



**Modelling the spatial variation of vital rates: an evaluation
of the strengths and weaknesses of correlative species
distribution models**

Journal:	<i>Diversity and Distributions</i>
Manuscript ID	DDI-2016-0432.R2
Manuscript Type:	Biodiversity Research
Date Submitted by the Author:	n/a
Complete List of Authors:	Suarez-Seoane, Susana; Universidad de Leon, Biodiversidad y Gestion Ambiental Alvarez-Martinez, Jose Manuel; National Museum of Natural Sciences, Spanish National Research Council (CSIC), Evolutionary Ecology; University of León, Biodiversisty and Environmental Management Wintle, Brendan; University of Melbourne, School of Botany Palacín, Carlos; Museo Nacional de Ciencias Naturales (CSIC), Ecología Evolutiva Alonso, Juan; Museo Nacional de Ciencias Naturales (CSIC), Ecología Evolutiva
Keywords:	breeding success, Great Bustard, otis tarda, species distribution models

SCHOLARONE™
Manuscripts

Running title: Spatial modelling of demographic parameters

Modelling the spatial variation of vital rates: an evaluation of the strengths and weaknesses of correlative species distribution models

Susana Suárez-Seoane^{1*}, Jose Manuel Álvarez-Martínez^{2,3}, Brendan A. Wintle⁴, Carlos Palacín² and Juan Carlos Alonso²

¹*Dep. of Biodiversity and Environmental Management. Facultad de Biología, Universidad de León, Campus de Vegazana s/n, 24071 León, Spain. s.seoane@unileon.es*

²*Dep. of Evolutionary Ecology. Museo Nacional de Ciencias Naturales (CSIC), José Gutiérrez Abascal 2, 28006 Madrid, Spain*

³*Environmental Hydraulics Institute IH Cantabria, C/ Isabel Torres 15, Parque Científico y Tecnológico de Cantabria, 39011 Santander, Spain*

⁴*School of Biosciences, University of Melbourne, Australia, 3010*

*Corresponding author (orcid.org/0000-0001-7656-4214)

1
2
3 26 **ABSTRACT**
4

5 27 **Aim.** Species distribution models based on breeding occurrence data allow identifying
6
7 28 both environmental drivers and geographic areas potentially relevant for breeding.
8
9 29 However, the interpretation of model predictions in terms of reproductive performance
10
11 30 should be further investigated, as this information is crucial for conservation planning.
12
13 31 We evaluated the strengths and weaknesses of a correlative modelling approach based
14
15 32 on breeding occurrence data (presence-absence) against another approach based on vital
16
17 33 rates' data (breeding success) for gaining insights on species persistence in the case of
18
19 34 Great Bustards (*Otis tarda*).
20
21

22
23 35 **Location.** Spain.
24

25 36 **Methods.** Breeding occurrence and breeding success were independently modelled
26
27 37 using generalized linear models and multi-model inference analyses. Sensitivities to the
28
29 38 way in which the population parameter (breeding success) was defined were explored
30
31 39 by building five versions of the dependent variable. We evaluated differences in model
32
33 40 performance and identified areas of congruence for breeding occurrence and breeding
34
35 41 success.
36
37

38 42 **Results.** The agreement between the spatial predictions achieved by breeding
39
40 43 occurrence and breeding success models differed substantially across databases, with
41
42 44 the largest differences in occupied breeding areas. The deviance explained by the
43
44 45 breeding occurrence model was 64.98% and ranged from 7.83% to 62.27% for the
45
46 46 breeding success models. Model performance was higher for models calibrated within
47
48 47 potential than within occupied breeding areas.
49
50

51 48 **Main conclusions.** The combination of data on both breeding occurrence and breeding
52
53 49 success into a species distribution modelling framework showed the limitations of
54
55 50 breeding occurrence models for inferring reproductive parameters. The definition of the
56
57
58
59
60

1
2
3 51 population parameter as dependent variable was a key factor that strongly affected the
4
5 52 inference of vital rates' models. The approach allowed for discriminating between areas
6
7 53 and landscape attributes necessary for the long-term species persistence from others that
8
9 54 may be relevant, but not so much for reproductive performance.
10
11 55 **Key words:** Breeding success, species distribution modelling, Great Bustard, *Otis*
12
13 56 *tarda*, population persistence.
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Review Only

1
2
3 57 **(A) INTRODUCTION**
4
5
6 58

7
8 59 Species Distribution Models (SDM; Guisan & Zimmerman, 2000) are useful tools for
9
10 60 exploring the factors driving species distribution. These techniques are increasingly
11
12 61 seen as crucial tools in species conservation and management (Suárez-Seoane *et al.*,
13
14 62 2002; Franklin, 2010; Austin & Van Niel 2011; Guisan *et al.*, 2014). However, the
15
16 63 reliability of SDM predictions and, therefore, their applicability in prioritizing
17
18 64 conservation efforts strongly depends on the data used for model calibration (Guisan *et*
19
20 65 *al.*, 2013).
21
22

23 66 In bird studies, SDM have been typically calibrated with species occurrence data
24
25 67 (presence-absence, presence-only and, less frequently, abundance) collected during the
26
27 68 breeding season, which allows producing breeding habitat suitability maps. Major
28
29 69 sources of occurrence data are, besides one's own field work, breeding bird atlases and
30
31 70 long-term monitoring programs, which are widely available at regional, national and
32
33 71 continental scales. Typical data from atlases allow discriminating, in a general grid
34
35 72 framework, between sites where species are likely (or even confirmed) to breed and
36
37 73 sites where breeding is unlikely. In the most recent approaches, this basic information is
38
39 74 complemented with data on species abundance at more detailed spatial resolution.
40
41 75 Numerous examples can be found across the literature where SDM are calibrated with
42
43 76 breeding bird atlas data; see, for example, Araújo *et al.* (2005), Virkkala *et al.* (2014),
44
45 77 Moudrý *et al.* (2017) or Howard *et al.* (2014). On the other hand, long-term monitoring
46
47 78 programs are primarily designed to provide data for evaluating population trends in
48
49 79 abundance, but can be also integrated in SDM to explore factors determining species
50
51 80 presence-absence and abundance (Brotons *et al.*, 2007).
52
53
54
55
56
57
58
59
60

1
2
3 81 However, despite the widespread use of SDM calibrated with breeding
4
5 82 occurrence data in conservation applications, this approach may have limitations. A
6
7 83 constraint arises from the fact that species occurrence data might be collected outside of
8
9 84 the species' reproductive niche (i.e., overall environmental requirements for successful
10
11 85 reproduction; Titeux *et al.*, 2007; Bykova *et al.*, 2012). When this occurs, SDM may
12
13 86 predict high suitability in areas not relevant for breeding or even in "sink" or "trap"
14
15 87 environments (Van Horne, 1983). Thus, the viability of the species is likely to be over-
16
17 88 estimated. In the case of species using different environments during the breeding
18
19 89 period, another problem emerges when occurrence data are collected in areas and
20
21 90 moments where species detectability is the highest (e.g., at the beginning of the
22
23 91 breeding season when birds arrive from migration and begin the occupation of
24
25 92 territories or when they are singing or displaying before the mating; Strebel *et al.*,
26
27 93 2014), but that are not central for offspring production. When the predictions of models
28
29 94 fitted to such data are used to prioritize conservation efforts, key biological
30
31 95 requirements may not be met by the resulting conservation strategy, as different parts of
32
33 96 a species' life cycle can only take place in certain environments.

34
35
36
37
38 97 Arguably, a main drawback of SDM calibrated with breeding occurrence data is
39
40 98 the interpretation of model predictions in terms of reproductive outputs. Identifying
41
42 99 areas where the breeding success of a given species is high, low or null is critical to
43
44 100 devise and implement effective conservation and management plans addressed to
45
46 101 guarantee species long-term persistence (Soga & Koike, 2013). In this sense, Brambilla
47
48 102 & Ficetola (2012) found that habitat suitability estimated through a presence-only SDM
49
50 103 correlated positively with two reproductive parameters of a passerine bird. However,
51
52 104 even if this approach can provide useful information for preliminary assessments of
53
54 105 breeding success at large-scale, it should be considered that the factors driving breeding
55
56
57
58
59
60

1
2
3 106 habitat selection (and therefore the estimates of habitat suitability) do not necessarily
4
5 107 influence in a similar way reproductive performance. In fact, a high level of mismatches
6
7 108 between observed avian breeding habitat preferences and fitness outcomes (breeding
8
9 109 success) have been identified across a wide variety of taxa (see Chalfoun & Schmidt,
10
11 110 2012 for a review on this topic).

11 111 An emerging response to these concerns is to seek mechanistic approaches (such
12
13 112 as metapopulation models) that are based on fundamental relationships and
14
15 113 dependencies and can provide a more robust way to predict species distribution than
16
17 114 correlative SDM. However, mechanistic models are highly parameterized and present
18
19 115 higher technical demands in terms of time, effort, resources and data for model
20
21 116 calibration and validation in comparison to correlative approaches (Kearny & Porter,
22
23 117 2009). Consequently, they are unable to compete with correlative SDM for widespread
24
25 118 application in conservation and management. In this context, hybrid
26
27 119 (mechanistic/correlative) approaches represent a good compromise between the
28
29 120 simplicity of correlative SDM and the benefits of being more complex mechanistic
30
31 121 models (Kearney *et al.* 2010; Michel, 2017). Fitting correlative models to spatial data on
32
33 122 population parameters, such as fecundity, and mapping the predictions of such models
34
35 123 across the landscape may provide a suitable compromise between simplicity and
36
37 124 robustness (*sensu* Falcucci *et al.*, 2009) when modelling reproductive performance.

38
39 125 However, modelling spatial variation in population parameters in the framework
40
41 126 of SDM is challenging. On one hand, collecting population data samples for model
42
43 127 calibration is much more time and effort consuming than collecting occurrence data.
44
45 128 Additionally, there are many technical choices to be made about the nature of the
46
47 129 dependent variable chosen to represent the population parameter of interest. For
48
49 130 example, in the case of birds, there are various ways to measure breeding success,
50
51
52
53
54
55
56
57
58
59
60

1
2
3 131 including: clutch size, hatching success, nesting success, breeding success per female or
4
5 132 group of females, general success of the breeding season and so on (Murray, 2000).
6
7 133 Indeed, the value of the dependent variable can take the form of counts, continuous
8
9 134 values and multiple or binary categorical values, which effectively increases model
10
11 135 structural uncertainty. While both the influence of predictor choice and model structural
12
13 136 uncertainty have received a great deal of attention in the SDM literature (e.g. Thuiller,
14
15 137 2003; Wintle *et al.*, 2003; Pearson *et al.*, 2006), we have find no studies investigating
16
17 138 how the definition of the dependent variable contributes to uncertainty in modelling
18
19 139 spatial variation in population parameters.
20
21

22
23 140 In this study, we aimed to compare the inference of a correlative species
24
25 141 distribution modelling approach based on species breeding occurrence data (i.e.,
26
27 142 presence-absence) against another correlative, but more mechanistic, approach based on
28
29 143 vital rates' data (i.e., breeding success) in order to gain insights on species persistence.
30
31 144 As study case, we used an endangered species, the Great Bustard (*Otis tarda*) across
32
33 145 Spain. We explored how the definition of the population parameter impacts on
34
35 146 ecological inference and predictions about key areas for conservation. We discussed the
36
37 147 implications of the modelling options and what the predictions and ecological inference
38
39 148 tell us about the biology of the species, as well as the best strategies for improving the
40
41 149 realism and applicability of species distribution modelling approaches to support
42
43 150 conservation decisions.
44
45

46
47 151

48 49 152 **(A) METHODS**

50
51 153

52 53 54 154 **(B) The study model: Great Bustards in Spain**

1
2
3 155 Great Bustards are large, lekking birds that live in highly fragmented populations in
4
5 156 cereal pseudo-steppes throughout the Palaearctic, from Morocco to eastern China
6
7 157 (Palacín & Alonso, 2008). Spain holds ca. 60-70% of the world population (Alonso &
8
9 158 Palacín, 2010). The species is globally threatened and classified as Vulnerable on the
10
11 159 Red List of Threatened Species (BirdLife International, 2015). Females nest on the
12
13 160 ground and rear their precocial chicks (usually one, sometimes two) alone over a period
14
15 161 of six to 12 months (Alonso *et al.*, 1998). Breeding success is highly variable, with
16
17 162 productivity values ranging between 0.04 and 0.53 chicks per female (Morales *et al.*,
18
19 163 2002; Martín *et al.*, 2007).
20
21
22
23

164

165 **(B) Great Bustard breeding occurrence and breeding success databases**

166 Great Bustard breeding occurrence was modelled using a database on presence-absence
167 records (PA database) compiled in Spring, when species detectability is maximal
168 (Alonso *et al.*, 2005), during the period 1987-2010 (Figure 1a). Presence data consisted
169 of all 350 lek centres known in Spain (Alonso *et al.*, 2012a). Absence data comprised an
170 equivalent sample of 350 points randomly distributed across Spain, excluding coastal
171 border areas, mountainous ranges and home ranges around lek centres (estimated with a
172 buffer of 2 km; Palacín *et al.*, 2012). These data can be considered as true absences,
173 since the distribution of the species at this time of the year is well known.

174 To estimate breeding success, we carried out population censuses across the
175 breeding range in Spain during the month of September (when the mortality peak of
176 juveniles is over and families can be detected due to their less elusive behaviour
177 compared to early summer; Martín *et al.*, 2007) of the same period 1987-2010 (Figure
178 1b, c). Breeding success was estimated annually as the young productivity (i.e., ratio of
179 the number of chicks to females) for each reproductive group (RG), which was made of

1
2
3 180 all flocks of females (FF) with or without chicks found in the same lek. Values $\geq 100\%$
4
5 181 were discarded because they were associated with very small groups consisting of one
6
7 182 or two females and their chicks, or corresponded to RG with a very low detectability
8
9 183 (i.e., where number of females counted in September was lower than 30% of those
10
11 184 counted in Spring). It should be considered that productivity data were heterogeneously
12
13 185 distributed across space and time during the study period (i.e., data were not available
14
15 186 for all RG all years). Therefore, the raw values were averaged for the whole series to
16
17 187 achieve a spatially and temporally consistent coverage of this population parameter,
18
19 188 reflecting the long-term trends of species persistence better than single measures made
20
21 189 for shorter (or even isolated) periods of time. See Alonso *et al.* (2005) and Álvarez-
22
23 190 Martínez *et al.* (2015) for more details on breeding success surveys and GIS database
24
25 191 preparation. We accounted for a total of 208 RGs.
26
27
28

29
30 192 In order to build dependent variables informing on breeding success, three
31
32 193 methodological criteria were applied on the original multi-temporal dataset, generating
33
34 194 five datasets on breeding success (BS databases; Table 1). Each database was based on
35
36 195 a particular combination of subjective and data-driven choices about the treatment of
37
38 196 dependent variables: (i) Dependent variables could be continuous (productivity values
39
40 197 ranged from 0 to 100%) or binary (productivity only had two values, which represent
41
42 198 high/low productivity, high/null productivity or positive/null productivity). (ii)
43
44 199 Continuous dependent variables were calculated by averaging annual productivity data
45
46 200 across the temporal series, using either the mean (database BS1) or the range of the
47
48 201 values (database BS2). (iii) Binary dependent variables were built by comparing: (iii.1)
49
50 202 RG with high productivity (locations where the productivity value was higher than the
51
52 203 averaged mean value for all RG across the whole study period; i.e., 0.15 young/female)
53
54 204 vs. RG with low productivity (locations collected across the occupied breeding area
55
56
57
58
59
60

1
2
3 205 where productivity was positive, but lower than the averaged mean value) (database
4
5 206 BS3). (iii.2) RG with high productivity (higher than the averaged mean) vs. locations
6
7 207 where productivity was null (i.e., a set of random points sampled across the potential
8
9 208 breeding distribution estimated by Suárez-Seoane *et al.*, 2002, avoiding a buffer of 2 km
10
11 209 -the species home range- around each FF) (database BS4). (iii.3) RG with positive
12
13 210 productivity (value greater than “0”; i.e., birds successfully bred, independently of the
14
15 211 number of chicks raised) vs. locations where productivity was null (database BS5).
16
17 212 Databases generated in occupied breeding areas (BS1, BS2 and BS3) are useful for
18
19 213 modelling “breeding performance”, as they allow for comparing different (but always
20
21 214 positive) values of breeding success. Databases generated also in potential breeding
22
23 215 areas (BS4 and BS5) are useful for modelling “breeding site selection”, as they allow
24
25 216 for comparing locations where birds had successfully bred from other potential sites
26
27 217 where breeding success was null.
28
29
30
31

32 218

33 219 **(B) Environmental variables**

34
35 220 On the basis of exploratory analyses, expert knowledge and published information
36
37 221 (Morales *et al.*, 2002; Alonso *et al.*, 2004; Pinto *et al.*, 2005; Martínez, 2008; Palacín *et*
38
39 222 *al.*, 2012), we selected a pool of 12 environmental GIS predictors potentially driving
40
41 223 both breeding occurrence and breeding success of Great Bustards in Spain (Table 2).
42
43 224 Predictors described topography, climate, primary production, landscape structure and
44
45 225 human disturbances. Scales ranged from 1:5000 to 1:200000 and pixel sizes from 25m
46
47 226 to 1km, depending on original data and methodological restrictions. All data were
48
49 227 rescaled to the same spatial resolution, matching the pixel size of 1km among variables.
50
51
52

53 228 The role of topographic predictors on breeding occurrence and breeding success
54
55 229 was evaluated through the slope and its variation among agricultural plots. Topography
56
57
58
59
60

1
2
3 230 largely influence visual communication with conspecifics, as the breeding system of
4
5 231 dispersed leks involves strong visual cues over long distances (Alonso *et al.*, 2012a).
6
7 232 Data sources were, respectively, a digital elevation model at 25m (CNIG;
8
9 233 <http://www.cnig.es>) and the Spanish Geographic Information System for Agricultural
10
11 234 Plots (SIGPAC; [http://www.magrama.gob.es/en/agricultura/temas/sistema-de-](http://www.magrama.gob.es/en/agricultura/temas/sistema-de-informacion-geografica-de-parcelas-agricolas-sigpac)
12
13 235 [informacion-geografica-de-parcelas-agricolas-sigpac](http://www.magrama.gob.es/en/agricultura/temas/sistema-de-informacion-geografica-de-parcelas-agricolas-sigpac)) at 1:5000 scale. Data on climate
14
15 236 and primary production were seasonally averaged (mean values and variation
16
17 237 coefficients) to assess the effect of these environmental factors during critical periods of
18
19 238 the year on breeding occurrence and breeding success. The effect of climate was
20
21 239 included through the maximum Summer temperature (temperature during the last days
22
23 240 of incubation and first days after hatching) and Autumn-Winter rainfall (precipitation
24
25 241 prior to breeding season), both extracted from the Climatic Map of the Iberian Peninsula
26
27 242 (Ninyerola *et al.*, 2005, 2007). The relevance of these climatic variables for Great
28
29 243 Bustards has been demonstrated previously by Morales *et al.* (2002) and Osborne *et al.*
30
31 244 (2007). According to these authors, winter precipitation controls the productivity of
32
33 245 annual plants during the following spring and, therefore, the availability of arthropods
34
35 246 during the period of chick maximum growth rate in early summer. The development of
36
37 247 herbaceous vegetation in early spring also contributes positively to the physiological
38
39 248 condition of females, which directly affects the percentage of females attempting to
40
41 249 breed, their clutch size and brood viability. The above-ground net primary production
42
43 250 was represented through the Normalized Difference Vegetation Index of Spring (NDVI;
44
45 251 Pettorelli *et al.*, 2007, 2011; Bro-Jørgensen *et al.* 2008; Hamel *et al.* 2009) obtained
46
47 252 from a temporal series of NOAA-AVHRR satellite imagery (1987-2010), that was
48
49 253 acquired from the SerGEO database from CCHS-CSIC
50
51 254 (<http://humanidades.cchs.csic.es/cchs/sig/sergeo.html>). Landscape structure was
52
53
54
55
56
57
58
59
60

1
2
3 255 quantified from the SIGPAC, that have been elaborated at 1:5000 scale for the year
4
5 256 2012. We created a grid of 1-km over the whole Spain that intersected with the SIGPAC
6
7 257 layer in order to calculate both the maximum perimeter of plots and the percentage of
8
9 258 arable land within each 1-km grid (authors' unpublished data). Human disturbances
10
11 259 were evaluated through the Euclidean-distance from each pixel to the nearest paved
12
13 260 road or highway (Suárez-Seoane *et al.*, 2002) and the land protection status. Data were
14
15 261 obtained, respectively, from the Spanish Centre of Geographic Information (CNIG;
16
17 262 <http://www.cnig.es>) at 1:200000 scale and the Nature 2000 Ecological Network
18
19 263 (<http://www.magrama.gob.es>). All GIS analyses were done in ArcGIS10.2 (ESRI,
20
21 264 2014).

22
23
24
25 265 Environmental features were gathered for each database on occurrence and
26
27 266 breeding success using a pixel-based approach. The spatial resolution of the analyses,
28
29 267 allowing for model inference across Spain, was 1km. In the case of BS databases, points
30
31 268 were spatially assigned to the location of the female flock with chicks (isolated family
32
33 269 or flock of females including at least one family) closest to the “centroid” of all female
34
35 270 flocks, with or without chicks, in a RG. Points were recalculated for each year during
36
37 271 the study period and averaged afterwards.

38
39
40 272 In order to avoid multi-collinearity problems that may lead to parameter bias
41
42 273 (Freckleton, 2011), we checked that Spearman's bivariate correlations among all
43
44 274 predictors were below 0.7 (Randin *et al.*, 2006), as well as that variance inflation factor
45
46 275 (VIF) in further modelling analyses was lower than 4. There is no formal VIF threshold,
47
48 276 but a value of 10 is commonly used as an indicator of severe multicollinearity (Neter *et*
49
50 277 *al.*, 1990; Graham, 2003; Zuur *et al.*, 2010).

51
52
53
54 278

55
56 279 **(B) Model building**
57
58
59
60

1
2
3 280 Both breeding occurrence and breeding success of Great Bustards in Spain were
4
5 281 independently modelled on the basis of the environmental features described in Table 2
6
7 282 using generalized linear models (GLM). Multi-model inference, based on AIC values,
8
9 283 and model averaging (Burnham & Anderson, 2002; Burnham *et al.*, 2011) were
10
11 284 implemented for each of the six datasets (PA, BS1 to BS5). This method allows for
12
13 285 selecting the best subset of approximating models (i.e., those the smallest AIC value,
14
15 286 indicating the most parsimonious models) among all possible candidates. Models with
16
17 287 $\Delta_i (AIC_{\text{best}} - AIC_i) \leq 2$ were considered substantially supported by the data and similar to
18
19 288 the best model in their empirical reliability. Using this subset of models, we estimated
20
21 289 the averaged standardised coefficients (β) for each predictor, as well as its significance
22
23 290 and relative importance. Model coefficients were standardised to allow comparisons
24
25 291 among predictors. The relative importance of each predictor was measured as the sum
26
27 292 of the Akaike weights of all models in the subset where that predictor was present. The
28
29 293 value of the summed Akaike weight of each predictor ranges from 0 (if it appears only
30
31 294 in the most unlikely models) to 1 (if it appears in all the best models) (Burnham and
32
33 295 Anderson 2002; Symonds & Moussalli, 2011). GLM were built using either a binomial
34
35 296 distribution with logit link, when the response variable was binomial, or a Gamma
36
37 297 distribution with log link, when the response variable was continuous. Model algorithms
38
39 298 were spatially projected across the study area using GIS tools, providing maps of habitat
40
41 299 suitability ranging from 0 to 1. Finally, variations in performance (deviance explained
42
43 300 by the best subset of variables against a null model), family of the most contributing
44
45 301 predictors and spatial predictions were evaluated across model outputs.
46
47
48
49
50

51 302 To assess the accuracy of model predictions, we built two validation datasets
52
53 303 consisting of 67 cases with the best reproductive performance across the study period
54
55 304 (RG with an average annual productivity higher than the average value for the study
56
57
58
59
60

1
2
3 305 period, with more than five years of available data) along with an equivalent sample of
4
5 306 points randomly chosen within either: (i) occupied breeding areas (home ranges in
6
7 307 Figure 1) or (ii) potential breeding areas (potential area defined by Suárez-Seoane *et al.*,
8
9 308 2002). We thus estimated, for each model output, the statistical significance of the
10
11 309 differences in habitat suitability across occupied and potential breeding areas (the larger
12
13 310 difference, the better is the discrimination capacity of a model).

14
15
16 311 We used these validation datasets to check the Pearson bivariate correlations
17
18 312 among prediction values achieved from PA and BS models in both occupied and
19
20 313 potential breeding areas. Finally, we mapped the areas of congruence between the
21
22 314 outputs achieved by the occurrence model (PA) and the breeding success models
23
24 315 calibrated in the potential area (BS4 and BS5). We have not compared PA with BS1,
25
26 316 BS2 and BS3 because the predictions made by these models are not applicable outside
27
28 317 the occupied area.

29
30
31 318 Analyses were done with the packages MASS and MuMIn from R 3.0.2
32
33 319 statistical software (R Development Core Team, 2014) and ArcGIS10.2 (ESRI, 2014).

34
35
36 320

37 38 321 **(A) RESULTS**

39
40
41 322

42
43 323 Table 3 summarises the results of the modelling approach. Performance varied widely
44
45 324 across model outcomes. In the case of the presence-absence (PA) model, the deviance
46
47 325 explained by the best subset of variables against a null model was 64.98%, being Great
48
49 326 Bustard occurrence significantly correlated with all families of predictors. In the case of
50
51 327 breeding success (BS) models, deviance ranged from 7.83% to 62.27%. The largest
52
53 328 differences were found between two groups of BS models. Models calibrated within
54
55 329 occupied breeding areas (BS1 to BS3) explained less deviance than those calibrated

1
2
3 330 within the potential distribution area (BS4 and BS5). The most relevant predictors of
4
5 331 breeding success also changed markedly between these two groups. In the former,
6
7 332 breeding success was significantly correlated with climate (temperature), primary
8
9 333 production and distance to IBAs. In the latter, breeding success was also related to
10
11 334 topography and landscape structure, instead of primary production.

12
13
14 335 The predictive performance of models fitted to mean (BS1) and range (temporal
15
16 336 variability) of breeding success (BS2) was similar. Nevertheless, mean breeding success
17
18 337 was driven by temperature, primary production and distance to IBAs, while breeding
19
20 338 success range was only significantly correlated with temperature. When modelling high
21
22 339 levels of breeding success (values above the mean for the period) in occupied (BS3)
23
24 340 *versus* potential (BS4) breeding areas, we found that the latter models had much higher
25
26 341 performance and comprised significant variables of different types, including
27
28 342 topography, temperature, landscape structure and human disturbances. In contrast, BS3
29
30 343 models only included primary production and human disturbances as significant
31
32 344 predictors. Models calibrated in potential breeding areas (BS4 and BS5) showed similar
33
34 345 predictive performance and were driven by the same families of predictors.

35
36
37
38 346 Spatial patterns of habitat suitability varied substantially across the pool of
39
40 347 models (Figure 2). The output of the PA model showed clear differences among suitable
41
42 348 and non-suitable areas at large scale. In the case of BS models, those calibrated in
43
44 349 occupied breeding areas (BS1 to BS3) presented homogeneous or even random patterns
45
46 350 of habitat suitability outside the reproductive areas, as predictions were only valid at
47
48 351 local scale (i.e., within home ranges; Figure 1b, c). Models calibrated in potential
49
50 352 breeding areas (BS4 and BS5) clearly identified differences between suitable and non-
51
52 353 suitable areas at large scale. Overall, the prediction values generated by BS4 and BS5
53
54
55
56
57
58
59
60

1
2
3 354 models were the most correlated to the predictions of PA model, in both occupied and
4
5 355 potential areas (Figures 3 and 4).
6

7 356 Model discrimination was lower in occupied than potential breeding areas. In the
8
9 357 former case, significant differences were only found for model BS3, while in the last
10
11 358 case, significant differences emerged for the PA model, as well as for all BS models
12
13 359 except for BS3 (Table 4).
14
15

16 360

17
18 361 **(A) DISCUSSION**
19

20 362

21
22
23 363 Many authors have highlighted the need to move beyond static correlative predictions of
24
25 364 species occurrence probability or relative likelihood of occurrence to model processes
26
27 365 that are more directly related to the long-term persistence of species (Guisan & Thuiller,
28
29 366 2005; Thuiller *et al.*, 2008; Zurrell *et al.*, 2009; Franklin, 2010; Álvarez-Martínez *et al.*,
30
31 367 2015). The anticipated benefits of modelling processes closely linked to species
32
33 368 persistence is that it will improve the robustness of predictions about species future
34
35 369 ranges in rapidly changing environments. In this sense, the current study pioneers a
36
37 370 comparison of the ecological inference arising from a classical correlative species
38
39 371 distribution modelling approach, based on presence-absence data, with that of another
40
41 372 approach in which correlative models are fitted to the spatial variation in a population
42
43 373 parameter, in this case, breeding success. Our results provided insights into the specific
44
45 374 strengths and weaknesses of SDM regarding their application in conservation biology.
46
47
48

49 375 Correlative models based on occurrence data have been formerly demonstrated as
50
51 376 valuable conservation tools for designing management actions aimed to promote patch
52
53 377 occupancy (e.g. Suárez-Seoane *et al.*, 2002) and density of focal species (García *et al.*,
54
55 378 2007). The application of such models implies several practical strengths, as they have
56
57
58
59
60

1
2
3 379 low input data needs, avoid the challenges of scaling up from individual to landscape
4
5 380 level and allow for evaluating niche tolerance limits at large scale (Peterson *et al.*,
6
7 381 2016). However, as stated by Oliver *et al.* (2012), these models overlook important
8
9 382 features for long-term population persistence, such as population stability and source-
10
11 383 sink dynamics.

12
13
14 384 In this sense, the incorporation of long temporal series of vital rates' data in SDM
15
16 385 allowed for achieving well performing models (deviance ranged from 7.83% to 62.27%;
17
18 386 see Donázar *et al.*, 2002 and Rodríguez & Bustamante, 2003 for comparison with other
19
20 387 studies dealing with breeding success modelling), that are useful for the identification of
21
22 388 landscape attributes contributing to population stability. We also detected that certain
23
24 389 environmental relationships affecting breeding success were overlooked in PA models.
25
26 390 For example, temporal variability in Summer maximum temperature was a significant
27
28 391 driver of breeding success, while it did not appear to strongly influence species
29
30 392 occurrence. This suggests that climatic stability is a demand much more evident for
31
32 393 breeding than for surviving. A similar scenario was found for the mean primary
33
34 394 production, which was not significantly correlated to species occurrence, but appeared
35
36 395 as one of the most important explanatory predictors of breeding success (BS1 and BS3
37
38 396 models). This indicates that primary production influences where and when to breed, as
39
40 397 it indirectly reflects the availability of food for chicks, but may be less important for
41
42 398 adult survival. These differences should be explicitly considered when planning
43
44 399 conservation measures that may fail if they are not undertaken in areas that nurture
45
46 400 long-term breeding success.

47
48
49
50
51 401 Indeed, the comparison between the habitat suitability patterns achieved by PA
52
53 402 and BS models allowed for the discrimination of critical areas for the long term
54
55 403 persistence of the species (i.e., areas potentially supporting high breeding performance
56
57
58
59
60

1
2
3 404 during the study period) from other areas that may be occupied periodically or
4
5 405 consistently in other life stages, but have a low suitability for reproduction and then, for
6
7 406 population maintenance. In this sense, the identification of marginal areas for species
8
9 407 persistence is a matter of conservation priority, due to the high risk of local extinction.
10
11 408 For example, the regions where BS models predicted much higher suitability than PA
12
13 409 models (highlighted in black in Figure 4) are those where most local extinctions have
14
15 410 occurred during recent decades (79% of 29 extinctions documented in 1960-1990 in the
16
17 411 whole of Spain have occurred in the northeastern and southwesternmost black patches),
18
19 412 due to hunting pressure and agricultural transformations (Alonso *et al.* 2003). The
20
21 413 message for conservation managers is that, with appropriate environmental measures,
22
23 414 the high potential breeding success in these areas would help restoring the original
24
25 415 populations. The larger geographic area predicted as suitable by the PA model, when
26
27 416 compared with BS models, could be associated to the fact that correlative models might
28
29 417 overestimate niche breadth by not constraining the niche to account for breeding needs
30
31 418 (Peterson *et al.*, 2016). Titeux *et al.* (2007) reached a similar conclusion when exploring
32
33 419 the role of incorporating fitness parameters (i.e., breeding success) in the definition of
34
35 420 niche boundaries for red-clacked shrike.

36
37 421 Modelling choices when building the dependent variable influenced the inference
38
39 422 and potential utility of the model outputs for environmental managers and decision-
40
41 423 makers. In this sense, Mostashari & Sussman (2005) propose a stakeholder-assisted
42
43 424 modelling process in which stakeholders participate through contributions of input and
44
45 425 feedback to the modelling process to improve the representation of focal systems.
46
47 426 Discussions should begin in the earliest stages of the ecological modelling process, as
48
49 427 they are essential for identifying the key choices for model building, mitigating many of
50
51 428 the subsequent problems that arise from inappropriate methodological decisions (Martin
52
53
54
55
56
57
58
59
60

1
2
3 429 *et al.*, 2012). For example, we showed here that models fitted using continuous variants
4
5 430 of breeding success within occupied breeding areas (BS1 and BS2) explained a greater
6
7 431 proportion of deviance than the binary model calibrated across occupied area (BS3),
8
9 432 being thus more indicative of species persistence. In addition, while binary models of
10
11 433 databases BS4 and BS5 explained much more deviance, we should consider that they
12
13 434 were calibrated in non-occupied areas and, therefore, they are actually more useful to
14
15 435 evaluate breeding sites' selection than species persistence. The ability of BS1, BS2 and
16
17 436 BS3 models to find even slight differences within currently occupied areas casts light on
18
19 437 critical parameters for species persistence linked to spatial and temporal environmental
20
21 438 restrictions and feeding resource availability (Alvarez-Martínez *et al.*, 2015).

22
23
24
25 439 A key modelling choice was whether the dependent variable should be continuous
26
27 440 (databases BS1 and BS2) or binary (BS3 to BS5). While the predictive performance of
28
29 441 BS1 (general pattern; mean value of breeding success across the temporal series) and
30
31 442 B2 (temporal variation; range value) models was similar, they were driven by a different
32
33 443 set of predictors. Areas of consistently good breeding performance over time, and
34
35 444 therefore of high species persistence, would be selected by Great Bustards' based on
36
37 445 many interacting environmental and social cues (Parejo *et al.*, 2006; Osborne *et al.*,
38
39 446 2007; Rieucan & Giraldeau, 2011). Temporal variation in breeding success was linked
40
41 447 to Summer maximum temperatures, which may be interpreted as the negative effect that
42
43 448 severe Summer droughts can have over breeding performance by limiting food
44
45 449 availability for the rearing chicks.

46
47
48
49 450 The choice of converting continuous values of breeding success into binary
50
51 451 dependent variables (BS3 to BS5) implies summarising the observed variability
52
53 452 according to unique thresholds, which involves a loss of information. This decision is
54
55 453 strongly linked to vagueness, a form of linguistic uncertainty which refers to the
56
57
58
59
60

1
2
3 454 possibility of borderline cases which neither satisfy a criterion, nor its negation, when
4
5 455 performing categorical classifications of data (Regan *et al.*, 2002). In such a situation,
6
7 456 the development of fuzzy logic sets and rules (Ascough II *et al.*, 2008; Álvarez-
8
9 457 Martínez *et al.*, 2010) is recommended to reduce the level of vagueness in decision-
10
11 458 making. If this is not possible, a deliberate consideration of the meaning and
12
13 459 implications of the chosen thresholds is advised.

14
15
16 460 The most influential methodological choice in terms of model performance,
17
18 461 selected predictors and spatial pattern of habitat suitability across the pool of results was
19
20 462 the area of calibration for binary BS variables. In fact, Jiménez-Valverde *et al.* (2009)
21
22 463 and Lobo *et al.* (2010) already highlighted the key importance of carefully choosing the
23
24 464 area of calibration depending on the question at hand. Models calibrated with binary
25
26 465 variables collected in occupied breeding areas (BS3) explained much less deviance than
27
28 466 models calibrated across non-occupied sites within the potential distribution area; BS4
29
30 467 and BS5). This is a common and expected result, since describing local differences
31
32 468 within suitable areas using models narrowly calibrated across space and environment
33
34 469 (i.e., comparing between areas of higher-than-average and those of lower-than-average
35
36 470 breeding success) is more demanding than assessing coarse differences between areas of
37
38 471 positive breeding performance and potential areas including many unsuitable pockets.
39
40 472 The lower deviance explained by BS3 model could be associated to the fact that short-
41
42 473 term environmental changes that may affect breeding success in suitable breeding areas
43
44 474 cannot be anticipated by birds at local scale. For example, the temporal variation in
45
46 475 Summer maximum temperature was significantly correlated with breeding success in
47
48 476 BS4 and BS5, but not in BS3 models. This fact suggest that Great Bustards might
49
50 477 “predict” general patterns of climatic stability across their range and use it as an
51
52 478 indicator of quality or suitability when looking for breeding areas through “public
53
54
55
56
57
58
59
60

1
2
3 479 information” that can be gained from the presence of conspecifics (Osborne *et al.*,
4
5 480 2007). However, birds cannot identify local differences in climatic stability within
6
7 481 reproductive areas. On the top of this, it should be considered that certain differences in
8
9 482 the variables selected by these models might be, at least partially, related to the extent of
10
11 483 the area where models were calibrated (Van Der Wal *et al.*, 2009). Nevertheless, model
12
13 484 calibration across different scenarios (implying different extents) was necessary to test
14
15 485 our hypothesis.

16
17
18 486 In this study, we have combined spatially explicit data from both species
19
20 487 occurrence and vital rates into a SDM framework to identify priority conservation areas
21
22 488 and landscape structures supporting population maintenance. The utility of this
23
24 489 approach would be greatly enhanced through the integration of temporal fluctuations in
25
26 490 other vital rates, such as survival and dispersal, using metapopulation models. Such an
27
28 491 approach would allow the identification of the threats most likely influencing
29
30 492 population persistence over time horizons relevant to management decision-making.
31
32 493 However, the extra data and expertise demands of such an approach preclude its
33
34 494 application in many instances (Franklin, 2010). We argue that our approach may
35
36 495 represent a suitable compromise between mechanistic models and the simpler
37
38 496 correlative SDM based on occurrence data.
39
40
41
42

43 497

44 45 498 **(A) ACKNOWLEDGMENTS**

46
47 499

48
49
50 500 The Spanish Ministry for Science and Innovation (project CGL2008-02567) supported
51
52 501 this research, although many censuses were funded by earlier projects from that
53
54 502 Ministry, as well as from ICONA, Junta de Castilla y León, Junta de Andalucía and
55
56 503 Gobierno de Navarra awarded to JCA since 1987. We thank M. Alcántara, J.A. Arranz,
57
58
59
60

1
2
3 504 A. Balmori, B. Campos, F.J. Carmona, J. Ezquerro, M.J. García-Baquero, M. Guerrero,
4
5 505 E. Izquierdo, J. Larumbe, A.I. Lasheras, J.E. Montero, A. Onrubia, J. Panadero, A.
6
7 506 Sánchez, C.A. Sánchez, C. Torralbo, A Torrijo and R. Ubaldo for providing additional
8
9 507 census data, as well as other anonymous persons who carried out censuses in several
10
11 508 areas. J.A. Alonso, C. Bravo, M. Magaña, B. Martín, C.A. Martín, E. Martín, M.
12
13 509 Morales, C. Ponce and A. Torres collaborated during many regional surveys. We are
14
15 510 also grateful to I. Prieto (University of León) and L. Mateos (ESRI) for their help with
16
17 511 environmental data handling. A. Cabria and J.M. Rodríguez (CNIG) provided the 5m-
18
19 512 DEM and GIS databases. I. del Bosque from CCHS-CSIC provided remote sensing data
20
21 513 and other GIS variables (SerGEO database collection). Thanks to P.E. Osborne
22
23 514 (University of Southampton) for his useful comments. J.M. Álvarez-Martínez was
24
25 515 supported by a contract from project CGL2008-02567 of the Spanish Ministry for
26
27 516 Science and Innovation and B.E Wintle by an Australian Research Council Future
28
29 517 Fellowship (FT100100819). This work was partially carried out at the Mixed Unit of
30
31 518 Research on Biodiversity (UMIB, University of Oviedo, Mieres, Spain). Finally, we
32
33 519 also thank to R. Heikkinen, Adam B. Smith and an anonymous referee for their useful
34
35 520 comments that helped to improve the quality of the manuscript.
36
37
38
39
40
41
42

522 BIOSKETCH

523 The main research topic of **Susana Suárez-Seoane** is the effect of land use change on
524 landscape dynamics and functioning. She is particularly focused on the analysis of
525 vertebrate species distribution patterns using remote sensing and species distribution
526 modelling techniques. The work presented in this manuscript is framed within the
527 project CGL2008-02567 entitled 'Human impacts on Great Bustard population
528 dynamics', funded by the Spanish Ministry for Science and Innovation and led by Prof.

1
2
3 529 Juan Carlos Alonso (www.proyectoavutarda.org). The overall aim of this project was to
4
5 530 assess human impacts on the viability of Great Bustards in the Iberian Peninsula, the
6
7 531 last stronghold of this globally endangered species.
8

9
10 532

11 533 Author contributions: S.S.S., J.M.A.M., B.W. and J.C.A. originally formulated the ideas
12
13 534 presented in this paper. J.C.A. and C.P. obtained the species data. J.M.A.M. was in
14
15 535 charge of data analysis. S.S.S., J.M.A.M. and B.W. wrote the first draft of this
16
17 536 manuscript and all authors contributed extensively to the preparation of the final
18
19
20 537 version.
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

538 **REFERENCES**

- 539 Alonso, J.C. & Palacín, C. (2010) The world status and population trends of the Great
540 Bustard: 2010 update. *Chinese Birds*, **1**, 141-147.
- 541 Alonso, J.C., Martín, E., Alonso, J.A. & Morales, M.B. (1998) Proximate and ultimate
542 causes of natal dispersal in the Great Bustard *Otis tarda*. *Behavioral Ecology*, **9**,
543 243-252.
- 544 Alonso J.C., Palacín, C. & Martín, C.A. (2003) Status and recent trends of the Great
545 Bustard (*Otis tarda*) population in the Iberian Peninsula. *Biological Conservation*,
546 **110**, 185-195.
- 547 Alonso, J.C., Martín, C.A., Alonso, J.A., Palacín, C., Magaña, M. & Lane, S.J. (2004)
548 Distribution dynamics of a Great Bustard metapopulation throughout a decade:
549 influence of conspecific attraction and recruitment. *Biodiversity and Conservation*,
550 **13**, 1659-1674.
- 551 Alonso, J.C., Palacín, C. & Martín, C.A. (2005) La avutarda común en la península
552 Ibérica: población actual y método de censo. SEO/BirdLife, Madrid.
- 553 Alonso, J.C., Álvarez-Martínez, J.M. & Palacín, C. (2012a) Leks in ground-displaying
554 birds: hotspots or safe places? *Behavioral Ecology*, **23**, 491-501.
- 555 Alonso, J.C., Magaña, M. & Álvarez-Martínez, J.M. (2012b) Male display areas in
556 exploded leks: the importance of food resources for male mating success.
557 *Behavioral Ecology*, **23**, 1296-1307.
- 558 Álvarez-Martínez, J.M., Stoorvogel, J.J., Suárez-Seoane, S. & de Luis, E. (2010)
559 Uncertainty analysis as a tool for refining land dynamics modelling on changing
560 landscapes: a case study in a Spanish Natural Park. *Landscape Ecology*, **25**, 1385-
561 1404.

- 1
2
3 562 Álvarez-Martínez, J.M., Suárez-Seoane, S., Palacín, C., Sanz, J. & Alonso, J.C. (2015)
4
5 563 Can Eltonian processes explain species distributions at large scale? A case study
6
7 564 with Great Bustard (*Otis tarda*). *Diversity and Distributions*, **21**, 123-138.
8
9 565 Araújo, M.B., Thuiller, W., Williams, P. H. & Reginster, I. (2005) Downscaling
10
11 566 European species atlas distributions to a finer resolution: implications for
12
13 567 conservation planning. *Global Ecology and Biogeography*, **14**, 17-30.
14
15
16 568 Ascough II, J.C., Maier, H.R., Ravalico J.K & Strudley, M.W. (2008) Future research
17
18 569 challenges for incorporation of uncertainty in environmental and ecological
19
20 570 decision-making. *Ecological Modelling*, **219**, 383-399.
21
22
23 571 Austin, M.P. & Van Niel, K.P. (2011) Improving species distribution models for climate
24
25 572 change studies: variable selection and scale. *Journal of Biogeography*, **38**, 1-8.
26
27 573 BirdLife International (2015) *Otis tarda*. The IUCN Red List of Threatened Species
28
29 574 2015: e.T22691900A85068116. Office for Official Publications of the European
30
31 575 Communities, Luxembourg. Downloaded on 20 July 2016.
32
33
34 576 Brambilla, M. & Ficetola, G.F. (2012) Species distribution models as a tool to estimate
35
36 577 reproductive parameters: a case study with a passerine bird species. *Journal of*
37
38 578 *Animal Ecology*, **81**, 781-787.
39
40
41 579 Bro-Jørgensen, J., Brown, M.E. & Pettorelli, N. (2008) Using the satellite-derived
42
43 580 normalized difference vegetation index (NDVI) to explain ranging patterns in a lek-
44
45 581 breeding antelope: the importance of scale. *Oecologia*, **158**, 177-182.
46
47 582 Brotons, L., Herrando, S. & Pla, M. (2007) Updating bird species distribution at large
48
49 583 spatial scales: applications of habitat modelling to data from long-term monitoring
50
51 584 programs. *Diversity and Distributions*, **13**, 276-288.
52
53
54 585 Burnham, K.P. & Anderson, D.R. (2002) Model selection and multi-model inference. A
55
56 586 practical information-theoretic approach. New York: Springer.
57
58
59
60

- 1
2
3 587 Burnham, K.P., Anderson, D.R. & Huyvaert, K.P. (2011) AIC model selection and
4
5 588 multimodel inference in behavioral ecology: some background, observations, and
6
7 589 comparisons. *Behavioral Ecology and Sociobiology*, **65**, 23-35.
8
9
10 590 Bykova, O., Chuine, I., Morin, X. & Higgins, S.I. (2012) Temperature dependence of
11
12 591 the reproduction niche and its relevance for plant species distributions. *Journal of*
13
14 592 *Biogeography*, **39**, 2191-2200.
15
16 593 Chalfoun, A.D. & Schmidt, K.A. (2012) Adaptive breeding-habitat selection: Is it for
17
18 594 the birds? *The Auk*, **129**, 589-599.
19
20 595 Donázar, J.A., Blanco, G., Hiraldo, F., Soto-Largo, E. & Oria, J. (2002) Effects of
21
22 596 forestry and other land use practices on the conservation of cinereous vultures.
23
24 597 *Ecological Applications*, **12**, 1445-1456.
25
26 598 ESRI (2014) ArcInfo desktop GIS 10.2. <http://www.esri.com>
27
28 599 Falcucci, A., Ciucci, P., Maiorano, L., Gentile, L. & Boitani, L. (2009) Assessing
30
31 600 habitat quality for conservation using an integrated occurrence-mortality model.
32
33 601 *Journal of Applied Ecology*, **46**, 600-609.
34
35 602 Franklin, J. (2010) Moving beyond static species distribution models in support of
36
37 603 conservation biogeography. *Diversity and Distributions*, **16**, 321-330.
38
39 604 Freckleton, R. (2011) Dealing with collinearity in behavioural and ecological data:
40
41 605 model averaging and the problems of measurement error. *Behavioral Ecology and*
42
43 606 *Sociobiology*, **65**, 91-101.
44
45 607 García, J., Suárez-Seoane, S., Miguélez, D., Osborne, P.E. & Zumalacárregui, C. (2007)
46
47 608 Spatial analysis of habitat quality in a fragmented population of little bustard (*Tetrax*
48
49 609 *tetrax*): implications for conservation. *Biological Conservation*, **137**, 45-56.
50
51 610 Graham, M.H. (2003) Confronting multicollinearity in ecological multiple regression.
52
53 611 *Ecology*, **84**, 2809-2815.
54
55
56
57
58
59
60

- 1
2
3 612 Guisan, A. & Thuiller, W. (2005) Predicting species distribution: offering more than
4
5 613 simple habitat models. *Ecology Letters*, **8**, 993-1009.
6
7 614 Guisan, A. & Zimmermann, N.E. (2000) Predictive habitat distribution models in
8
9 615 ecology. *Ecological Modelling*, **135**, 147-186.
10
11 616 Guisan, A. *et al.* (2013) Predicting species distributions for conservation decisions.
12
13 617 *Ecology Letters*, **16**, 1424-1435.
14
15
16 618 Guisan, A., Petitpierre, B., Broennimann, O., Daehler, C. & Kueffer, C. (2014)
17
18 619 Unifying niche shift studies: insights from biological invasions. *Trends in Ecology*
19
20 620 *& Evolution*, **29**, 260-269.
21
22
23 621 Hamel, S., Garel, M., Festa-Bianchet M., Gaillard, J.M. & Côté, S.D. (2009) Spring
24
25 622 Normalized Difference Vegetation Index (NDVI) predicts annual variation in timing
26
27 623 of peak faecal crude protein in mountain ungulates. *Journal of Applied Ecology*, **46**,
28
29 624 582-589.
30
31
32 625 Howard, C., Stephens, P.A., Pearce-Higgins, J.W., Gregory, R.D. & Willis, S.G. (2014)
33
34 626 Improving species distribution models: the value of data on abundance. *Methods in*
35
36 627 *Ecology and Evolution*, **5**, 506-513.
37
38 628 Jiménez-Valverde, A., Peterson, A.T., Soberón, J., Overton, J.M., Aragón, P. & Lobo,
39
40 629 J.M. (2011) Use of niche models in invasive species risk assessments. *Biological*
41
42 630 *Invasions*, **13**, 2785-2797.
43
44
45 631 Kearney, M. & Porter, W. (2009) Mechanistic niche modelling: combining
46
47 632 physiological and spatial data to predict species' ranges. *Ecology Letters*, **12**, 334-
48
49 633 350.
50
51
52 634 Kearney, M., Simpson, S.J., Raubenheimer, D. & Helmuth, B. (2010) Modelling the
53
54 635 ecological niche from functional traits. *Philosophical Transactions of the Royal*
55
56 636 *Society - Series B*, **365**, 3469-3483.
57
58
59
60

- 1
2
3 637 Lobo, J.M., Jiménez-Valverde, A. & Hortal, J. (2010) The uncertain nature of absences
4
5 638 and their importance in species distribution modelling. *Ecography*, **33**, 103-14.
6
7 639 Martín, C.A., Alonso, J.C., Alonso, J.A., Palacín, C., Magaña, M. & Martín, B. (2007)
8
9 640 Sex-biased juvenile survival in a bird with extreme size dimorphism, the Great
10
11 641 Bustard. *Journal of Avian Biology*, **38**, 335-346.
12
13 642 Martínez, C. (2008) Distribution, density and productivity of great bustards in
14
15 643 northwestern Spain: a regional approach. *Journal of Ornithology*, **149**, 507-14.
16
17 644 Michel, M.J., Chien, H., Beachum, C.E., Bennett, M.G. & Knouft, J.H. (2017) Climate
18
19 645 change, hydrology, and fish morphology: predictions using phenotype-environment
20
21 646 associations. *Climatic Change*, **140**, 563-576.
22
23 647 Morales, M.B., Alonso, J.C. & Alonso, J.A. (2002) Annual productivity and individual
24
25 648 female reproductive success in a Great Bustard population. *Ibis*, **144**, 293-300.
26
27 649 Mostashari, A. & Sussman, J. (2005) Stakeholder-assisted modelling and policy design
28
29 650 process for environmental decision-making. *Journal of Environmental Assessment*
30
31 651 *Policy and Management*, **7**, 355-386.
32
33 652 Moudrý, V., Komárek, J. & Šimová, P. (2017) Which breeding bird categories should
34
35 653 we use in models of species distribution? *Ecological Indicators*, **74**, 526-529.
36
37 654 Murray Jr, B.G. (2000) Measuring annual reproductive success in birds. *The Condor*,
38
39 655 **102**, 470-473.
40
41 656 Neter, J., Wasserman, W. & Kutner, M.H. (1990) Applied linear statistical models:
42
43 657 regression, analysis of variance, and experimental designs. Irwin, Homewood, IL.
44
45 658 Ninyerola, M., Pons, X. & Roure, J.M. (2005) Atlas Climático Digital de la
46
47 659 Península Ibérica. Metodología y aplicaciones en bioclimatología y geobotánica.
48
49 660 ISBN 932560-8-7. Universidad Autónoma de Barcelona, Bellaterra.
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 661 Ninyerola, M., Pons, X. & Roure, J.M. (2007) Monthly precipitation mapping of the
4
5 662 Iberian Peninsula using spatial interpolation tools implemented in a Geographic
6
7 663 Information System. *Theoretical and Applied Climatology*, **89**, 195-209.
8
9
10 664 Oliver, T. H., Gillings, S., Girardello, M., Rapacciuolo, G., Brereton, T.M.,
11
12 665 Siriwardena, G.M., Roy, David B., Pywell, R. & Fuller, R.J. (2012) Population
13
14 666 density but not stability can be predicted from species distribution models. *Journal*
15
16 667 *of Applied Ecology*, **49**, 581-590.
17
18 668 Osborne, P.E., Suárez-Seoane S. & Alonso, J.C. (2007) Behavioural mechanisms that
19
20 669 undermine species envelope models: the causes of patchiness in the distribution of
21
22 670 Great Bustards *Otis tarda* L. in Spain. *Ecography*, **6**, 819-829.
23
24
25 671 Palacín, C. & Alonso, J.C. (2008) An updated estimate of the world status and
26
27 672 population trends of the Great Bustard *Otis tarda*. *Ardeola*, **55**, 13-25.
28
29 673 Palacín, C., Alonso, J.C., Martin, C.A. & Alonso, J.A. (2012) The importance of
30
31 674 traditional farmland areas for steppe birds: a case study of migrant female Great
32
33 675 Bustards *Otis tarda* in Spain. *Ibis*, **154**, 85-95.
34
35
36 676 Parejo, D., Oro, D. & Danchin, E. (2006) Testing habitat copying in breeding habitat
37
38 677 selection in a species adapted to variable environments. *Ibis*, **148**, 146-154.
39
40 678 Pearson, R.G., Thuiller, W., Araújo, M.B., Martínez-Meyer, E., Brotons, L, McClean,
41
42 679 C., Miles, L., Segurado, P., Dawson, T.P. & Lees, D.C. (2006) Model-based
43
44 680 uncertainty in species range prediction. *Journal of Biogeography*, **33**, 1704-1711.
45
46
47 681 Peterson, A.T., Papeş, M. & Soberón, J. (2016) Mechanistic and Correlative Models of
48
49 682 Ecological Niches. *European Journal of Ecology*, **1**, 28-38.
50
51 683 Pettorelli, N., Pelletier, F., Hardenberg, A.V., Festa-Bianchet, M. & Côté, S.D. (2007)
52
53 684 Early onset of vegetation growth vs. rapid green-up: impacts on juvenile mountain
54
55 685 ungulates. *Ecology*, **88**, 381-390.
56
57
58
59
60

- 1
2
3 686 Pettorelli, N., Ryan, S., Mueller, T., Bunnefeld, N., Jedrzejewska, B., Lima, M. &
4
5 687 Kausrud, K. (2011) The Normalized Difference Vegetation Index (NDVI):
6
7 688 unforeseen successes in animal ecology. *Climate Research*, **46**, 15-27.
8
9
10 689 Pinto, M., Rocha, P. & Moreira, F. (2005) Long-term trends in Great Bustard
11
12 690 populations in Portugal suggest concentration in single high quality area. *Biological*
13
14 691 *Conservation*, **124**, 415-23.
15
16 692 R Development Core Team. (2011) R: A language and environment for statistical
17
18 693 computing. R Foundation for Statistical Computing, Vienna (Austria).
19
20 694 <http://www.R-project.org>
21
22
23 695 Randin, C.F., Dirnböck, T., Dullinger, S., Zimmermann, N.E., Zappa, M. & Guisan, A.
24
25 696 (2006) Are niche-based species distribution models transferable in space? *Journal of*
26
27 697 *Biogeography*, **33**, 1689-1703.
28
29
30 698 Regan, H.M., Colyvan, M. & Burgman, M.A. (2002) A taxonomy and treatment of
31
32 699 uncertainty for ecology and conservation biology. *Ecological Applications*, **12**, 618-
33
34 700 628.
35
36 701 Rieucou, G. & Giraldeau, L.A. (2011) Exploring the costs and benefits of social
37
38 702 information use: an appraisal of current experimental evidence. *Phil. Trans. R. Soc.*
39
40 703 *B*, **366**, 949-957.
41
42
43 704 Rodríguez, C. & Bustamante, J. (2003) The effect of weather on lesser kestrel breeding
44
45 705 success: can climate change explain historical population declines? *Journal of*
46
47 706 *Animal Ecology*, **72**, 793-810.
48
49
50 707 Soga M. & Koike S. (2013) Large forest patches promote breeding success of a
51
52 708 terrestrial mammal in urban landscapes. *PLoS ONE*, **8**, e51802.
53
54
55
56
57
58
59
60

- 1
2
3 709 Strebel, N., Kéry, M., Schaub, M. & Schmid, H. (2014) Studying phenology by flexible
4
5 710 modelling of seasonal detectability peaks. *Methods in Ecology and Evolution*, **5**,
6
7 711 483-490.
- 8
9 712 Symonds, M. R. & Moussalli, A. (2011) A brief guide to model selection, multimodel
10
11 713 inference and model averaging in behavioural ecology using Akaike's information
12
13 714 criterion. *Behavioral Ecology and Sociobiology*, **65**, 13-21.
- 15
16 715 Suárez-Seoane, S., Osborne, P.E. & J.C. Alonso (2002) Large-scale habitat selection by
17
18 716 agricultural steppe birds in Spain: identifying species-habitat responses using
19
20 717 generalized additive models. *Journal of Applied Ecology*, **39**, 755-771.
- 22
23 718 Thuiller, W. (2003) BIOMOD-optimizing predictions of species distributions and
24
25 719 projecting potential future shifts under global change. *Global Change Biology*, **9**,
26
27 720 1353-62.
- 28
29 721 Thuiller, W., Albert, C., Araújo, M. B., Berry, P. M., Cabeza, M., Guisan, A., Hickler,
30
31 722 T., Midgley, G.F., Paterson, J., Schurr, F.M., Sykes, M.T. & Zimmermann, N.E.
32
33 723 (2008) Predicting global change impacts on plant species' distributions: future
34
35 724 challenges. *Perspectives in Plant Ecology, Evolution and Systematics*, **9**, 137-152.
- 36
37 725 Titeux, N., Dufrene, M., Radoux, J., Hirzel, A. H. & Defourny, P. 2007. Fitness-related
38
39 726 parameters improve presence-only distribution modelling for conservation practice:
40
41 727 the case of the red-backed shrike. *Biological Conservation*, **138**, 207-223.
- 42
43 728 Van Der Wal, J., Shoo, L.P., Graham, C. & Williams, S.E. 2009. Selecting pseudo-
44
45 729 absence data for presence-only distribution modeling: How far should you stray
46
47 730 from what you know? *Ecological Modeling*, **220**, 589-594.
- 48
49 731 Van Horne, B. (1983) Density as a misleading indicator of habitat quality. *The Journal*
50
51 732 *of Wildlife Management*, **47**, 893-901.
- 52
53
54
55
56
57
58
59
60

- 1
2
3 733 Virkkala, R., Pöyry, J., Heikkinen, R. K., Lehikoinen, A., & Valkama, J. (2014)
4
5 734 Protected areas alleviate climate change effects on northern bird species of
6
7 735 conservation concern. *Ecology and Evolution*, **4**, 2991-3003.
8
9 736 Wintle, B.A., McCarthy, M.A., Volinsky, C.T. & Kavanagh, R.P. (2003) The use of
10
11 737 Bayesian Model Averaging to better represent the uncertainty in ecological models.
12
13 738 *Conservation Biology*, **17**, 1579-1590.
14
15
16 739 Zuur, A.F., Ieno, E.N. & Elphick, C.S. (2010) A protocol for data exploration to avoid
17
18 740 common statistical problems. *Methods in Ecology and Evolution*, **1**, 3-14.
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1. Methodological criteria used to define dependent variables informing on the population parameter to be modelled. Each column represents a different database on breeding success (BS) elaborated from the original dataset by applying different decisions. For example, in the case of database BS1, the dependent variable consisted of continuous values and the statistical parameter used to average the temporal series of productivity data was the mean. The table also includes the sample size of each dataset.

Methodological criteria	BS1	BS2	BS3	BS4	BS5
(i) General type of dependent variable	Continuous		Binary		
(ii) Continuous dependent variable	Mean (<i>n</i> =208)	Range (<i>n</i> =208)			
(iii) Binary dependent variable:			High productivity (> averaged mean for the whole period; 0.15 young/female) (<i>n</i> =98)	Positive productivity (>0) (<i>n</i> =208)	
-High or positive productivity			Occupied breeding sites Low productivity (0< productivity < averaged mean for the period) (<i>n</i> =110)	Potential breeding sites Null productivity (=0) (<i>n</i> = 110)	
-Low or null productivity				(n=208)	

Table 2. Environmental predictors used to model both breeding occurrence and breeding success of Great Bustard. Spring (SP) includes March, April and May; Summer (SU): June, July and August; and Autumn-Winter (WI): September to February. CV is the coefficient of variation = $[(SD/\bar{x}) * 100]$, where SD is the standard deviation.

Family	Variable	Code	Units	Source
Topography	Slope of the terrain	SLO	degrees	Digital elevation model (DEM) from the Spanish Centre of Geographic Information (CNIG) at 25m of spatial resolution
	Slope variation among agricultural plots	SLOSPcv	degrees	GIS database of agricultural plots (SIGPAC) at 1:5000 scale, year 2012
Climate	Maximum temperature of Summer (mean)	TMAmSU	°C	Iberian Climatic Map at 200m of spatial resolution, period 1950-1999 (Ninyerola <i>et al.</i> 2005, 2007)
	Maximum temperature of Summer (CV)	TMAcvSU	°C	
	Rainfall of Autumn-Winter (mean)	PPsmWI	mm	
	Rainfall of Autumn-Winter (CV)	PPcvWI	mm	
Primary production	NDVI of Spring (mean)	NDVI _m SP	dimensionless (-1, 1)	NOAA-AVHRR at 1km of spatial resolution, period 1987-2010 (SerGEO database from CCHS-CSIC)
	NDVI of Spring (CV)	NDVI _c vSP	dimensionless (-1, 1)	
Landscape structure	Maximum perimeter of plots	PERIM _{max}	m	GIS database of agricultural plots (SIGPAC) at 1:5000 scale, year 2012
	Arable land (% occupation)	ARLAND	%	
Human disturbances	Distance to paved roads and highways	DISTRO	m	Spanish Centre of Geographic Information (CNIG) at 1:200000 scale
	Distance to IBAs (Important Bird Areas)	DISTIBA	m	

Table 3. Results of multi-model averaging for presence-absence (PA) and breeding success (BS1 to BS5) models (see definitions of variables in Table 2). Each cell shows the sign, the full model-averaged standardised coefficients (β) (with shrinkage) \pm their standard errors multiplied by 100, the significance categories (*** $p < 0.001$, ** $p < 0.01$, * $p < 0.1$) and, in parenthesis, the relative importance of each variable estimated as the sum of the AIC weights for each predictor included in n subsets of models having Δ_i (AIC_{best}-AIC_i) ≤ 2 . Deviance explained in relation to the null model was calculated for a final model including the best subset of variables. We also show the values corresponding to a final GLM model including variables retained after multi-model inference.

Family	Variable	PA	BS1	BS2	BS3	BS4	BS5
Topography	SLO	+0.65 \pm 12.94*** (1)	+0.01 \pm 0.09 (0.05)	+0.45 \pm 0.44 (0.67)	+0.87 \pm 7.7 (0.05)	-415 \pm 157.88** (1)	-449.45 \pm 140.94** (1)
	SLOSPcv	-4.01 \pm 71.83*** (1)	-	+0.69 \pm 0.43 (0.87)	-2.72 \pm 12.58 (0.11)	-90.79 \pm 45.07* (1)	-112.19 \pm 35.73** (1)
Climate	TMAmSU	+0.02 \pm 0.87* (0.1)	+0.7 \pm 0.43* (0.92)	+1.18 \pm 0.47* (1)	+6.18 \pm 18.5 (0.17)	-25.84 \pm 51.27 (0.33)	-97.42 \pm 41.6* (1)
	TMAcvSU	-0.55 \pm 30.29 (0.64)	+0.06 \pm 0.21 (0.18)	-1.35 \pm 0.45** (1)	-1.38 \pm 9.05 (0.06)	-238.34 \pm 69.87*** (1)	-214.49 \pm 46.37*** (1)
	PPsmWI	-0.01 \pm 0.02** (1)	-	-0.23 \pm 0.41 (0.36)	-	-3.19 \pm 25.77 (0.07)	-
	PPcvWI	-0.10 \pm 3.48** (1)	-0.32 \pm 0.38 (0.58)	+0.18 \pm 0.35 (0.31)	-0.49 \pm 6.93 (0.05)	-13.16 \pm 39.12 (0.24)	-8.73 \pm 24.66 (0.22)
Primary production	NDVImSP	-	-0.98 \pm 0.35** (1)	-	-73.63 \pm 30.32* (1)	-	+1.73 \pm 15.32 (0.13)
	NDVIcvSP	+15.63 \pm 500.90** (1)	-0.01 \pm 0.1 (0.06)	-0.07 \pm 0.22 (0.16)	-5.34 \pm 17.11 (0.16)	-44.06 \pm 64.83 (0.46)	-3.99 \pm 18.86 (0.15)
Landscape structure	PERIMmax	+0.02 \pm 1.43 (0.47)	+0.53 \pm 0.41 (0.82)	+0.02 \pm 0.13 (0.07)	+3.72 \pm 15.77 (0.12)	+5.89 \pm 37.25 (0.08)	-1.91 \pm 17.34 (0.13)
	ARLAND	+0.04 \pm 0.56*** (1)	+0.05 \pm 0.19 (0.14)	+0.01 \pm 0.08 (0.03)	-1.07 \pm 8.39 (0.06)	+582.53 \pm 108.23*** (1)	+479.32 \pm 64.64*** (1)
Human disturbances	DISTRO	+0.01 \pm 0.01 (0.59)	+0.5 \pm 0.36 (0.88)	-	+31.06 \pm 33.06 (0.65)	+127.9 \pm 61.13* (1)	+82.81 \pm 34.01* (1)
	DISTIBA	+0.01 \pm 0.01*** (1)	+0.8 \pm 0.35* (1)	+0.07 \pm 0.23 (0.16)	+85.31 \pm 33.02* (1)	-116.5 \pm 48.94* (1)	-195.65 \pm 36.06*** (1)
Global Model	Nb of predictors	11	10	10	11	11	11
	Dev. explained	64.98	19.88	21.11	7.83	62.27	59.02
	Adj.r.squared	0.79	0.21	0.24	0.14	0.77	0.74

Table 4. Spatial verification of the models in occupied vs. potential breeding areas (i.e., breeding areas with successful reproduction vs. potential breeding areas not used for reproduction). The values shown are the mean (\pm standard deviation) habitat suitability for highly productive RG and random points, as well as the difference in suitability between them (the larger this difference, the better is the discrimination capacity of the models) and the statistical significance of these differences (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$) when the value is positive. PA represents the presence-absence model and BS the breeding success models.

	PA	BS1	BS2	BS3	BS4	BS5
Occupied breeding areas						
Suitability of highly productive RG	0.812 \pm 0.220	0.059 \pm 0.011	0.050 \pm 0.019	0.455 \pm 0.14	0.781 \pm 0.272	0.786 \pm 0.266
Suitability of random points	0.834 \pm 0.220	0.063 \pm 0.009	0.057 \pm 0.017	0.401 \pm 0.108	0.825 \pm 0.233	0.84 \pm 0.219
Mean difference of suitability	-0.023	-0.004	-0.007	0.054**	-0.044	-0.055
Potential breeding areas						
Suitability of highly productive RG	0.811 \pm 0.220	0.059 \pm 0.011	0.05 \pm 0.019	0.455 \pm 0.14	0.781 \pm 0.272	0.786 \pm 0.266
Suitability of random points	0.355 \pm 0.353	0.049 \pm 0.021	0.043 \pm 0.022	0.629 \pm 0.19	0.236 \pm 0.328	0.211 \pm 0.297
Mean difference of suitability	0.457***	0.010*	0.007*	-0.174	0.544***	0.575***

1
2
3 773 **Figure legends**

4
5 774

6
7 775 Figure 1. Geographic range of Great Bustard in Spain during the period 1987-2010 at
8
9
10 776 different seasons. Back dots correspond to: (a) Lek centres identified in Spring (Alonso
11
12 777 et al. 2012b), (b) flocks of females with chicks (isolated family or flock of females with
13
14 778 at least one family) detected in September, (c) flocks of females with a number of
15
16 779 chicks higher than the averaged mean value for the study period (Álvarez-Martínez *et*
17
18 780 *al.* 2015). In (b) and (c), points were defined by the location of the female flock with
19
20 781 chicks closest to the “centroid” of all female flocks, with or without chicks, constituting
21
22 782 a reproductive group. Grey buffers represent the species home range, estimated as a
23
24 783 buffer of 2 km (Palacín *et al.* 2012) around either lek centres or family flocks for either
25
26 784 Spring or September.
27
28

29 785

30
31
32 786 Figure 2. Spatial patterns of habitat suitability achieved from occurrence (PA) and
33
34 787 breeding success (BS) models.
35

36 788

37
38 789 Figure 3. Pearson correlation matrices between habitat suitability for breeding
39
40 790 occurrence (PA) and breeding success (BS1 to BS5) in: (a) occupied breeding area, (b)
41
42 791 potential breeding area. The graph also shows histograms and scattered plots.
43
44

45 792

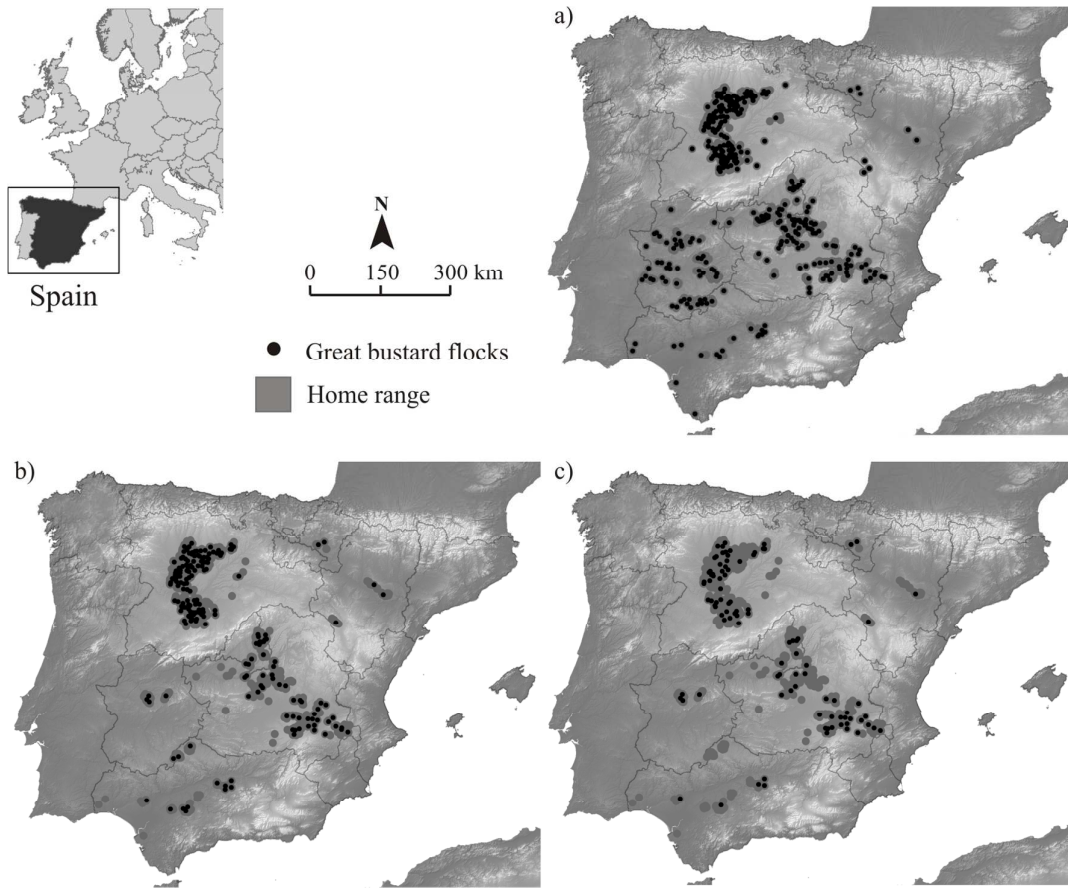
46
47 793 Figure 4. Spatial agreement between model outputs for occurrence (PA model) and
48
49 794 breeding success (BS4 and BS5 models).
50

51 795

52 796
53
54
55
56
57
58
59
60

797 **Figure 1**

798



799

800

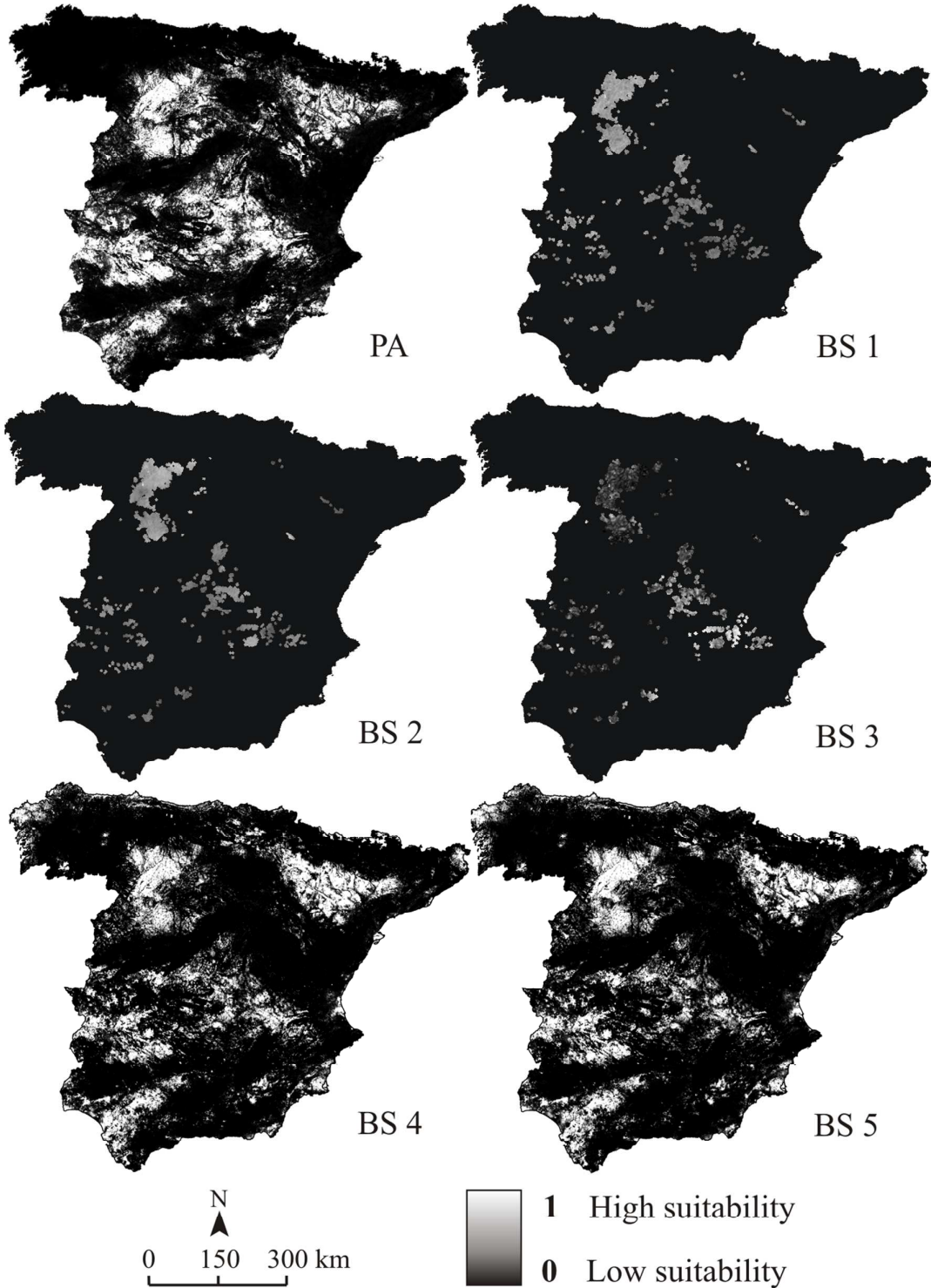
801

Only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

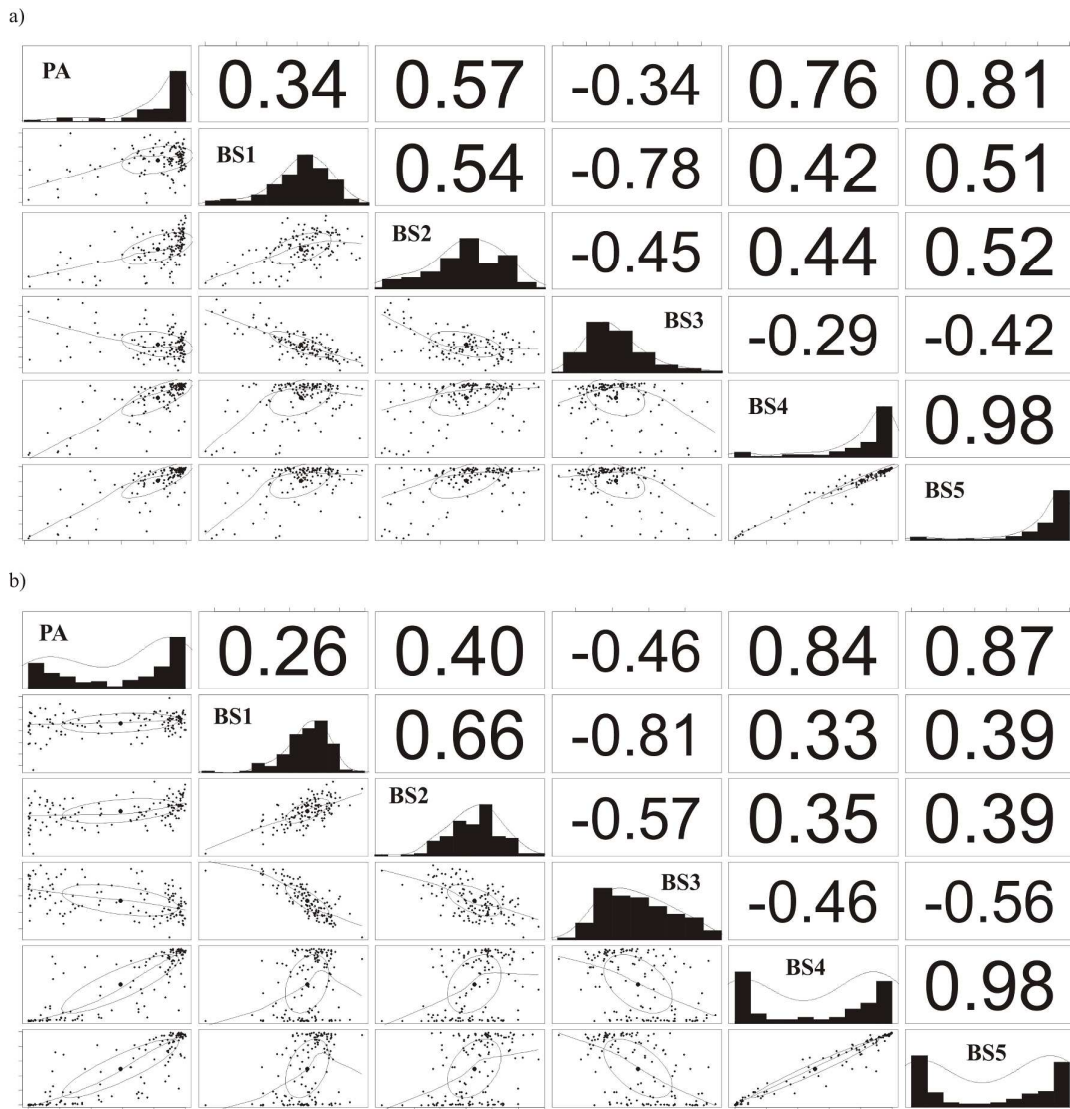
802 Figure 2



803

804 **Figure 3**

805



806

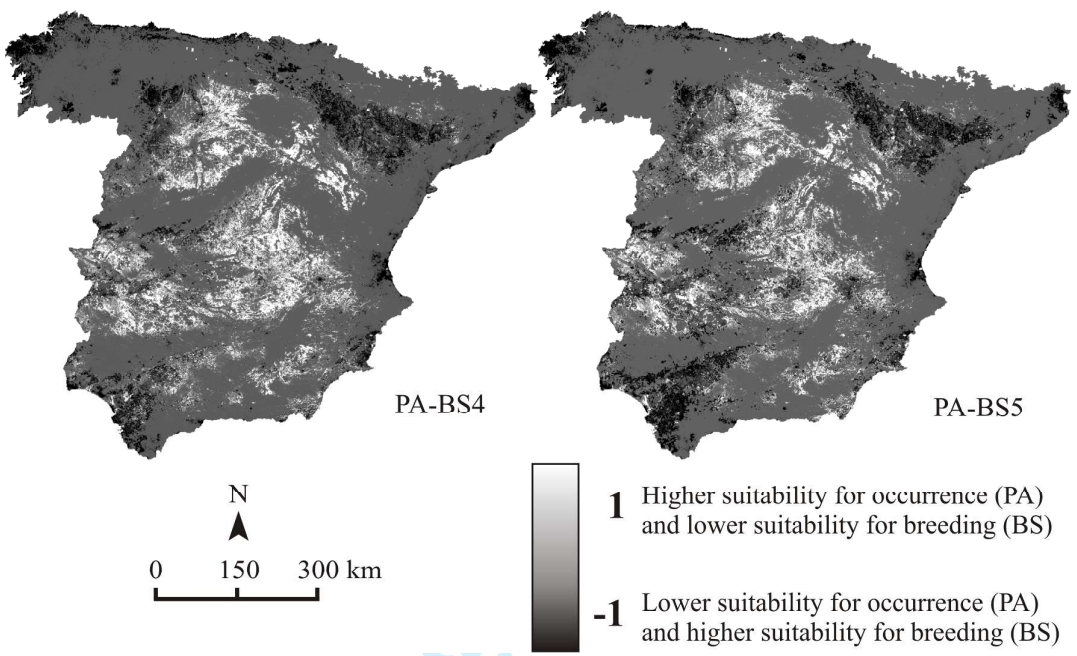
807

808

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

809 **Figure 4**

810



811

812

view Only

26 **ABSTRACT**

27 **Introduction**~~**Aim.**~~ Basic Species distribution models based on breeding occurrence data
28 allow ~~for exploring the factors driving species distribution. They can be used to~~
29 identifying both environmental drivers and geographic areas potentially relevant for
30 breeding. ~~However, they do not allow for discriminating between areas of different~~
31 ~~reproductive performance, which may limit their usefulness for conservation and~~
32 ~~management decisions.~~ However, the interpretation of model predictions in terms of
33 reproductive performance should be further investigated, as this information is crucial
34 for conservation planning. ~~**Aim.**~~ We evaluated the strengths and weaknesses of a
35 correlative ~~species distribution~~ modelling approach based on breeding occurrence data
36 (presence-absence) against another approach based on vital rates' data (breeding
37 success) for gaining insights on species persistence in the case of Great Bustards (*Otis*
38 *tarda*) ~~in Spain.~~

39 **Location.** Spain.

40 **Methods.** Breeding occurrence and breeding success were independently modelled
41 using generalized linear models and multi-model inference analyses. Sensitivities to the
42 way in which the population parameter (breeding success) was defined were explored
43 by building five versions of the dependent variable. We evaluated differences in model
44 performance and identified areas of congruence for breeding occurrence and breeding
45 success.

46 **Results.** The agreement between the spatial predictions achieved by breeding
47 occurrence and breeding success models differed substantially across databases, with
48 the largest differences ~~between models calibrated within~~ occupied ~~vs. potential~~
49 breeding areas. The deviance explained by the breeding occurrence model was 64.98%

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50 and ranged from 7.83% to 62.27% for the breeding success models. Model performance
51 was higher for models calibrated within potential than within occupied breeding areas.
52 **Conclusions**Main conclusions. The combination of ~~spatially explicit~~ data on both
53 breeding occurrence and ~~vital rates~~breeding success into a species distribution
54 modelling framework showed the limitations of breeding occurrence models for
55 inferring reproductive parameters. The definition of the population parameter as
56 dependent variable was a key factor that strongly affected the inference of vital rates'
57 models. The approach allowed for discriminating between areas and landscape
58 attributes necessary for the long-term species persistence from others that may be
59 relevant ~~in certain parts of the life cycle~~, but not so much for ~~reproduction~~reproductive
60 performance.
61 **Key words:** Breeding success, ~~SDM~~species distribution modelling, Great Bustard, *Otis*
62 *tarda*, population persistence.

1
2
3 63 | **(A) INTRODUCTION**
4
5
6 64

7
8 65 Species Distribution Models (SDM; Guisan & Zimmerman, 2000) are useful tools for
9
10 66 exploring the factors driving species distribution. These techniques are increasingly
11
12 67 seen as crucial tools in species conservation and management (Suárez-Seoane *et al.*,
13
14 68 2002; Franklin, 2010; Austin & Van Niel 2011; Guisan *et al.*, 2014). However, the
15
16 69 reliability of SDM predictions and, therefore, their applicability in prioritizing
17
18 70 conservation efforts strongly depends on the data used for model calibration (Guisan *et*
19
20
21 71 *al.*, 2013).
22

23
24 72 In bird studies, SDM have been typically calibrated with species occurrence data
25
26 73 (presence-absence, presence-only and, less frequently, abundance) collected during the
27
28 74 breeding season, which allows producing breeding habitat suitability maps. Major
29
30 75 sources of occurrence data are, besides one's own field work, breeding bird atlases and
31
32 76 long-term monitoring programs, which are widely available at regional, national and
33
34 77 continental scales. Typical data from atlases allow discriminating, in a general grid
35
36 78 framework, between sites where species are likely (or even confirmed) to breed and
37
38 79 sites where breeding is unlikely. In the most recent approaches, this basic information is
39
40 80 complemented with data on species abundance at more detailed spatial resolution.
41
42
43 81 Numerous examples can be found across the literature where SDM are calibrated with
44
45 82 breeding bird atlas data; see, for example, Araújo *et al.* (2005), Virkkala *et al.* (2014),
46
47 83 Moudrý *et al.* (2017) or Howard *et al.* (2014). On the other hand, long-term
48
49 84 monitoring programs are primarily designed to provide data for evaluating population
50
51
52 85 trends in abundance, but can be also integrated in SDM to explore factors determining
53
54 86 species presence-absence and abundance (Brotons *et al.*, 2007).
55
56
57
58
59
60

1
2
3 87 However, despite the widespread use of SDM calibrated with breeding
4
5 88 occurrence data in conservation applications, this approach may have limitations. A
6
7 89 constraint arises from the fact that species occurrence data might be collected outside of
8
9
10 90 the species' reproductive niche (i.e., overall environmental requirements for successful
11
12 91 reproduction; Titeux *et al.*, 2007; Bykova *et al.*, 2012). When this occurs, SDM may
13
14 92 predict high suitability in areas not relevant for breeding or even in "sink" or "trap"
15
16 93 environments (Van Horne, 1983). Thus, the viability of the species is likely to be over-
17
18 94 estimated. In the case of species using different environments during the breeding
19
20 95 period, another problem emerges ~~In such situations, the viability of the species is likely~~
21
22 96 ~~to be over-estimated.~~ (Pulliam, 2000; Titeux *et al.*, 2007; Álvarez-Martínez *et al.*, 2015).
23
24 97 ~~When this occurs, SDM may predict areas of high suitability for breeding in what could~~
25
26 98 ~~be considered "sink" or "trap" environments (Van Horne, 1983). This problem is~~
27
28 99 ~~particularly acute~~ when occurrence data are collected in areas and moments where
29
30
31
32 100 species detectability is the highest (e.g., at the beginning of the breeding season when
33
34 101 birds arrive from migration and begin the occupation of territories or when they are
35
36 102 singing or displaying before the mating; Strebel *et al.*, 2014), but that are not central for
37
38 103 offspring production~~breeding success~~. When the predictions of models fitted to such
39
40 104 data are used to prioritize conservation efforts, key biological requirements may not be
41
42 105 met by the resulting conservation strategy, as different parts of a species' life cycle can
43
44 106 only take place in certain environments. ~~In such situations, the viability of the species is~~
45
46 107 ~~likely to be over-estimated.~~

48
49 108 Arguably, ~~the a main limitation~~drawback for conservation applications of SDM
50
51 109 calibrated with breeding occurrence data is the interpretation of model predictions in
52
53 110 terms of reproductive outputs. Identifying areas where the breeding success of a given
54
55 111 species is high, low or null ~~In fact, this approach ignores the key component of species~~
56
57
58
59
60

1
2
3 112 ~~environment modelling, as it does not allow for identifying areas where the breeding~~
4
5 113 ~~success of a given species is high, low or null. Understanding~~ In fact, this approach
6
7 114 ignores the key component of species-environment modelling, as it does not allow for
8
9 115 identifying areas where the breeding success of a given species is high, low or null. the
10
11 116 ~~effects of landscape parameters on species breeding success~~ is critical to devise and
12
13 117 implement effective conservation and management plans addressed to guarantee species
14
15 118 long-term ~~species~~ persistence (Soga & Koike, 2013). ~~In fact, this approach ignores the~~
16
17 119 ~~key component of species-environment modelling, as it does not allow for identifying~~
18
19 120 ~~areas where the breeding success of a given species is high, low or null. In this sense,~~
20
21 121 Brambilla & Ficetola (2012) found that habitat suitability estimated through a presence-
22
23 122 only SDM correlated positively with two reproductive parameters of a passerine bird.
24
25 123 However, even if this approach can provide useful information for preliminary
26
27 124 assessments of breeding success at large-scale, it should be considered that the factors
28
29 125 driving breeding habitat selection (and therefore the estimates of habitat suitability) do
30
31 126 not necessarily influence in a similar way reproductive performance. In fact, a high
32
33 127 level of mismatches between observed avian breeding habitat preferences and fitness
34
35 128 outcomes (breeding success) have been identified across a wide variety of taxa (see
36
37 129 Chalfoun & Schmidt, 2012 for a review on this topic).

38
39
40
41
42
43 130 An emerging response to ~~this~~ these concerns is to seek ~~more~~ mechanistic
44
45 131 approaches (such as metapopulation models) that are based on fundamental
46
47 132 relationships and dependencies and can provide a more robust way to predict species
48
49 133 distribution than correlative SDM. However, mechanistic models are highly
50
51 134 parameterized and present higher technical demands in terms of time, effort, resources
52
53 135 and data for model calibration and validation in comparison to correlative approaches
54
55 136 (Kearny & Porter, 2009). Consequently, they are unable to compete with correlative

1
2
3 137 SDM for widespread application in conservation and management. In this context,
4
5 138 hybrid (mechanistic/correlative) approaches represent a good compromise between the
6
7 139 simplicity of correlative SDM and the benefits of being more complex mechanistic
8
9 140 models (Kearney *et al.* 2010; Michel, 2017). Fitting correlative models to spatial data on
10
11 141 population parameters, such as fecundity, and mapping the predictions of such models
12
13 142 ~~to variation in correlates~~ across the landscape may provide a suitable compromise
14
15 143 between simplicity and robustness (*sensu* Falcucci *et al.*, 2009) when modelling
16
17 144 reproductive performance.

18
19
20 145 However, modelling spatial variation in population parameters in the framework
21
22 146 of SDM is challenging. On one hand, collecting population data samples for model
23
24 147 calibration is much more time and effort consuming than collecting occurrence data.
25
26 148 Additionally, there are many technical choices to be made about the nature of the
27
28 149 dependent variable chosen to represent the population parameter of interest. For
29
30 150 example, in the case of birds, there are various ways to measure breeding success,
31
32 151 including: clutch size, hatching success, nesting success, breeding success per female or
33
34 152 group of females, general success of the breeding season and so on (Murray, 2000).
35
36 153 Indeed, the value of the dependent variable can take the form of counts, continuous
37
38 154 values and multiple or binary categorical values, which effectively increases model
39
40 155 structural uncertainty. While both ~~;~~ the influence of predictor choice and model
41
42 156 structural uncertainty ~~;~~ have received a great deal of attention in the SDM literature (e.g.
43
44 157 Thuiller, 2003; Wintle *et al.*, 2003; Pearson *et al.*, 2006), we ~~could have~~ find no studies
45
46 158 investigating how the definition of the dependent variable contributes to uncertainty in
47
48 159 modelling spatial variation in population parameters.

49
50
51 160 In this study, we aimed to compare the inference of a correlative species
52
53 161 distribution modelling approach based on species breeding occurrence data (i.e.,
54
55
56
57
58
59
60

1
2
3 162 presence-absence) against another correlative, but more mechanistic, approach based on
4
5 163 vital rates' data (i.e., breeding success) ~~in~~ in order to gain insights on species
6
7 164 persistence. As study case, we used an endangered species, the Great Bustard (*Otis*
8
9
10 165 *tarda*) across Spain. We explored how the definition of the population parameter
11
12 166 impacts on ecological inference and predictions about key areas for conservation. We
13
14 167 discussed the implications of the modelling options and what the predictions and
15
16 168 ecological inference tell us about the biology of the species, as well as the best strategies
17
18 169 for improving the realism and applicability of ~~modelling~~-species distribution modelling
19
20 170 approaches to support conservation decisions.
21
22
23
24

25 172 (A) METHODS

26
27 173

28 29 174 (B) The study model: Great Bustards in Spain

30
31
32 175 Great Bustards are large, lekking birds that live in highly fragmented populations in
33
34 176 cereal pseudo-steppes throughout the Palaearctic, from Morocco to eastern China
35
36 177 (Palacín & Alonso, 2008). Spain holds ca. 60-70% of the world population (Alonso &
37
38 178 Palacín, 2010). The species is globally threatened and classified as Vulnerable on the
39
40 179 Red List of Threatened Species (BirdLife International, 2015). Females nest on the
41
42 180 ground and rear their precocial chicks (usually one, sometimes two) alone over a period
43
44 181 of six to 12 months (Alonso *et al.*, 1998). Breeding success is highly variable, with
45
46 182 productivity values ranging between 0.04 and 0.53 chicks per female (Morales *et al.*,
47
48 183 2002; Martín *et al.*, 2007).
49
50

51
52 184

53 54 185 (B) Great Bustard breeding occurrence and breeding success databases

1
2
3 186 Great Bustard breeding occurrence was modelled using a database on presence-absence
4
5 187 records (PA database) compiled in Spring, when species detectability is maximal
6
7 188 (Alonso *et al.*, 2005), during the period 1987-2010 (Figure 1a). Presence data consisted
8
9 189 of all 350 lek centres known in Spain (Alonso *et al.*, 2012a). Absence data comprised an
10
11 190 equivalent sample of 350 points randomly distributed across Spain, excluding ~~both the~~
12
13 191 coastal border areas, mountainous ranges and ~~the~~ home ranges around lek centres
14
15
16 192 (estimated with a buffer of 2 km; Palacín *et al.*, 2012). These data can be considered as
17
18 193 true absences, since the distribution of the species at this time of the year is well known.

19 194 To estimate breeding success, we carried out population censuses across the
20
21 195 breeding range in Spain during the month of September (when the mortality peak of
22
23 196 juveniles is over and families can be detected due to their less elusive behaviour
24
25 197 compared to early summer; Martín *et al.*, 2007) of the same period 1987-2010 (Figure
26
27 198 1b, c). Breeding success was estimated annually as the young productivity (i.e., ratio of
28
29 199 the number of chicks to females) for each reproductive group (RG), which was made of
30
31 200 all flocks of females (FF) with or without chicks (~~FF~~) found in the same lek. Values
32
33 201 $\geq 100\%$ were discarded because they were associated with very small groups consisting
34
35 202 of one or two females and their chicks, or corresponded to RG with a very low
36
37 203 detectability (i.e., where number of females counted in September was lower than 30%
38
39 204 of those counted in Spring). It should be considered that productivity data were
40
41 205 heterogeneously distributed across space and time during the study period (i.e., data
42
43 206 were not available for all RG all years). Therefore, the raw values were averaged for the
44
45 207 whole series to achieve a spatially and temporally consistent coverage of this population
46
47 208 parameter, reflecting the long-term trends of species persistence better than single
48
49 209 measures made for shorter (or even isolated) periods of time. See Alonso *et al.* (2005)
50
51
52
53
54
55
56
57
58
59
60

1
2
3 210 and Álvarez-Martínez *et al.* (2015) for more details on breeding success surveys and
4
5 211 GIS database preparation. We accounted for a total of 208 RGs.

6
7 212 In order to build dependent variables informing on breeding success ~~for further~~
8
9 213 ~~modelling analyses and compare outcomes of methodological choices~~, three
10
11 214 methodological criteria were applied on the original multi-temporal dataset, generating
12
13 215 five datasets on breeding success (BS databases; Table 1). Each database was based on
14
15 216 a particular combination of subjective and data-driven choices about the treatment of
16
17 217 dependent variables: (i) Dependent variables could be continuous (productivity values
18
19 218 ranged from 0 to 100%) or binary (productivity ~~take~~ only had two values, which
20
21 219 represent high/low productivity, high/null productivity or positive/null productivity). (ii)
22
23 220 Continuous dependent variables were calculated by averaging annual productivity data
24
25 221 across the temporal series, using either the mean (database BS1) or the range of the
26
27 222 values (database BS2). (iii) Binary dependent variables were built by comparing: (iii.1)
28
29 223 RG with high productivity (locations where the productivity value was higher than the
30
31 224 averaged mean value for all RG across the whole study period; i.e., 0.15 young/female)
32
33 225 *vs.* RG with low productivity (locations collected across the occupied breeding area
34
35 226 where productivity was positive, but lower than the averaged mean value) (database
36
37 227 BS3); ~~);~~ (iii.2) RG with high productivity (a productivity value higher than the averaged
38
39 228 mean) *vs.* locations where productivity was null (i.e., a set of random points sampled
40
41 229 across the potential breeding distribution, ~~that have been formerly~~ estimated by Suárez-
42
43 230 Seoane *et al.*, 2002, avoiding a buffer of 2 km ~~-the species home range-~~ around each FF)
44
45 231 (database BS4). (iii.3) RG with positive productivity (value greater than “0”; i.e., birds
46
47 232 successfully bred, independently of the number of chicks raised) *vs.* locations where
48
49 233 productivity was null (database BS5). ~~The application of these three criteria generated~~
50
51 234 ~~five datasets on breeding success (BS databases), each based on a particular~~
52
53
54
55
56
57
58
59
60

1
2
3 235 ~~combination of subjective and data-driven choices about the treatment of dependent~~
4 ~~variables.~~ Databases generated in occupied breeding areas (BS1, BS2 and BS3) are
5 236
6
7 237 useful for modelling “breeding performance”, as they allow for comparing different (but
8
9 238 always positive) values of breeding success. ~~D;~~ ~~while~~ databases generated also in
10 239 potential breeding areas (BS4 and BS5) are useful for modelling “breeding site
11 240 selection”, as they allow for comparing locations where birds had successfully bred
12 241 from other potential sites where breeding success was ~~equal to 0~~ null.
13
14
15
16
17
18
19

20 243 **(B) Environmental variables**

21
22 244 On the basis of exploratory analyses, expert knowledge and published information
23 245 (Morales *et al.*, 2002; Alonso *et al.*, 2004; Pinto *et al.*, 2005; Martínez, 2008; Palacín *et*
24 246 *al.*, 2012), we selected a pool of 12 environmental GIS predictors potentially driving
25 247 both breeding occurrence and breeding success of Great Bustards in Spain (Table 2).
26 248 Predictors described topography, climate, primary production, landscape structure and
27 249 human disturbances. Scales ranged from 1:25000 to 1:200000 and pixel sizes from
28 250 ~~200m~~ 25m to 1km, depending on original data and methodological restrictions. All data
29 251 were ~~interpolated-rescaled using a natural neighbour method~~ to the same spatial
30 252 resolution, matching the pixel size of 1km among variables.
31
32
33
34
35
36
37
38
39
40
41
42

43 253 The role of topographic predictors on breeding occurrence and breeding success
44 254 was evaluated through the slope and its variation among agricultural plots. Topography
45 255 largely influence visual communication with conspecifics, as the breeding system of
46 256 dispersed leks involves strong visual cues over long distances (Alonso *et al.*, 2012a).
47 257 Data sources were, respectively, a digital elevation model at 25m (CNIG;
48 258 <http://www.cnig.es>) and the Spanish Geographic Information System for Agricultural
49 259 Plots (SIGPAC; [11](http://www.magrama.gob.es/en/agricultura/temas/sistema-de-</u>
50
51
52
53
54
55
56
57
58
59
60</p></div><div data-bbox=)

1
2
3 260 | [informacion-geografica-de-parcelas-agricolas-sigpac](#)) at 1:5000 scale. Data on climate
4
5 261 | and primary production were seasonally averaged (mean values and variation
6
7 262 | coefficients) to assess the effect of [these](#) environmental factors during critical periods of
8
9 263 | the year on breeding [occurrence and breeding](#) success. The effect of climate was
10
11 264 | included through the maximum Summer temperature (temperature during the last days
12
13 265 | of incubation and first days after hatching) and Autumn-Winter rainfall (precipitation
14
15 266 | prior to breeding season), both extracted from the Climatic Map of the Iberian Peninsula
16
17 267 | (Ninyerola *et al.*, 2005, 2007). The relevance of these climatic variables ~~has been~~
18
19 268 | ~~demonstrated for many species of birds and, specifically, for Great Bustards (Morales *et*~~
20
21 269 | ~~*al.*, 2002; Osborne *et al.*, 2007).~~ [for Great Bustards has been demonstrated previously](#)
22
23 270 | [by Morales *et al.* \(2002\) and Osborne *et al.* \(2007\).](#) According to these authors, winter
24
25 271 | precipitation controls the productivity of annual plants during the following spring and,
26
27 272 | therefore, the availability of arthropods during the period of chick maximum growth
28
29 273 | rate in early summer. ~~The~~ ~~is~~ ~~early-spring~~ development of herbaceous vegetation [in early](#)
30
31 274 | [spring](#) also contributes positively to the physiological condition of females, which
32
33 275 | directly affects the percentage of females attempting to breed, their clutch size and
34
35 276 | brood viability. The above-ground net primary production was represented through the
36
37 277 | Normalized Difference Vegetation Index of Spring (NDVI; Pettorelli *et al.*, 2007, 2011;
38
39 278 | ~~Bro-Jørgensen~~ ~~Bro-Jørgensen~~ *et al.* 2008; Hamel *et al.* 2009) obtained from a temporal
40
41 279 | series of NOAA-AVHRR satellite imagery (1987-2010), that was acquired from the
42
43 280 | SerGEO database from CCHS-CSIC
44
45 281 | (<http://humanidades.cchs.csic.es/cchs/sig/sergeo.html>). Landscape structure was
46
47 282 | quantified from the ~~Spanish Geographic Information System for Agricultural Plots~~
48
49 283 | ~~(SIGPAC; [geografica-de-parcelas-agricolas-sigpac](http://www.magrama.gob.es/en/agricultura/temas/sistema-de-informacion-
50
51 284 | <a href=))~~, that have been elaborated at 1:5000 scale for
52
53
54
55
56
57
58
59
60

1
2
3 285 the year 2012. We created a grid of 1-km over the whole Spain that intersected with the
4
5 286 SIGPAC layer in order to calculate both the maximum perimeter of plots and the
6
7 287 percentage of arable land within each 1-km grid (authors' unpublished data). Human
8
9 288 disturbances were evaluated through the Euclidean-distance from each pixel to the
10
11 289 nearest paved road or highway (Suárez-Seoane *et al.*, 2002) and the land protection
12
13 290 status. Data were obtained, respectively, from the Spanish Centre of Geographic
14
15 291 Information (CNIG; ~~http://www.cnig.es~~) at 1:200000 scale and the Nature 2000
16
17 292 Ecological Network (<http://www.magrama.gob.es>). All GIS analyses were done in
18
19 293 ArcGIS10.2 (ESRI, ~~2013~~2014).

20
21
22 294 Environmental features were gathered for each ~~of the six~~ databases on
23
24 295 occurrence and breeding success using a pixel-based approach. The spatial resolution of
25
26 296 the analyses, allowing for model inference across Spain, was 1km. In the case of BS
27
28 297 databases, points were spatially assigned to the location of the female flock with chicks
29
30 298 (isolated family or flock of females including at least one family) closest to the
31
32 299 “centroid” of all female flocks, with or without chicks, in a RG. Points were
33
34 300 recalculated for each year during the study period and averaged afterwards.

35
36
37 301 In order to avoid multi-collinearity problems that may lead to parameter bias
38
39 302 (Freckleton, 2011), we checked that Spearman's bivariate correlations among all
40
41 303 predictors were below 0.7 (Randin *et al.*, 2006), as well as ~~and that~~ variance inflation
42
43 304 factor (VIF) in further modelling analyses was lower than 4. There is no formal VIF
44
45 305 threshold, but a value of 10 is commonly used as an indicator of severe multicollinearity
46
47 306 (Neter *et al.*, 1990; Graham, 2003; Zuur *et al.*, 2010).

48
49
50 307

51 308 **(B) Model building**

52
53
54
55
56
57
58
59
60

1
2
3 309 Both breeding occurrence and breeding success of Great Bustards in Spain were
4
5 310 independently modelled on the basis of the environmental features described in Table 2
6
7 311 using generalized linear models (GLM). Multi-model inference, based on AIC values,
8
9 312 and model averaging (Burnham & Anderson, 2002; Burnham *et al.*, 2011) were
10
11 313 implemented for each of the six datasets (PA, BS1 to BS5). This method allows for
12
13 314 selecting the best subset of approximating models (i.e., those the smallest AIC value,
14
15 315 indicating the most parsimonious models) among all possible candidates. Models with
16
17 316 $\Delta_i (AIC_{best} - AIC_i) \leq 2$ were considered substantially supported by the data and similar to
18
19 317 the best model in their empirical reliability. Using this subset of models, we estimated
20
21 318 the averaged standardised coefficients (β) for each predictor, as well as its significance
22
23 319 and relative importance. Model coefficients were standardised to allow comparisons
24
25 320 among predictors. The relative importance of each predictor was measured as the sum
26
27 321 of the Akaike weights of all models in the subset where that predictor was present. The
28
29 322 value of the summed Akaike weight of each predictor ranges from 0 (if it appears only
30
31 323 in the most unlikely models) to 1 (if it appears in all the best models) (Burnham and
32
33 324 Anderson 2002; Symonds & Moussalli, 2011). GLMs were built using either a binomial
34
35 325 distribution with logit link, when the response variable was binomial, or a Gamma
36
37 326 distribution with log link, when the response variable was continuous. Model algorithms
38
39 327 were spatially projected across the study area using GIS tools, providing maps of habitat
40
41 328 suitability ranging from 0 to 1. Finally, variations in performance (deviance explained
42
43 329 by the best subset of variables against a null model), ~~type-family~~ of the most
44
45 330 contributing predictors and spatial predictions were evaluated across ~~the~~ model outputs.
46
47
48
49
50

51 To ~~evaluate-assess~~ the accuracy of model predictions, we built two validation
52
53 datasets consisting of used the next validation datasets: 67 cases showing with the best
54
55 333 reproductive performance across the study period (RG with an average annual
56
57
58
59
60

1
2
3 334 | productivity higher than the average value ~~of~~for the study period, with more than five
4
5 335 | years of available data) along with an equivalent sample of points randomly chosen
6
7 336 | within either: (i) occupied breeding areas (home ranges in Figure 1) or (ii) potential
8
9 337 | breeding areas (potential area defined by Suárez-Seoane *et al.*, 2002). We thus
10
11 338 | estimated, for each model output, the statistical significance of the differences in habitat
12
13 339 | suitability across occupied and potential breeding areas (the larger difference, the better
14
15 340 | is the discrimination capacity of a model).

16
17
18 341 | We used ~~this~~these validation datasets to check the Pearson bivariate correlations
19
20 342 | among prediction values achieved from PA and BS models in both occupied and
21
22 343 | potential breeding areas. Finally, we mapped the areas of congruence between the
23
24 344 | outputs achieved by the occurrence model (PA) and the breeding success models
25
26 345 | calibrated in the potential area (BS4 and BS5). We have not compared PA with BS1,
27
28 346 | BS2 and BS3 because the predictions made by these models are not applicable outside
29
30 347 | the occupied area.

31
32
33
34 348 | Analyses were done with the packages MASS and MuMIn from R 3.0.2
35
36 349 | statistical software (R Development Core Team, 2014) and ArcGIS10.2 (ESRI, 2014).

37
38
39 350 |

40 351 | **(A) RESULTS**

41
42
43 352 |

44
45 353 | Table 3 summarises the results of the modelling approach. Performance varied widely
46
47 354 | across model outcomes. In the case of the presence-absence (PA) model, the deviance
48
49 355 | explained by the best subset of variables against a null model was 64.98%, being Great
50
51 356 | Bustard occurrence significantly correlated with all families of predictors. In the case of
52
53 357 | breeding success (BS) models, deviance ranged from 7.83% to 62.27%. The largest
54
55 358 | differences were found between two groups of BS models. Models calibrated within
56
57
58
59
60

1
2
3 359 occupied breeding areas (BS1 to BS3) explained less deviance than those calibrated
4
5 360 within the potential distribution area (BS4 and BS5). The most relevant predictors of
6
7 361 breeding success also changed markedly between these two groups. In the former,
8
9 362 breeding success was significantly correlated with climate (temperature), primary
10
11 363 production and distance to IBAs. In the latter, breeding success was also related to
12
13 364 topography and landscape structure, instead of primary production.

14
15
16 365 The predictive performance of models fitted to mean (BS1) and range (temporal
17
18 366 variability) of breeding success (BS2) was similar. Nevertheless, mean breeding success
19
20 367 was driven by temperature, primary production and distance to IBAs, while breeding
21
22 368 success range was only significantly correlated with temperature. When modelling high
23
24 369 levels of breeding success (values above the mean for the period) in occupied (BS3)
25
26 370 *versus* potential (BS4) breeding areas, we found that the latter models had much higher
27
28 371 performance and comprised significant variables of different types, including
29
30 372 topography, temperature, landscape structure and human disturbances. In contrast, BS3
31
32 373 models only included primary production and human disturbances as significant
33
34 374 predictors. Models calibrated in potential breeding areas (BS4 and BS5) showed similar
35
36 375 predictive performance and were driven by the same families of predictors.

37
38
39
40
41 376 Spatial patterns of habitat suitability varied substantially across the pool of
42
43 377 models (Figure 2). The output of the PA model showed clear differences among suitable
44
45 378 and non-suitable areas at large scale. In the case of BS models, those calibrated in
46
47 379 occupied breeding areas (BS1 to BS3) presented homogeneous or even random patterns
48
49 380 of habitat suitability outside the reproductive areas, as predictions were only valid at
50
51 381 local scale (i.e., within home ranges; Figure 1b, c) ~~(i.e., within reproductive areas; see~~
52
53 382 ~~home range in Figure 1b, c)~~. Models calibrated in potential breeding areas (BS4 and
54
55 383 BS5) clearly identified differences between suitable and non-suitable areas at large
56
57
58
59
60

1
2
3 384 scale. Overall, the prediction values generated by BS4 and BS5 models were the most
4
5 385 correlated to the predictions of PA model, in both occupied and potential areas (Figures
6
7 386 3 and 4).

8
9 387 Model discrimination was lower in occupied than potential breeding areas. In the
10
11 388 former case, significant differences were only found for model BS3, while in the last
12
13 389 case, significant differences emerged for the PA model, as well as for all BS models
14
15 390 except for BS3 (Table 4).

16
17 391

18
19
20
21 392 **(A) DISCUSSION**

22
23 393

24
25 394 Many authors have highlighted the need to move beyond static correlative predictions of
26
27 395 species occurrence probability or relative likelihood of occurrence to model processes
28
29 396 that are more directly related to the long-term persistence of species (Guisan & Thuiller,
30
31 397 2005; ~~Araujo & Guisan, 2006; Heikkinen et al., 2006;~~ Thuiller et al., 2008; Zurrell et
32
33 398 al., 2009; Franklin, 2010; ~~Álvarez-Martínez et al., 2015~~). The anticipated benefits of
34
35 399 modelling processes closely linked to species persistence is that it will improve the
36
37 400 robustness of predictions about species future ranges in rapidly changing environments.
38
39 401 In this sense, the current study pioneers a comparison of the ecological inference arising
40
41 402 from a classical correlative species distribution modelling approach, based on presence-
42
43 403 absence data, with that of another approach in which correlative models are fitted to the
44
45 404 spatial variation in a population parameter, in this case, breeding success. Our results
46
47 405 provided insights into the specific strengths and weaknesses of ~~correlative models of~~
48
49 406 ~~occupancy~~SDM regarding their application in conservation biology.

50
51
52
53
54 407 Correlative ~~SDM models~~ based on occurrence data have been formerly
55
56 408 demonstrated as valuable conservation tools for designing management actions aimed to
57
58
59
60

1
2
3 409 promote patch occupancy (e.g. Suárez-Seoane *et al.*, 2002) and density of focal species
4
5 410 (García *et al.*, 2007). The application of such models implies several practical strengths,
6
7 411 as they have low input data needs, avoid the challenges of scaling up from individual to
8
9 412 landscape level and allow for evaluating niche tolerance limits at large scale (Peterson
10
11 413 *et al.*, 2016). However, as stated by Oliver *et al.* (2012), these models overlook
12
13 414 important features for long-term population persistence, such as population stability and
14
15 415 source-sink dynamics.

16
17
18 416 In this sense, ~~we found that~~ the incorporation of long temporal series of vital rates'
19
20 417 data in SDM allowed for achieving well performing models (deviance ranged from
21
22 418 7.83% to 62.27%; ~~being these values high in relation to other studies dealing with~~
23
24 ~~breeding success modelling; e.g. Donázar *et al.*, 2002 and Rodríguez & Bustamante,~~
25
26 419 2003 for comparison with other studies dealing with breeding success modelling), that
27
28 420 are useful for the identification of landscape attributes contributing to population
29
30 421 stability. We also detected that certain environmental relationships affecting breeding
31
32 422 success were overlooked in PA models. For example, temporal variability in Summer
33
34 423 maximum temperature was a significant driver of breeding success, while it did not
35
36 424 appear to strongly influence species occurrence. This suggests that climatic stability is a
37
38 425 demand much more evident for breeding than for surviving. A similar scenario was
39
40 426 found for the mean primary production, which was not significantly correlated to
41
42 427 species occurrence, but appeared as one of the most important explanatory predictors of
43
44 428 breeding success (BS1 and BS3 models). This indicates that primary production
45
46 429 influences where and when to breed, as it indirectly reflects the availability of food for
47
48 430 chicks, but may be less important for adult survival. These differences should be
49
50 431 explicitly considered when planning conservation measures that may fail if they are not
51
52 432 undertaken in areas that nurture long-term breeding success.
53
54
55
56
57
58
59
60

1
2
3 434 Indeed, the comparison between the habitat suitability patterns achieved by PA
4
5 435 and BS models allowed for the discrimination of critical areas for the long term
6
7 436 persistence of the species (i.e., areas potentially supporting high breeding performance
8
9 437 during the study period) from other areas that may be occupied periodically or
10
11 438 consistently in other life stages, but have a low suitability for reproduction and then, for
12
13 439 population maintenance. In this sense, the identification of marginal areas for species
14
15 440 persistence is a matter of conservation priority, due to the high risk of local extinction.
16
17 441 For example, the regions where BS models predicted much higher suitability than PA
18
19 442 models (highlighted in black in Figure 4) are those where most local extinctions have
20
21 443 occurred during recent decades (79% of 29 extinctions documented in 1960-1990 in the
22
23 444 whole of Spain have occurred in the northeastern and southwesternmost black patches),
24
25 445 due to hunting pressure and agricultural transformations (Alonso *et al.* 2003). The
26
27 446 message for conservation managers is that, with appropriate environmental measures,
28
29 447 the high potential breeding success in these areas would help restoring the original
30
31 448 populations. The larger geographic area predicted as suitable by the PA model, when
32
33 449 compared with BS models, could be associated to the fact that correlative models might
34
35 450 overestimate niche breadth by not constraining the niche to account for breeding needs
36
37 451 (Peterson *et al.*, 2016). Titeux *et al.* (2007) reached a similar conclusion when exploring
38
39 452 the role of incorporating fitness parameters (i.e., breeding success) in the definition of
40
41 453 niche boundaries for red-clacked shrike.
42
43
44
45
46

47 454 Modelling choices when building the dependent variable influenced the inference
48
49 455 and potential utility of the model outputs for environmental managers and decision-
50
51 456 makers. In this sense, Mostashari & Sussman (2005) propose a stakeholder-assisted
52
53 457 modelling process in which stakeholders participate through contributions of input and
54
55 458 feedback to the modelling process to improve the representation of focal systems.
56
57
58
59
60

1
2
3 459 Discussions should begin in the earliest stages of the ecological modelling process, as
4
5 460 they are essential for identifying the key choices for ~~the purpose of the~~ model building,
6
7 461 mitigating many of the subsequent problems that arise from inappropriate
8
9 462 methodological decisions (Martin *et al.*, 2012). For example, we showed here that
10
11 463 models fitted using continuous variants of breeding success within occupied breeding
12
13 464 areas (BS1 and BS2) explained a greater proportion of deviance than the binary model
14
15 465 calibrated across occupied area (BS3), being ~~were thus~~ more indicative of species
16
17 466 persistence ~~than the binary model calibrated in occupied area (BS3), with a greater~~
18
19 467 ~~proportion of explained deviance~~. In addition, while binary models of databases BS4
20
21 468 and BS5 explained much more deviance, we should consider that they were calibrated
22
23 469 in non-occupied areas and, therefore, they are actually more useful to evaluate breeding
24
25 470 sites' selection than species persistence. The ability of BS1, BS2 and BS3 models to
26
27 471 find even slight differences within currently occupied areas casts light on critical
28
29 472 parameters for species persistence linked to spatial and temporal environmental
30
31 473 restrictions and feeding resource availability (Alvarez-Martínez *et al.*, 2015).
32
33
34
35

36 474 A key modelling choice was whether the dependent variable should be continuous
37
38 475 (databases BS1 and BS2) or binary (BS3 to BS5). While the predictive performance of
39
40 476 BS1 (general pattern; mean value of breeding success across the temporal series) and
41
42 477 B2 (temporal variation; range value) models was similar, they were driven by a different
43
44 478 set of predictors. Areas of consistently good breeding performance over time, and
45
46 479 therefore of high species persistence, would be selected by Great Bustards' based on
47
48 480 many interacting environmental and social cues (Parejo *et al.*, 2006; Osborne *et al.*,
49
50 481 2007; Rieucan & Giraldeau, 2011). Temporal variation in breeding success was linked
51
52 482 to Summer maximum temperatures, which may be interpreted as the negative effect that
53
54
55
56
57
58
59
60

1
2
3 483 severe Summer droughts can have over breeding performance by limiting food
4
5 484 availability for the rearing chicks.
6

7 485 The choice of converting continuous values of breeding success into binary
8
9 486 dependent variables (BS3 to BS5) implies summarising the observed variability
10
11 487 according to unique thresholds, which involves a loss of information. This decision is
12
13 488 strongly linked to vagueness, a form of linguistic uncertainty which refers to the
14
15 489 possibility of borderline cases which neither satisfy a criterion, nor its negation, when
16
17 490 performing categorical classifications of data (Regan *et al.*, 2002). In such a situation,
18
19 491 the development of fuzzy logic sets and rules (Ascough II *et al.*, 2008; Álvarez-
20
21 492 Martínez *et al.*, 2010) is recommended to reduce the level of vagueness in decision-
22
23 493 making. If this is not possible, a deliberate consideration of the meaning and
24
25 494 implications of the chosen thresholds is advised.
26
27
28

29 495 The most influential methodological choice in terms of model performance,
30
31 496 selected predictors and spatial pattern of habitat suitability across the pool of results was
32
33 497 the area of calibration for binary BS variables. In fact, Jiménez-Valverde *et al.* (2009)
34
35 498 and Lobo *et al.* (2010) already highlighted the key importance of carefully choosing the
36
37 499 area of calibration depending on the question at hand. Models calibrated with binary
38
39 500 variables collected in occupied breeding areas (BS3) explained much less deviance than
40
41 501 models calibrated across non-occupied sites within the potential distribution area; BS4
42
43 502 and BS5). This is a common and expected result, since describing local differences
44
45 503 within suitable areas using models narrowly calibrated across space and environment
46
47 504 (i.e., comparing between areas of higher-than-average and those of lower-than-average
48
49 505 breeding success) is more demanding than assessing coarse differences between areas of
50
51 506 positive breeding performance and potential areas including many unsuitable pockets.
52
53 507 The lower deviance explained by BS3 model could be associated to the fact that short-
54
55
56
57
58
59
60

1
2
3 508 | term environmental changes ~~which~~that may affect breeding success in suitable breeding
4
5 509 | areas cannot be anticipated by birds at local scale. For example, the temporal variation
6
7 510 | in Summer maximum temperature was significantly correlated with breeding success in
8
9 511 | BS4 and BS5, but not in BS3 models. This fact suggest that Great Bustards might
10
11 512 | “predict” general patterns of climatic stability across their range and use it as an
12
13 513 | indicator of quality or suitability when looking for breeding areas through “public
14
15 514 | information” that can be gained from the presence of conspecifics (Osborne *et al.*,
16
17 515 | 2007). However, birds cannot identify local differences in climatic stability within
18
19 516 | reproductive areas. On the top of this, it should be considered that certain differences in
20
21 517 | the variables selected by these models might be, at least partially, related to the extent of
22
23 518 | the area where models were calibrated (Van Der Wal *et al.*, 2009). Nevertheless, model
24
25 519 | calibration across different scenarios (implying different extents) was necessary to test
26
27 520 | our hypothesis.

28
29
30
31 521 | In this study, we have combined spatially explicit data from both species
32
33 522 | occurrence and vital rates into a SDM framework to identify priority conservation areas
34
35 523 | and landscape structures supporting population maintenance. The utility of this
36
37 524 | approach would be greatly enhanced through the integration of temporal fluctuations in
38
39 525 | other vital rates, such as survival and dispersal, using metapopulation models. Such an
40
41 526 | approach would allow the identification of the threats most likely ~~to~~influence
42
43 527 | population persistence over time horizons relevant to management decision-making.
44
45 528 | However, the extra data and expertise demands of such an approach preclude its
46
47 529 | application in many instances (Franklin, 2010). We argue that our approach may
48
49 530 | represent a suitable compromise between mechanistic models and the simpler
50
51 531 | correlative SDM based on occurrence data~~on occupancy~~.

52
53
54
55
56
57
58
59
60

1
2
3 533 | **(A) ACKNOWLEDGMENTS**
4
5
6

7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

534

535 | The Spanish Ministry for Science and Innovation (project CGL2008-02567) ~~funded~~
536 | supported this research, although many censuses were funded by earlier projects from
537 | that Ministry, as well as from ICONA, Junta de Castilla y León, Junta de Andalucía and
538 | Gobierno de Navarra awarded to JCA since 1987. We thank M. Alcántara, J.A. Arranz,
539 | A. Balmori, B. Campos, F.J. Carmona, J. Ezquerro, M.J. García-Baquero, M. Guerrero,
540 | E. Izquierdo, J. Larumbe, A.I. Lasheras, J.E. Montero, A. Onrubia, J. Panadero, A.
541 | Sánchez, C.A. Sánchez, C. Torralbo, A Torrijo and R. Ubaldo for providing additional
542 | census data, as well as other anonymous persons who carried out censuses in several
543 | areas. J.A. Alonso, C. Bravo, M. Magaña, B. Martín, C.A. Martín, E. Martín, M.
544 | Morales, C. Ponce and A. Torres collaborated during many regional surveys. We are
545 | also grateful to I. Prieto (University of León) and L. Mateos (ESRI) for their help with
546 | environmental data handling. A. Cabria and J.M. Rodríguez (CNIG) provided the 5m-
547 | DEM and GIS databases. I. del Bosque from CCHS-CSIC provided remote sensing data
548 | and other GIS variables (SerGEO database collection). Thanks to P.E. Osborne
549 | (University of Southampton) for his useful comments. J.M. Álvarez-Martínez was
550 | supported by a contract from project CGL2008-02567 of the Spanish Ministry for
551 | Science and Innovation and B.E Wintle by an Australian Research Council Future
552 | Fellowship (FT100100819). This work was partially carried out at the Mixed Unit of
553 | Research on Biodiversity (UMIB, University of Oviedo, Mieres, Spain). Finally, we
554 | also thank to R. Heikkinen, Adam B. Smith and an anonymous referee for their useful
555 | comments that helped to improve the quality of the manuscript.
556 |

557 | **BIOSKETCH**

1
2
3 558 The main research topic of **Susana Suárez-Seoane** is the effect of land use change on
4
5 559 landscape dynamics and functioning. She is particularly focused on the analysis of
6
7 560 vertebrate species distribution patterns using remote sensing and species distribution
8
9 561 modelling techniques. The work presented in this manuscript is framed within the
10
11 562 project CGL2008-02567 entitled ‘Human impacts on Great Bustard population
12
13 563 dynamics’, funded by the Spanish Ministry for Science and Innovation and led by Prof.
14
15 564 Juan Carlos Alonso (www.proyectoavutarda.org). The overall aim of this project was to
16
17 565 assess human impacts on the viability of Great Bustards in the Iberian Peninsula, the
18
19 566 last stronghold of this globally endangered species.
20
21
22
23

24
25 568 Author contributions: S.S.S., J.M.A.M., B.W. and J.C.A. originally formulated the ideas
26
27 569 presented in this paper. J.C.A. and C.P. obtained the species data. J.M.A.M. was in
28
29 570 charge of data analysis. S.S.S., J.M.A.M. and B.W. wrote the first draft of this
30
31 571 manuscript and all authors contributed extensively to the preparation of the final
32
33 572 version.
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

573 **REFERENCES**

- 574 Alonso, J.C. & Palacín, C. (2010) The world status and population trends of the Great
575 Bustard: 2010 update. *Chinese Birds*, **1**, 141-147.
- 576 Alonso, J.C., Martín, E., Alonso, J.A. & Morales, M.B. (1998) Proximate and ultimate
577 causes of natal dispersal in the Great Bustard *Otis tarda*. *Behavioral Ecology*, **9**,
578 243-252.
- 579 Alonso J.C., Palacín, C. & Martín, C.A. (2003) Status and recent trends of the Great
580 Bustard (*Otis tarda*) population in the Iberian Peninsula. *Biological Conservation*,
581 **110**, 185-195.
- 582 Alonso, J.C., Martín, C.A., Alonso, J.A., Palacín, C., Magaña, M. & Lane, S.J. (2004)
583 Distribution dynamics of a Great Bustard metapopulation throughout a decade:
584 influence of conspecific attraction and recruitment. *Biodiversity and Conservation*,
585 **13**, 1659-1674.
- 586 Alonso, J.C., Palacín, C. & Martín, C.A. (2005) La avutarda común en la península
587 Ibérica: población actual y método de censo. SEO/BirdLife, Madrid.
- 588 Alonso, J.C., Álvarez-Martínez, J.M. & Palacín, C. (2012a) Leks in ground-displaying
589 birds: hotspots or safe places? *Behavioral Ecology*, **23**, 491-501.
- 590 Alonso, J.C., Magaña, M. & Álvarez-Martínez, J.M. (2012b) Male display areas in
591 exploded leks: the importance of food resources for male mating success.
592 *Behavioral Ecology*, **23**, 1296-1307.
- 593 Álvarez-Martínez, J.M., Stoorvogel, J.J., Suárez-Seoane, S. & de Luis, E. (2010)
594 Uncertainty analysis as a tool for refining land dynamics modelling on changing
595 landscapes: a case study in a Spanish Natural Park. *Landscape Ecology*, **25**, 1385-
596 1404.

- 1
2
3 597 Álvarez-Martínez, J.M., Suárez-Seoane, S., Palacín, C., Sanz, J. & Alonso, J.C. (2015)
4
5 598 Can Eltonian processes explain species distributions at large scale? A case study
6
7 599 with Great Bustard (*Otis tarda*). *Diversity and Distributions*, **21**, 123-138.
8
9 600 Araújo, M.B., Thuiller, W., Williams, P. H. & Reginster, I. (2005) Downscaling
10
11 601 European species atlas distributions to a finer resolution: implications for
12
13 602 conservation planning. *Global Ecology and Biogeography*, **14**, 17-30.
14
15
16 603 Ascough II, J.C., Maier, H.R., Ravalico J.K & Strudley, M.W. (2008) Future research
17
18 604 challenges for incorporation of uncertainty in environmental and ecological
19
20 605 decision-making. *Ecological Modelling*, **219**, 383-399.
21
22
23 606 Austin, M.P. & Van Niel, K.P. (2011) Improving species distribution models for climate
24
25 607 change studies: variable selection and scale. *Journal of Biogeography*, **38**, 1-8.
26
27
28 608 BirdLife International (2015) *Otis tarda*. The IUCN Red List of Threatened Species
29
30 609 2015: e.T22691900A85068116. Office for Official Publications of the European
31
32 610 Communities, Luxembourg. Downloaded on 20 July 2016.
33
34 611 [Brambilla, M. & Ficetola, G.F. \(2012\) Species distribution models as a tool to estimate](#)
35
36 612 [reproductive parameters: a case study with a passerine bird species. *Journal of*](#)
37
38 613 [Animal Ecology](#), **81**, 781-787.
39
40
41 614 Bro-Jørgensen, J., Brown, M.E. & Pettorelli, N. (2008) Using the satellite-derived
42
43 615 normalized difference vegetation index (NDVI) to explain ranging patterns in a lek-
44
45 616 breeding antelope: the importance of scale. *Oecologia*, **158**, 177-182.
46
47
48 617 Brotons, L., Herrando, S. & Pla, M. (2007) Updating bird species distribution at large
49
50 618 spatial scales: applications of habitat modelling to data from long-term monitoring
51
52 619 programs. *Diversity and Distributions*, **13**, 276-288.
53
54
55 620 Burnham, K.P. & Anderson, D.R. (2002) Model selection and multi-model inference. A
56
57 621 practical information-theoretic approach. New York: Springer.
58
59
60

1
2
3 622 Burnham, K.P., Anderson, D.R. & Huyvaert, K.P. (2011) AIC model selection and
4
5 623 multimodel inference in behavioral ecology: some background, observations, and
6
7 624 comparisons. *Behavioral Ecology and Sociobiology*, **65**, 23-35.

8
9 625 ~~Bykova, O., Chuine, I., Morin, X. & Higgins, S.I. (2012) Temperature dependence of~~
10
11 626 ~~the reproduction niche and its relevance for plant species distributions. *Journal of*~~
12
13 627 ~~*Biogeography*, **39**, 2191-2200.~~

14
15 628 ~~Chalfoun, A.D. & Schmidt, K.A. (2012) Adaptive breeding-habitat selection: Is it for~~
16
17 629 ~~the birds? *The Auk*, **129**, 589-599.~~

18
19 630 ~~Dobrowski, S.Z., Thorne, J.H., Greenberg, J.A., Safford, H.D., Mynsberge, A.R.,~~
20
21 631 ~~Crimmins, S.M. & Swanson, A.K. (2011) Modeling plant ranges over 75 years of~~
22
23 632 ~~climate change in California, USA: temporal transferability and species traits.~~
24
25 633 ~~*Ecological Monographs*, **81**, 241-257.~~

26
27 634 Donázar, J.A., Blanco, G., Hiraldo, F., Soto-Largo, E. & Oria, J. (2002) Effects of
28
29 635 forestry and other land use practices on the conservation of cinereous vultures.
30
31 636 *Ecological Applications*, **12**, 1445-1456.

32
33 637 ESRI (2014) ArcInfo desktop GIS 10.2. <http://www.esri.com>

34
35 638 Falcucci, A., Ciucci, P., Maiorano, L., Gentile, L. & Boitani, L. (2009) Assessing
36
37 639 habitat quality for conservation using an integrated occurrence-mortality model.
38
39 640 *Journal of Applied Ecology*, **46**, 600-609.

40
41 641 Franklin, J. (2010) Moving beyond static species distribution models in support of
42
43 642 conservation biogeography. *Diversity and Distributions*, **16**, 321-330.

44
45 643 ~~Franklin, J., Regan, H.M. & Syphard, A.D. (2014) Linking spatially explicit species~~
46
47 644 ~~distribution and population models to plan for the persistence of plant species under~~
48
49 645 ~~global change. *Environmental Conservation*, **41**, 97-109.~~

- 1
2
3 646 Freckleton, R. (2011) Dealing with collinearity in behavioural and ecological data:
4
5 647 model averaging and the problems of measurement error. *Behavioral Ecology and*
6
7 648 *Sociobiology*, **65**, 91-101.
- 9
10 649 ~~Frederiksen, M., Lebreton, J. D., Pradel, R., Choquet, R. & Gimenez, O. (2014)~~
11
12 650 ~~Review: Identifying links between vital rates and environment: a toolbox for the~~
13
14 651 ~~applied ecologist. *Journal of Applied Ecology*, **51**, 71-81.~~
- 15
16 652 García, J., Suárez-Seoane, S., Miguélez, D., Osborne, P.E. & Zumalacárregui, C. (2007)
17
18 653 Spatial analysis of habitat quality in a fragmented population of little bustard (*Tetrax*
19
20 654 *tetrax*): implications for conservation. *Biological Conservation*, **137**, 45-56.
- 21
22 655 ~~Goodechild, M.F. (1994) Integrating GIS and remote sensing for vegetation analysis and~~
23
24 656 ~~modeling: methodological issues. *Journal of Vegetation Science*, **5**, 615-626.~~
- 25
26 657 Graham, M.H. (2003) Confronting multicollinearity in ecological multiple regression.
27
28 658 *Ecology*, **84**, 2809-2815.
- 29
30 659 Guisan, A. & Thuiller, W. (2005) Predicting species distribution: offering more than
31
32 660 simple habitat models. *Ecology Letters*, **8**, 993-1009.
- 33
34 661 Guisan, A. & Zimmermann, N.E. (2000) Predictive habitat distribution models in
35
36 662 ecology. *Ecological Modelling*, **135**, 147-186.
- 37
38 663 Guisan, A. *et al.* (2013) Predicting species distributions for conservation decisions.
39
40 664 *Ecology Letters*, **16**, 1424-1435.
- 41
42 665 Guisan, A., Petitpierre, B., Broennimann, O., Daehler, C. & Kueffer, C. (2014)
43
44 666 Unifying niche shift studies: insights from biological invasions. *Trends in Ecology*
45
46 667 *& Evolution*, **29**, 260-269.
- 47
48 668 Hamel, S., Garel, M., Festa-Bianchet M., Gaillard, J.M. & Côté, S.D. (2009) Spring
49
50 669 Normalized Difference Vegetation Index (NDVI) predicts annual variation in timing
51
52
53
54
55
56
57
58
59
60

- 1
2
3 670 of peak faecal crude protein in mountain ungulates. *Journal of Applied Ecology*, **46**,
4
5 671 582-589.
6
7 672 ~~Heikkinen, R.K., Luoto, M. & Virkkala, R. (2006) Does seasonal fine-tuning of climatic~~
8
9 ~~variables improve the performance of bioclimatic envelope models for migratory~~
10 673
11 ~~birds? *Diversity and Distributions*, **12**, 502-510.~~
12 674
13
14 675 Howard, C., Stephens, P.A., Pearce-Higgins, J.W., Gregory, R.D. & Willis, S.G. (2014)
15
16 676 Improving species distribution models: the value of data on abundance. *Methods in*
17
18 677 *Ecology and Evolution*, **5**, 506-513.
19
20 678 Jiménez-Valverde, A., Peterson, A.T., Soberón, J., Overton, J.M., Aragón, P. & Lobo,
21
22 J.M. (2011) Use of niche models in invasive species risk assessments. *Biological*
23 679
24 *Invasions*, **13**, 2785-2797.
25
26 680
27 681 Kearney, M. & Porter, W. (2009) Mechanistic niche modelling: combining
28
29 682 physiological and spatial data to predict species' ranges. *Ecology Letters*, **12**, 334-
30
31 683 350.
32
33 684 Kearney, M., Simpson, S.J., Raubenheimer, D. & Helmuth, B. (2010) Modelling the
35
36 685 ecological niche from functional traits. *Philosophical Transactions of the Royal*
37
38 686 *Society - Series B*, **365**, 3469-3483.
39
40 687 ~~Leseroë, A., Dugger, K.M., Ballard, G. & Ainley, D.G. (2009) Effects of individual~~
41
42 ~~quality, reproductive success and environmental variability on survival of a long-~~
43 688
44 ~~lived seabird. *Journal of Animal Ecology*, **78**, 798-806.~~
45 689
46
47 690 Lobo, J.M., Jiménez-Valverde, A. & Hortal, J. (2010) The uncertain nature of absences
48
49 691 and their importance in species distribution modelling. *Ecography*, **33**, 103-14.
50
51 692 Martín, C.A., Alonso, J.C., Alonso, J.A., Palacín, C., Magaña, M. & Martín, B. (2007)
52
53 693 Sex-biased juvenile survival in a bird with extreme size dimorphism, the Great
54
55 694 Bustard. *Journal of Avian Biology*, **38**, 335-346.
56
57
58
59
60

- 1
2
3 695 ~~Martin, T.G., Burgman, M.A., Fidler, F., Kuhnert, P.M., Low-Choy, S., McBride, M. &~~
4
5 696 ~~Mengersen, K. (2012) Eliciting expert knowledge in conservation science.~~
6
7 697 ~~*Conservation Biology*, **26**, 29-38.~~
8
9
10 698 Martínez, C. (2008) Distribution, density and productivity of great bustards in
11
12 699 northwestern Spain: a regional approach. *Journal of Ornithology*, **149**, 507-14.
13
14 700 Michel, M.J., Chien, H., Beachum, C.E., Bennett, M.G. & Knouft, J.H. (2017) Climate
15
16 701 change, hydrology, and fish morphology: predictions using phenotype-environment
17
18 702 associations. *Climatic Change*, **140**, 563-576.
19
20
21 703 ~~Moilanen, A., Franco, A.M., Early, R.I., Fox, R., Wintle, B. & Thomas, C.D. (2005)~~
22
23 704 ~~Prioritizing multiple-use landscapes for conservation: methods for large multi-~~
24
25 705 ~~species planning problems. *Proceedings of the Royal Society of London B-*~~
26
27 706 ~~*Biological Sciences*, **272**, 1885-1891.~~
28
29
30 707 Morales, M.B., Alonso, J.C. & Alonso, J.A. (2002) Annual productivity and individual
31
32 708 female reproductive success in a Great Bustard population. *Ibis*, **144**, 293-300.
33
34 709 Mostashari, A. & Sussman, J. (2005) Stakeholder-assisted modelling and policy design
35
36 710 process for environmental decision-making. *Journal of Environmental Assessment*
37
38 711 *Policy and Management*, **7**, 355-386.
39
40
41 712 Moudrý, V., Komárek, J. & Šimová, P. (2017) Which breeding bird categories should
42
43 713 we use in models of species distribution? *Ecological Indicators*, **74**, 526-529.
44
45 714 Murray Jr, B.G. (2000) Measuring annual reproductive success in birds. *The Condor*,
46
47 715 **102**, 470-473.
48
49
50 716 Neter, J., Wasserman, W. & Kutner, M.H. (1990) Applied linear statistical models:
51
52 717 regression, analysis of variance, and experimental designs. Irwin, Homewood, IL.
53
54
55
56
57
58
59
60

- 1
2
3 718 Ninyerola, M., Pons, X. & Roure, J.M. (2005) Atlas Climático Digital de la
4
5 719 Península Ibérica. Metodología y aplicaciones en bioclimatología y geobotánica.
6
7 720 ISBN 932560-8-7. Universidad Autónoma de Barcelona, Bellaterra.
8
9
10 721 Ninyerola, M., Pons, X. & Roure, J.M. (2007) Monthly precipitation mapping of the
11
12 722 Iberian Peninsula using spatial interpolation tools implemented in a Geographic
13
14 723 Information System. *Theoretical and Applied Climatology*, **89**, 195-209.
15
16 724 Oliver, T. H., Gillings, S., Girardello, M., Rapacciuolo, G., Brereton, T.M.,
17
18 725 Siriwardena, G.M., Roy, David B., Pywell, R. & Fuller, R.J. (2012) Population
19
20 726 density but not stability can be predicted from species distribution models. *Journal*
21
22 727 *of Applied Ecology*, **49**, 581-590.
23
24
25 728 Osborne, P.E., Suárez-Seoane S. & Alonso, J.C. (2007) Behavioural mechanisms that
26
27 729 undermine species envelope models: the causes of patchiness in the distribution of
28
29 730 Great Bustards *Otis tarda* L. in Spain. *Ecography*, **6**, 819-829.
30
31
32 731 Palacín, C. & Alonso, J.C. (2008) An updated estimate of the world status and
33
34 732 population trends of the Great Bustard *Otis tarda*. *Ardeola*, **55**, 13-25.
35
36 733 Palacín, C., Alonso, J.C., Martín, C.A. & Alonso, J.A. (2012) The importance of
37
38 734 traditional farmland areas for steppe birds: a case study of migrant female Great
39
40 735 Bustards *Otis tarda* in Spain. *Ibis*, **154**, 85-95.
41
42
43 736 Parejo, D., Oro, D. & Danchin, E. (2006) Testing habitat copying in breeding habitat
44
45 737 selection in a species adapted to variable environments. *Ibis*, **148**, 146-154.
46
47 738 Pearson, R.G., Thuiller, W., Araújo, M.B., Martínez-Meyer, E., Brotons, L., McClean,
48
49 739 C., Miles, L., Segurado, P., Dawson, T.P. & Lees, D.C. (2006) Model-based
50
51 740 uncertainty in species range prediction. *Journal of Biogeography*, **33**, 1704-1711.
52
53
54 741 Peterson, A.T., Papeş, M. & Soberón, J. (2016) Mechanistic and Correlative Models of
55
56 742 Ecological Niches. *European Journal of Ecology*, **1**, 28-38.
57
58
59
60

- 1
2
3 743 Pettorelli, N., Pelletier, F., Hardenberg, A.V., Festa-Bianchet, M. & Côté, S.D. (2007)
4
5 744 Early onset of vegetation growth vs. rapid green-up: impacts on juvenile mountain
6
7 745 ungulates. *Ecology*, **88**, 381-390.
8
9 746 Pettorelli, N., Ryan, S., Mueller, T., Bunnefeld, N., Jedrzejewska, B., Lima, M. &
10
11 747 Kausrud, K. (2011) The Normalized Difference Vegetation Index (NDVI):
12
13 748 unforeseen successes in animal ecology. *Climate Research*, **46**, 15-27.
14
15
16 749 Pinto, M., Rocha, P. & Moreira, F. (2005) Long-term trends in Great Bustard
17
18 750 populations in Portugal suggest concentration in single high quality area. *Biological*
19
20 751 *Conservation*, **124**, 415-23.
21
22
23 752 R Development Core Team. (2011) R: A language and environment for statistical
24
25 753 computing. R Foundation for Statistical Computing, Vienna (Austria).
26
27 754 <http://www.R-project.org>
28
29
30 755 Randin, C.F., Dirnböck, T., Dullinger, S., Zimmermann, N.E., Zappa, M. & Guisan, A.
31
32 756 (2006) Are niche-based species distribution models transferable in space? *Journal of*
33
34 757 *Biogeography*, **33**, 1689-1703.
35
36 758 Regan, H.M., Colyvan, M. & Burgman, M.A. (2002) A taxonomy and treatment of
37
38 759 uncertainty for ecology and conservation biology. *Ecological Applications*, **12**, 618-
39
40 760 628.
41
42
43 761 Rieucou, G. & Giraldeau, L.A. (2011) Exploring the costs and benefits of social
44
45 762 information use: an appraisal of current experimental evidence. *Phil. Trans. R. Soc.*
46
47 763 *B*, **366**, 949-957.
48
49 764 Rodríguez, C. & Bustamante, J. (2003) The effect of weather on lesser kestrel breeding
50
51 765 success: can climate change explain historical population declines? *Journal of*
52
53 766 *Animal Ecology*, **72**, 793-810.
54
55
56
57
58
59
60

- 1
2
3 767 ~~Ryder, T.B., Parker, P.G., Blake, J.G. & Loiselle, B.A. (2009) It takes two to tango:~~
4 ~~reproductive skew and social correlates of male mating success in a lek breeding~~
5 768 ~~bird. *Proc. R. Soc. B*, **276**, 2377-2384.~~
6
7
8
9
10 770 Soga M. & Koike S. (2013) Large forest patches promote breeding success of a
11 terrestrial mammal in urban landscapes. *PLoS ONE*, **8**, e51802.
12
13 771
14 772 Strebel, N., Kéry, M., Schaub, M. & Schmid, H. (2014) Studying phenology by flexible
15 modelling of seasonal detectability peaks. *Methods in Ecology and Evolution*, **5**,
16 773 483-490.
17
18
19
20 774
21 775 Symonds, M. R. & Moussalli, A. (2011) A brief guide to model selection, multimodel
22 inference and model averaging in behavioural ecology using Akaike's information
23 776 criterion. *Behavioral Ecology and Sociobiology*, **65**, 13-21.
24
25 777
26
27 778 Suárez-Seoane, S., Osborne, P.E. & J.C. Alonso (2002) Large-scale habitat selection by
28 agricultural steppe birds in Spain: identifying species-habitat responses using
29 779 generalized additive models. *Journal of Applied Ecology*, **39**, 755-771.
30
31 780
32
33 781 ~~Suárez Seoane, S., García de la Morena, E., Morales, M. B., Osborne, P.E. & de Juana,~~
34 ~~E. (2008) Maximum entropy niche based modelling of seasonal changes in little~~
35 782 ~~bustard (*Tetrax tetrax*) distribution. *Ecological Modelling*, **219**, 17-29.~~
36
37
38 783
39
40 784 ~~Taylor, R.L., Walker, B.L., Naugle, D.E. & Mills, L.S. (2012) Managing multiple vital~~
41 ~~rates to maximize greater sage grouse population growth. *The Journal of Wildlife*~~
42 785 ~~*Management*, **76**, 336-347.~~
43
44
45 786
46
47 787 Thuiller, W. (2003) BIOMOD-optimizing predictions of species distributions and
48 projecting potential future shifts under global change. *Global Change Biology*, **9**,
49 788 1353-62.
50
51 789
52
53 790 Thuiller, W., Albert, C., Araújo, M. B., Berry, P. M., Cabeza, M., Guisan, A., Hickler,
54
55 791 T., Midgley, G.F., Paterson, J., Schurr, F.M., Sykes, M.T. & Zimmermann, N.E.

- 1
2
3 792 (2008) Predicting global change impacts on plant species' distributions: future
4
5 793 challenges. *Perspectives in Plant Ecology, Evolution and Systematics*, **9**, 137-152.
6
7 794 Titeux, N., Dufrene, M., Radoux, J., Hirzel, A. H. & Defourny, P. 2007. Fitness-related
8
9 795 parameters improve presence-only distribution modelling for conservation practice:
10
11 796 the case of the red-backed shrike. *Biological Conservation*, **138**, 207-223.
12
13 797 Van Der Wal, J., Shoo, L.P., Graham, C. & Williams, S.E. 2009. Selecting pseudo-
14
15 798 absence data for presence-only distribution modeling: How far should you stray
16
17 799 from what you know? *Ecological Modeling*, **220**, 589-594.
18
19 800 Van Horne, B. (1983) Density as a misleading indicator of habitat quality. *The Journal*
20
21 801 *of Wildlife Management*, **47**, 893-901.
22
23 802 Virkkala, R., Pöyry, J., Heikkinen, R. K., Lehikoinen, A., & Valkama, J. (2014)
24
25 803 Protected areas alleviate climate change effects on northern bird species of
26
27 804 conservation concern. *Ecology and Evolution*, **4**, 2991-3003.
28
29 805 Wintle, B.A., McCarthy, M.A., Volinsky, C.T. & Kavanagh, R.P. (2003) The use of
30
31 806 Bayesian Model Averaging to better represent the uncertainty in ecological models.
32
33 807 *Conservation Biology*, **17**, 1579-1590.
34
35 808 Zuur, A.F., Ieno, E.N. & Elphick, C.S. (2010) A protocol for data exploration to avoid
36
37 809 common statistical problems. *Methods in Ecology and Evolution*, **1**, 3-14.
38
39 810 ~~Zurell, D., Jeltsch, F., Dormann, C. F. & Schröder, B. (2009) Static species distribution~~
40
41 ~~models in dynamically changing systems: how good can predictions really be?~~
42
43 ~~*Ecography*, **32**, 733-744.~~
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1. Methodological criteria used to define dependent variables informing on the population parameter to be modelled. Each column represents a different database on breeding success (BS) elaborated from the original dataset by applying different decisions. For example, in the case of database BS1, the dependent variable consisted of continuous values and the statistical parameter used to average the temporal series of productivity data was the mean. The table also includes the sample size of each dataset.

Methodological criteria	BS1	BS2	BS3	BS4	BS5
(i) General type of dependent variable	Continuous		Binary		
(ii) Continuous dependent variable	Mean (<i>n</i> =208)	Range (<i>n</i> =208)			
(iii) Binary dependent variable:			High productivity (> averaged mean for the whole period; 0.15 young/female) (<i>n</i> =98)	Positive productivity (>0) (<i>n</i> =208)	
-High or positive productivity			Occupied breeding sites Low productivity (0< productivity < averaged mean for the period) (<i>n</i> =110)	Potential breeding sites Null productivity (=0) (<i>n</i> = 110)	
-Low or null productivity				(<i>n</i> =208)	

Table 2. Environmental predictors used to model both breeding occurrence and breeding success of Great Bustard. Spring (SP) includes March, April and May; Summer (SU): June, July and August; and Autumn-Winter (WI): September to February. CV is the coefficient of variation = $[(SD/\bar{x}) * 100]$, where SD is the standard deviation.

Family	Variable	Code	Units	Source
Topography	Slope of the terrain	SLO	degrees	Digital elevation model (DEM) from the Spanish Centre of Geographic Information (CNIG) at 25m of spatial resolution
	Slope variation among agricultural plots	SLOSPcv	degrees	GIS database of agricultural plots (SIGPAC) at 1:5000 scale, year 2012
Climate	Maximum temperature of Summer (mean)	TMAmSU	°C	Iberian Climatic Map at 200m of spatial resolution, period 1950-1999 (Ninyerola <i>et al.</i> 2005, 2007)
	Maximum temperature of Summer (CV)	TMAcvSU	°C	
	Rainfall of Autumn-Winter (mean)	PPsmWI	mm	
	Rainfall of Autumn-Winter (CV)	PPcvWI	mm	
Primary production	NDVI of Spring (mean)	NDVI _m SP	dimensionless (-1, 1)	NOAA-AVHRR at 1km of spatial resolution, period 1987-2010 (SerGEO database from CCHS-CSIC)
	NDVI of Spring (CV)	NDVI _c vSP	dimensionless (-1, 1)	
Landscape structure	Maximum perimeter of plots	PERIM _{max}	m	GIS database of agricultural plots (SIGPAC) at 1:5000 scale, year 2012
	Arable land (% occupation)	ARLAND	%	
Human disturbances	Distance to paved roads and highways	DISTRO	m	Spanish Centre of Geographic Information (CNIG) at 1:200000 scale
	Distance to IBAs (Important Bird Areas)	DISTIBA	m	

Table 3. Results of multi-model averaging for presence-absence (PA) and breeding success (BS1 to BS5) models (see definitions of variables in Table 2). Each cell shows the sign, the full model-averaged standardised coefficients (β) (with shrinkage) \pm their standard errors multiplied by 100, the significance categories (*** $p < 0.001$, ** $p < 0.01$, * $p < 0.1$) and, in parenthesis, the relative importance of each variable estimated as the sum of the AIC weights for each predictor included in n subsets of models having Δ_i (AIC_{best}-AIC_i) ≤ 2 . Deviance explained in relation to the null model was calculated for a final model including the best subset of variables. We also show the values corresponding to a final GLM model including variables retained after multi-model inference.

Family	Variable	PA	BS1	BS2	BS3	BS4	BS5
Topography	SLO	+0.65 \pm 12.94*** (1)	+0.01 \pm 0.09 (0.05)	+0.45 \pm 0.44 (0.67)	+0.87 \pm 7.7 (0.05)	-415 \pm 157.88** (1)	-449.45 \pm 140.94** (1)
	SLOSPcv	-4.01 \pm 71.83*** (1)	-	+0.69 \pm 0.43 (0.87)	-2.72 \pm 12.58 (0.11)	-90.79 \pm 45.07* (1)	-112.19 \pm 35.73** (1)
Climate	TMAmSU	+0.02 \pm 0.87* (0.1)	+0.7 \pm 0.43* (0.92)	+1.18 \pm 0.47* (1)	+6.18 \pm 18.5 (0.17)	-25.84 \pm 51.27 (0.33)	-97.42 \pm 41.6* (1)
	TMAcvSU	-0.55 \pm 30.29 (0.64)	+0.06 \pm 0.21 (0.18)	-1.35 \pm 0.45** (1)	-1.38 \pm 9.05 (0.06)	-238.34 \pm 69.87*** (1)	-214.49 \pm 46.37*** (1)
	PPsmWI	-0.01 \pm 0.02** (1)	-	-0.23 \pm 0.41 (0.36)	-	-3.19 \pm 25.77 (0.07)	-
	PPcvWI	-0.10 \pm 3.48** (1)	-0.32 \pm 0.38 (0.58)	+0.18 \pm 0.35 (0.31)	-0.49 \pm 6.93 (0.05)	-13.16 \pm 39.12 (0.24)	-8.73 \pm 24.66 (0.22)
Primary production	NDVImSP	-	-0.98 \pm 0.35** (1)	-	-73.63 \pm 30.32* (1)	-	+1.73 \pm 15.32 (0.13)
	NDVIevSP	+15.63 \pm 500.90** (1)	-0.01 \pm 0.1 (0.06)	-0.07 \pm 0.22 (0.16)	-5.34 \pm 17.11 (0.16)	-44.06 \pm 64.83 (0.46)	-3.99 \pm 18.86 (0.15)
Landscape structure	PERIMmax	+0.02 \pm 1.43 (0.47)	+0.53 \pm 0.41 (0.82)	+0.02 \pm 0.13 (0.07)	+3.72 \pm 15.77 (0.12)	+5.89 \pm 37.25 (0.08)	-1.91 \pm 17.34 (0.13)
	ARLAND	+0.04 \pm 0.56*** (1)	+0.05 \pm 0.19 (0.14)	+0.01 \pm 0.08 (0.03)	-1.07 \pm 8.39 (0.06)	+582.53 \pm 108.23*** (1)	+479.32 \pm 64.64*** (1)
Human disturbances	DISTRO	+0.01 \pm 0.01 (0.59)	+0.5 \pm 0.36 (0.88)	-	+31.06 \pm 33.06 (0.65)	+127.9 \pm 61.13* (1)	+82.81 \pm 34.01* (1)
	DISTIBA	+0.01 \pm 0.01*** (1)	+0.8 \pm 0.35* (1)	+0.07 \pm 0.23 (0.16)	+85.31 \pm 33.02* (1)	-116.5 \pm 48.94* (1)	-195.65 \pm 36.06*** (1)
Global Model	Nb of predictors	11	10	10	11	11	11
	Dev. explained	64.98	19.88	21.11	7.83	62.27	59.02
	Adj.r.squared	0.79	0.21	0.24	0.14	0.77	0.74

Table 4. Spatial verification of the models in occupied vs. potential breeding areas (i.e., breeding areas with successful reproduction vs. potential breeding areas not used for reproduction). The values shown are the mean (\pm standard deviation) habitat suitability for highly productive RG and random points, as well as the difference in suitability between them (the larger this difference, the better is the discrimination capacity of the models) and the statistical significance of these differences (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$) when the value is positive. PA represents the presence-absence model and BS the breeding success models.

	PA	BS1	BS2	BS3	BS4	BS5
Occupied breeding areas						
Suitability of highly productive RG	0.812 \pm 0.220	0.059 \pm 0.011	0.050 \pm 0.019	0.455 \pm 0.14	0.781 \pm 0.272	0.786 \pm 0.266
Suitability of random points	0.834 \pm 0.220	0.063 \pm 0.009	0.057 \pm 0.017	0.401 \pm 0.108	0.825 \pm 0.233	0.84 \pm 0.219
Mean difference of suitability	-0.023	-0.004	-0.007	0.054**	-0.044	-0.055
Potential breeding areas						
Suitability of highly productive RG	0.811 \pm 0.220	0.059 \pm 0.011	0.05 \pm 0.019	0.455 \pm 0.14	0.781 \pm 0.272	0.786 \pm 0.266
Suitability of random points	0.355 \pm 0.353	0.049 \pm 0.021	0.043 \pm 0.022	0.629 \pm 0.19	0.236 \pm 0.328	0.211 \pm 0.297
Mean difference of suitability	0.457***	0.010*	0.007*	-0.174	0.544***	0.575***

1
2
3 845 **Figure legends**
4

5 846
6

7 847 Figure 1. Geographic range of Great Bustard in Spain during the period 1987-2010 at
8
9 848 different seasons. Back dots correspond to: (a) Lek centres identified in Spring (Alonso
10 849 et al. 2012b), (b) flocks of females with chicks (isolated family or flock of females with
11
12 850 at least one family) detected in September, (c) flocks of females with a number of
13
14 851 chicks higher than the averaged mean value for the study period (Álvarez-Martínez *et*
15
16 852 *al.* 2015). In (b) and (c), points were defined by the location of the female flock with
17
18 853 chicks closest to the “centroid” of all female flocks, with or without chicks, constituting
19
20 854 a reproductive group. Grey buffers represent the species home range, estimated as a
21
22 855 buffer of 2 km (Palacín *et al.* 2012) around either lek centres or family flocks for either
23
24 856 Spring or September.
25
26
27
28
29

30 857

31
32 858 Figure 2. Spatial patterns of habitat suitability achieved from occurrence (PA) and
33
34 859 breeding success (BS) models.
35

36 860

37
38 861 Figure 3. Pearson ~~pair-wise~~ correlation matrices between habitat suitability for breeding
39
40 862 occurrence (PA) and breeding success (BS1 to BS5) in: (a) occupied breeding area, (b)
41
42 863 potential breeding area. The graph also shows histograms and scattered plots.
43
44

45 864

46
47 865 Figure 4. Spatial agreement between model outputs for occurrence (PA model) and
48
49 866 breeding success (BS4 ~~to~~ and BS5 models).
50
51

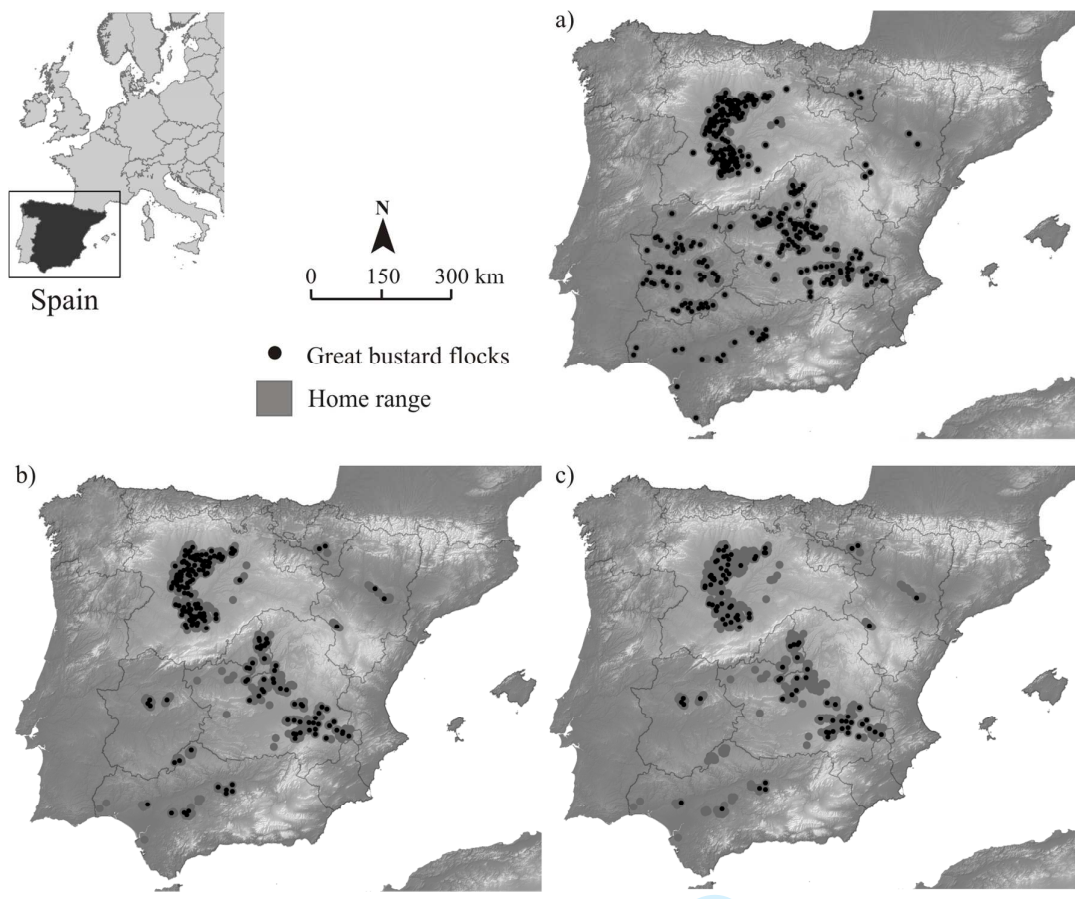
52 867
53

54 868
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

869 **Figure 1**

870



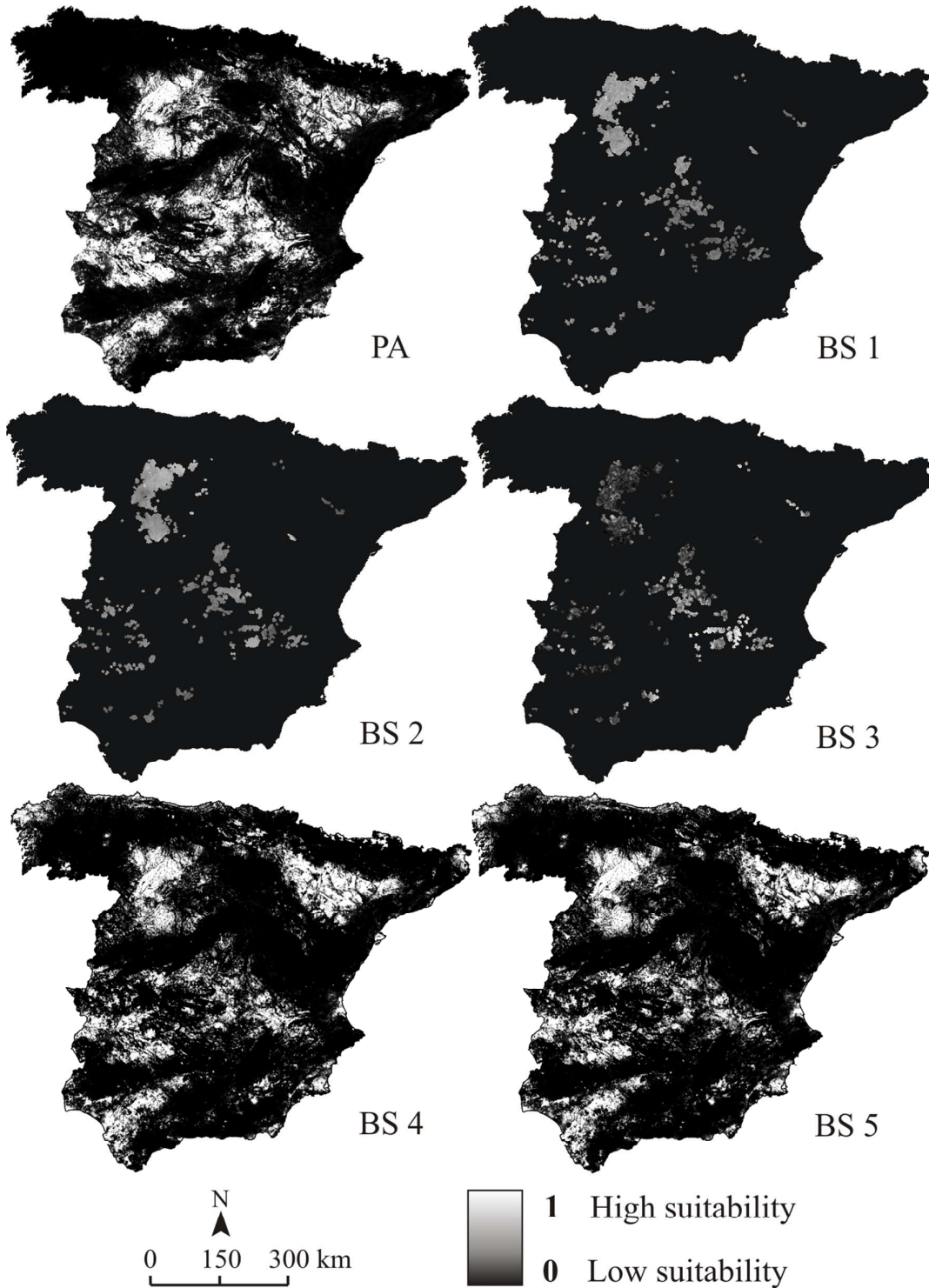
871

872

873

Only

874 **Figure 2**

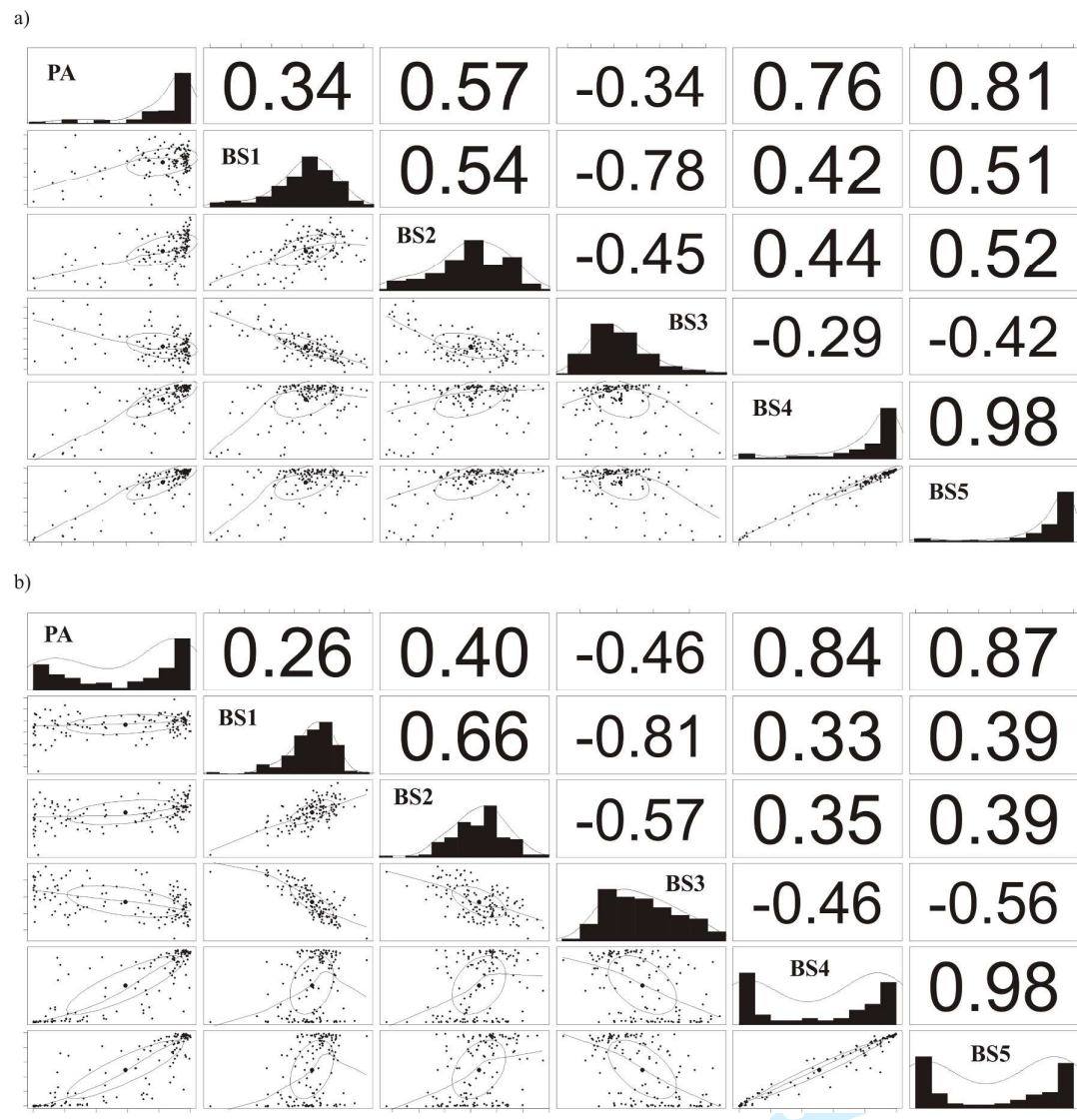


875

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

876 **Figure 3**

877



878

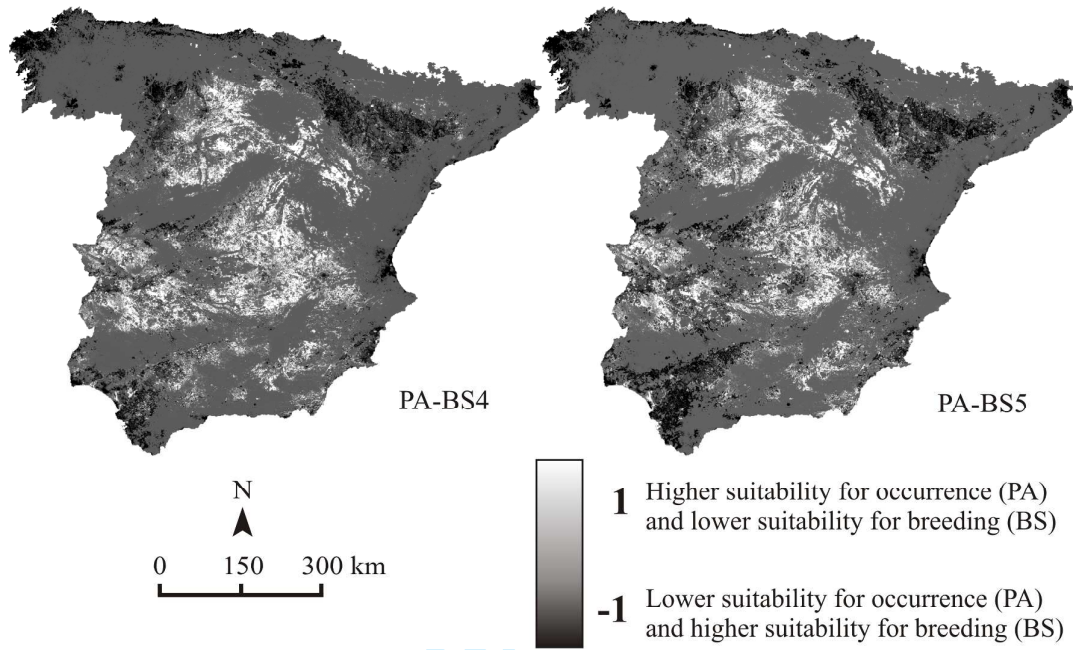
879

880

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

881 **Figure 4**

882



883

884

view Only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60