two centuries (Purser et al. 2010), and the diffuse populations along the
416 eastern coast correspond to the area where the first recorded introduction of red deer into
417 Ireland took place in 1246 (McDevitt et al. 2009).

418 The sika deer model showed two very clear hotspots in the E and SW of the island,
419 and two less dense populations in the NW, reflecting the history of their introduction in
420 Ireland (Purser et al. 2010). There was considerable overlap between the populations of red
421 and sika deer, which could merit further study on their habitat and diet preferences to
422 investigate the possible niche, spatial, or temporal segregation that might facilitate
423 coexistence. From our covariate effects, sika seemed to differ in habitat preferences with red
424 deer (non-overlapping CIs) in tree cover density and small woody feature density, where sika
425 deer seemed to prefer denser cover than red deer, and particularly in slope, where sika
426 seemed to prefer steeper slopes than the other two species. This difference might be reflecting
427 some habitat or space use partitioning due to competition, but it also might be related to the
428 fact that sika deer seem to prefer more acidic soils, which would allow them to exploit young
429 conifer plantations (Alfredsson et al. 1998). In addition, the distribution overlap of the two
430 species causes concerns with regards to the hybridisation between the two, which has been
431 observed both in captivity and in the wild (Abernethy 1994) and which could be a threat to
432 the genetic purity of the Kerry herd (Smith et al. 2014).

433 Fallow deer were predicted to be the most widespread species, distributed mostly over
434 the areas from where the other two species were largely absent. This might be due to different
435 habitat and food preferences, since fallow deer are known to be more obligated grazers than
436 either red or sika deer (Obidziński et al. 2013), or due to competitive exclusion, but it could
437 also be a reflection of the founder effect since fallow deer seem to have slow range expansion
438 rates from where their populations are first established (Ward 2005). Nevertheless, since the
439 last published distribution in 2008 (Carden et al. 2011), fallow deer distribution seems to

440 have expanded northward, now displaying a continuous distribution from the SE coast
441 through the midlands and the west and all the way up to the NW coast.

442 4.3 Joint likelihood models as a tool for management in data-scarce scenarios

443 Our predicted distributions described an island where deer of at least one species were
444 omnipresent, with some regions where two species spatially overlap. The covariates showed
445 that although the three species preferred areas with dense tree cover and within forests or
446 small woody features, that did not necessarily mean that deer shy away from human presence,
447 reflected in our models by a positive effect of human footprint index. That is, however, more
448 reflective of Ireland's natural habitats than of deer preferences: Ireland and NI have a large
449 proportion of heavily modified habitat (approximately 69% of Ireland and 76% of NI are
450 covered by farmland, (2021b, 2021a), with most of their agricultural land devoted to
451 permanent and rough grazing grasslands, very attractive to deer (Drennan et al. 2005, O'Mara
452 2012), The forests, small and patchily distributed, are mostly non-native and are present
453 within mosaics dominated by human modified habitats, making it almost impossible for deer
454 to avoid anthropomorphised environments. This has obvious consequences for human-
455 wildlife coexistence, since deer have more opportunities to interact heavily with human
456 resources such as roads, commercial forestry and farms. Thus, these results constitute a
457 starting point for management, by providing information on areas where the relative densities
458 of the relevant deer species are higher, and where targeted actions would be most effective.

459 With this research, we have demonstrated the use of joint Bayesian spatial models
460 fitted through INLA methods to obtain accurate distributions and relative abundances of
461 species. Our models have been validated with independent data, proving their accuracy even
462 with low quality, patchy data, which makes them a useful tool for the management and
463 conservation of wildlife in most contexts where a data collection protocol has not been

464 established. Our work now opens new exciting future scenarios, because the same type of
465 model can be adapted to estimate actual abundances by including data on the number of
466 individuals (e.g. group sizes) and sampling effort, leading ISDMs to produce even more
467 accurate information on species abundances which are so essential for science-informed
468 management.

469 **5 – Acknowledgements**

470 We are thankful to all the agencies that provided data (complete list in methods
471 section), and to the people that facilitated the datasets and were available to answer enquiries.
472 Additional thanks to Sarah Keenan and Charles Harper for their contribution to the
473 SMARTDEER project, and to Andrew McCullagh and Tony Quinn (DAFM) for their
474 continued support throughout the development of the project.

475 **6 – Funding statement**

476 This work is a part of a nationally-coordinated project entitled "A smart and open-
477 science approach to monitor and analyse deer populations in the Republic of Ireland and set
478 the scene for evidence-based deer management" funded by DAFM, grant 2019R417 [2019
479 Research Call Instruments (II - V)]. VM-P was additionally funded by the Department of
480 Further and Higher Education, Research, Innovation and Science of the government of
481 Ireland, through the Higher Education Authority's call to support research disrupted by
482 COVID-19.

483 **7 – Author contributions:**

484 VM-P: conceptualization, data curation, methodology, formal analysis, validation,
485 visualization, writing - original draft, writing - review & editing; PSM: software, writing -
486 review & editing; KM: conceptualization, data curation, writing - review & editing; TB: data
487 curation, writing - review & editing; BC: data curation, writing - review & editing; BJM:

488 conceptualization, project administration, writing - review & editing; MN: conceptualization,
489 project administration, writing - review & editing; KM: software, writing - review & editing;
490 AW: data curation, writing - review & editing; SC: conceptualization, data curation,
491 supervision, funding acquisition, project administration, writing - review and editing.

492 **References**

- 493 Aarts, G., J. Fieberg, and J. Matthiopoulos. 2012. Comparative interpretation of count,
494 presence-absence and point methods for species distribution models. *Methods in*
495 *Ecology and Evolution* **3**:177-187.
- 496 Abernethy, K. 1994. The establishment of a hybrid zone between red and sika deer (genus
497 *Cervus*). *Molecular Ecology* **3**:551-562.
- 498 Ahmad Suhaimi, S. S., G. S. Blair, S. G. Jarvis, and B. Leroy. 2021. Integrated species
499 distribution models: A comparison of approaches under different data quality
500 scenarios. *Diversity and Distributions* **27**:1066-1075.
- 501 Alfredsson, H., L. M. Condrón, M. Clarholm, and M. R. Davis. 1998. Changes in soil acidity
502 and organic matter following the establishment of conifers on former grassland in
503 New Zealand. *Forest Ecology and Management* **112**:245-252.
- 504 Apollonio, M., R. Andersen, and R. J. Putman. 2010. European ungulates and their
505 management in the 21st century. Cambridge University Press.
- 506 Bachl, F. E., F. Lindgren, D. L. Borchers, and J. B. Illian. 2019. inlabru: an R package for
507 Bayesian spatial modelling from ecological survey data. *Methods in Ecology and*
508 *Evolution* **10**:760-766.

- 509 Bakka, H., H. Rue, G. A. Fuglstad, A. Riebler, D. Bolin, J. B. Illian, E. Krainski, D. P.
510 Simpson, and F. K. Lindgren. 2018. Spatial modeling with R-INLA: A review. Wiley
511 Interdisciplinary Reviews: Computational Statistics:1-24.
- 512 Beglane, F., K. Baker, R. F. Carden, A. R. Hoelzel, A. L. Lamb, R. M. Fhionnhaile, H.
513 Miller, and N. Sykes. 2018. Ireland's fallow deer: their historical, archaeological and
514 biomolecular records. *Proceedings of the Royal Irish Academy: Archaeology,
515 Culture, History, Literature* **118C**:141-165.
- 516 Bowler, D. E., E. B. Nilsen, R. Bischof, R. B. O'Hara, T. T. Yu, T. Oo, M. Aung, and J. D. C.
517 Linnell. 2019. Integrating data from different survey types for population monitoring
518 of an endangered species: the case of the Eld's deer. *Scientific Reports* **9**:7766.
- 519 Boyce, M. S., and R. Corrigan. 2017. Moose Survey App for Population Monitoring. *Wildlife
520 Society Bulletin (2011-)* **41**:125-128.
- 521 Carden, R. F., C. M. Carlin, F. Marnell, D. McElholm, J. Hetherington, and M. P. Gammell.
522 2011. Distribution and range expansion of deer in Ireland. *Mammal Review* **41**:313-
523 325.
- 524 Carden, R. F., A. D. McDevitt, F. E. Zachos, P. C. Woodman, P. O'Toole, H. Rose, N. T.
525 Monaghan, M. G. Campana, D. G. Bradley, and C. J. Edwards. 2012.
526 Phylogeographic, ancient DNA, fossil and morphometric analyses reveal ancient and
527 modern introductions of a large mammal: the complex case of red deer (*Cervus*
528 *elaphus*) in Ireland. *Quaternary Science Reviews* **42**:74-84.
- 529 Central Statistics Office. 2021. Environmental Indicators Ireland 2021. Retrieved from
530 <https://www.cso.ie>

- 531 Chadwick, A. H., P. R. Ratcliffe, and K. Abernethy. 1996. Sika deer in Scotland: density,
532 population size, habitat use and fertility - some comparisons with red deer. *Scottish*
533 *Forestry* **50**:8-16.
- 534 Dick, J. T., M. Freeman, J. Provan, and N. Reid. 2010. First record of free-living Reeves'
535 muntjac deer (*Muntiacus reevesi* (Ogilby 1839)) in Northern Ireland. *The Irish*
536 *Naturalists' Journal* **31**:151-151.
- 537 Dorazio, R. M. 2014. Accounting for imperfect detection and survey bias in statistical
538 analysis of presence-only data. *Global Ecology and Biogeography* **23**:1472-1484.
- 539 Drennan, M., A. Carson, and S. Crosse. 2005. Overview of animal production from pastures
540 in Ireland. Utilisation of grazed grass in temperate animal systems. *.in a satellite*
541 *workshop of the XXth International Grassland Congress, Cork.*
- 542 Elith, J., C. H. Graham, R. P. Anderson, M. Dudík, S. Ferrier, A. Guisan, R. J. Hijmans, F.
543 Huettmann, J. R. Leathwick, A. Lehmann, J. Li, L. G. Lohmann, B. A. Loiselle, G.
544 Manion, C. Moritz, M. Nakamura, Y. Nakazawa, J. McC. M. Overton, A. Townsend
545 Peterson, S. J. Phillips, K. Richardson, R. Scachetti-Pereira, R. E. Schapire, J.
546 Soberón, S. Williams, M. S. Wisz, and N. E. Zimmermann. 2006. Novel methods
547 improve prediction of species' distributions from occurrence data. *Ecography* **29**:129-
548 151.
- 549 Elith, J., and J. R. Leathwick. 2009. Species Distribution Models: Ecological Explanation and
550 Prediction Across Space and Time. *Annual Review of Ecology, Evolution, and*
551 *Systematics* **40**:677-697.
- 552 European Union, Copernicus Land Monitoring Service, 2018. European Environment
553 Agency.

- 554 Forsyth, D. M., S. Comte, N. E. Davis, A. J. Bengsen, S. D. Côté, D. G. Hewitt, N. Morellet,
555 and A. Mysterud. 2022. Methodology matters when estimating deer abundance: a
556 global systematic review and recommendations for improvements. *The Journal of*
557 *Wildlife Management*.
- 558 Frans, V. F., A. A. Augé, J. Fyfe, Y. Zhang, N. McNally, H. Edelhoff, N. Balkenhol, and J.
559 O. Engler. 2021. Integrated SDM database: Enhancing the relevance and utility of
560 species distribution models in conservation management. *Methods in Ecology and*
561 *Evolution* **13**:243-261.
- 562 Gortázar, C., R. J. Delahay, R. A. McDonald, M. Boadella, G. J. Wilson, D. Gavier-Widen,
563 and P. Acevedo. 2012. The status of tuberculosis in European wild mammals.
564 *Mammal Review* **42**:193-206.
- 565 Guillera-Aroita, G., J. J. Lahoz-Monfort, J. Elith, A. Gordon, H. Kujala, P. E. Lentini, M. A.
566 McCarthy, R. Tingley, and B. A. Wintle. 2015. Is my species distribution model fit
567 for purpose? Matching data and models to applications. *Global Ecology and*
568 *Biogeography* **24**:276-292.
- 569 Guisan, A., and W. Thuiller. 2005. Predicting species distribution: offering more than simple
570 habitat models. *Ecology Letters* **8**:993-1009.
- 571 Guisan, A., and N. E. Zimmermann. 2000. Predictive habitat distribution models in ecology.
572 *Ecological Modelling* **135**:147-186.
- 573 Hijmans, R. J. 2021. raster: geographic data analysis and modeling. R package version 3.4-
574 10.
- 575 Hortal, J., and J. M. Lobo. 2005. An ED-based Protocol for Optimal Sampling of
576 Biodiversity. *Biodiversity and Conservation* **14**:2913-2947.

- 577 Isaac, N. J. B., M. A. Jarzyna, P. Keil, L. I. Dambly, P. H. Boersch-Supan, E. Browning, S.
578 N. Freeman, N. Golding, G. Guillera-Arroita, P. A. Henrys, S. Jarvis, J. Lahoz-
579 Monfort, J. Pagel, O. L. Pescott, R. Schmucki, E. G. Simmonds, and R. B. O’Hara.
580 2019. Data Integration for Large-Scale Models of Species Distributions. *Trends in*
581 *Ecology & Evolution* **0**:1-12.
- 582 Krausman, P. R., and V. C. Bleich. 2013. Conservation and management of ungulates in
583 North America. *International Journal of Environmental Studies* **70**:372-382.
- 584 Langbein, J., R. Putman, and B. Pokorny. 2011. Traffic collisions involving deer and other
585 ungulates in Europe and available measures for mitigation. Pages 215-259 in M.
586 Apollonio, R. Andersen, and R. Putman, editors. *Ungulate Management in Europe:*
587 *Problems and Practices*. Cambridge University Press, Cambridge.
- 588 Lezama-Ochoa, N., M. G. Pennino, M. A. Hall, J. Lopez, and H. Murua. 2020. Using a
589 Bayesian modelling approach (INLA-SPDE) to predict the occurrence of the Spinetail
590 Devil Ray (*Mobular mobular*). *Scientific Reports* **10**:18822.
- 591 Lindgren, F., H. Rue, and J. Lindström. 2011. An explicit link between Gaussian fields and
592 Gaussian Markov random fields: the stochastic partial differential equation approach.
593 *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **73**:423-
594 498.
- 595 Linnell, J. D. C., B. Cretois, E. B. Nilsen, C. M. Rolandsen, E. J. Solberg, V. Veiberg, P.
596 Kaczensky, B. Van Moorter, M. Panzacchi, G. R. Rauset, and B. Kaltenborn. 2020.
597 The challenges and opportunities of coexisting with wild ungulates in the human-
598 dominated landscapes of Europe's Anthropocene. *Biological Conservation* **244**.
- 599 Linnell, J. D. C., and F. E. Zachos. 2011. Status and distribution patterns of European
600 ungulates: genetics, population history and conservation. Pages 12-53 in M.

- 601 Apollonio, R. Andersen, and R. Putman, editors. Ungulate Management in Europe:
602 Problems and Practices. Cambridge University Press, Cambridge.
- 603 Liu, Y., and M. Nieuwenhuis. 2014. An analysis of habitat-use patterns of fallow and sika
604 deer based on culling data from two estates in Co. Wicklow. *Irish Forestry* **71**:27 - 49.
- 605 Liu, Y., and M. Nieuwenhuis. 2018. The effect of roadside land-use on the occurrence of deer
606 vehicle collisions. *Irish Forestry* **75**:8-25.
- 607 Lobo, J. M., A. Jiménez-Valverde, and J. Hortal. 2010. The uncertain nature of absences and
608 their importance in species distribution modelling. *Ecography* **33**:103-114.
- 609 McDevitt, A. D., C. J. Edwards, P. O'Toole, P. O'Sullivan, C. O'Reilly, and R. F. Carden.
610 2009. Genetic structure of, and hybridisation between, red (*Cervus elaphus*) and sika
611 (*Cervus nippon*) deer in Ireland. *Mammalian Biology* **74**:263-273.
- 612 Miller, D. A. W., K. Pacifici, J. S. Sanderlin, B. J. Reich, and B. Gardner. 2019. The recent
613 past and promising future for data integration methods to estimate species'
614 distributions. *Methods in Ecology and Evolution* **10**:22-37.
- 615 Millsbaugh, J. J., J. R. Skalski, R. L. Townsend, D. R. Diefenbach, M. S. Boyce, L. P.
616 Hansen, and K. Kammermeyer. 2009. An Evaluation of Sex-Age-Kill (SAK) Model
617 Performance. *Journal of Wildlife Management* **73**:442-451.
- 618 Milner, J. M., C. Bonenfant, A. Mysterud, J.-M. Gaillard, S. Csányi, and N. C. Stenseth.
619 2006. Temporal and spatial development of red deer harvesting in Europe: biological
620 and cultural factors. *Journal of Applied Ecology* **43**:721-734.
- 621 Nagy-Reis, M., J. R. Reimer, M. A. Lewis, W. F. Jensen, and M. S. Boyce. 2021. Aligning
622 population models with data: Adaptive management for big game harvests. *Global
623 Ecology and Conservation* **26**.

- 624 Northern Ireland Statistics and Research Agency. 2021. Statistical Review of Northern
625 Ireland Agriculture 2020. Department of Agriculture, Environment and Rural Affairs.
- 626 Nyhus, P. J. 2016. Human–Wildlife Conflict and Coexistence. *Annual Review of*
627 *Environment and Resources* **41**:143-171.
- 628 O'Mara, F. P. 2012. The role of grasslands in food security and climate change. *Annals of*
629 *Botany* **110**:1263-1270.
- 630 O'Mara, F. P. 2012. The role of grasslands in food security and climate change. *Annals of*
631 *Botany* **110**:1263-1270.
- 632 Obidziński, A., P. Kiełtyk, J. Borkowski, L. Bolibok, and K. Remuszko. 2013. Autumn-
633 winter diet overlap of fallow, red, and roe deer in forest ecosystems, Southern Poland.
634 *Open Life Sciences* **8**:8-17.
- 635 Open Street Map contributors, 2017. Planet dump. Retrieved in January 2022 from
636 <https://planet.osm.org>
- 637 Pacifici, K., B. J. Reich, D. A. W. Miller, B. Gardner, G. Stauffer, S. Singh, A. McKerrow,
638 and J. A. Collazo. 2017. Integrating multiple data sources in species distribution
639 modeling: a framework for data fusion. *Ecology* **98**:840-850.
- 640 Paradinas, I., D. Conesa, A. López-Quílez, and J. M. Bellido. 2017. Spatio-Temporal model
641 structures with shared components for semi-continuous species distribution
642 modelling. *Spatial Statistics* **22**:434-450.
- 643 Pebesma, E. J. 2018. Simple Features for R: standardized support for spatial vector data. *The*
644 *R Journal* **10**:439-446.
- 645 Powerscourt, V. 1884. On the Acclimatization of the Japanese Deer at Powerscourt.
646 *Proceedings of the Zoological Society of London* **52**:207-209.

- 647 Purser, P., F. Wilson, and R. F. Carden. 2010. Deer and forestry in Ireland: a review of
648 current status and management requirements., Coillearnacha Dúchasacha (Woodlands
649 of Ireland) steering committee, Dublin, Ireland.
- 650 Putman, R., M. Apollonio, and R. Andersen. 2011. Ungulate Management in Europe:
651 Problems and Practices. Cambridge University Press, Cambridge.
- 652 Richardson, S., A. C. Mill, D. Davis, D. Jam, and A. I. Ward. 2020. A systematic review of
653 adaptive wildlife management for the control of invasive, non-native mammals, and
654 other human–wildlife conflicts. *Mammal Review* **50**:147-156.
- 655 Rue, H., S. Martino, and N. Chopin. 2009. Approximate Bayesian inference for latent
656 Gaussian models by using integrated nested Laplace approximations. *Journal of the*
657 *Royal Statistical Society. Series B: Statistical Methodology* **71**:319-392.
- 658 Simmonds, E. G., S. G. Jarvis, P. A. Henrys, N. J. B. Isaac, and R. B. O'Hara. 2020. Is more
659 data always better? A simulation study of benefits and limitations of integrated
660 distribution models. *Ecography* **43**:1413-1422.
- 661 Simpson, D. P., H. Rue, T. G. Martins, A. Riebler, and S. H. Sørbye. 2014. Penalising model
662 component complexity: A principled, practical approach to constructing priors.
663 *Statistical Science* **32**:1-42.
- 664 Smith, S. L., R. F. Carden, B. Coad, T. Birkitt, and J. M. Pemberton. 2014. A survey of the
665 hybridisation status of *Cervus* deer species on the island of Ireland. *Conservation*
666 *genetics* **15**:823-835.
- 667 Spake, R., C. Bellamy, R. Gill, K. Watts, T. Wilson, B. Ditchburn, F. Eigenbrod, and J.
668 Bennett. 2020. Forest damage by deer depends on cross-scale interactions between

- 669 climate, deer density and landscape structure. *Journal of Applied Ecology* **57**:1376-
670 1390.
- 671 Venter, O., E. W. Sanderson, A. Magrath, J. R. Allan, J. Beher, K. R. Jones, H. P.
672 Possingham, W. F. Laurance, P. Wood, B. M. Fekete, M. A. Levy, and J. E. Watson.
673 2016. Sixteen Years of Change in the Global Terrestrial Human Footprint and
674 Implications for Biodiversity Conservation. *Nature Communications*:12558.
- 675 Venter, O., E. W. Sanderson, A. Magrath, J. R. Allan, J. Beher, K. R. Jones, H. P.
676 Possingham, W. F. Laurance, P. Wood, B. M. Fekete, M. A. Levy, and J. E. Watson.
677 2018. Last of the Wild Project, Version 3 (LWP-3): 2009 Human Footprint, 2018
678 Release. NASA Socioeconomic Data and Applications Center (SEDAC), Palisades,
679 NY.
- 680 Ward, A. I. 2005. Expanding ranges of wild and feral deer in Great Britain. *Mammal Review*
681 **35**:165-173.
- 682 Ward, A. I., S. Richardson, and J. Mergeay. 2021. Reeves' muntjac populations continue to
683 grow and spread across Great Britain and are invading continental Europe. *European*
684 *Journal of Wildlife Research* **67**.
- 685 Williams, B. K. 2011. Adaptive management of natural resources--framework and issues. *J*
686 *Environ Manage* **92**:1346-1353.
- 687 Yackulic, C. B., R. Chandler, E. F. Zipkin, J. A. Royle, J. D. Nichols, E. H. Campbell Grant,
688 S. Veran, and R. B. O'Hara. 2013. Presence-only modelling using MAXENT: when
689 can we trust the inferences? *Methods in Ecology and Evolution* **4**:236-243.

690

691