

8 **Abstract**

9 Predicting distribution has become a fundamental component in conservation or wildlife
10 management. Modelling is increasingly used to identify important areas (e.g. those
11 areas more suitable for a species or more likely to hold high densities). Models often
12 use presence/absence rather than abundance data, partly because measuring abundance
13 is more difficult than measuring presence. We aimed to test if the relationship between
14 occurrence models and predicted abundance varied for two sibling species that differ in
15 the level of nest aggregation: the Montagu's harrier (a semi-colonial raptor species) and
16 the hen harrier (more territorial). We modelled presence/absence distribution and the
17 number of pairs of each species with GLM and large-scale environmental variables, and
18 compared predicted results of both sets of models. In the case of the hen harrier,
19 predictions of the presence/absence model reliably identified areas with highest
20 densities for the species. In contrast, in the Montagu's harrier, there were large
21 apparently favourable areas where predicted breeding density was low. Our results
22 indicate that breeding system is likely to shape the relationship between
23 presence/absence vs density models. In species that are randomly or evenly spaced,
24 even if spatial variations in density occur, using results of presence/absence models is
25 likely to be adequate for population monitoring. In contrast, in the case of semi-colonial
26 species, it is necessary to take into account both occurrence and abundance models to
27 identify areas of conservation importance or concern. There are a considerable number
28 of birds which are semi-colonial or aggregated species, thus these results have general
29 implications.

30 **Keywords:** *Circus pygargus*, *Circus cyaneus*, favourability function, hen harrier,
31 Montagu's harrier, predictive models.

32

33 **1. Introduction**

34 The relationship between geographic distribution of species and abundance, and the
35 patterns and environmental conditions which determine them, is a central theme of
36 ecology and conservation (Brown, 1984; Nielsen et al., 2005). At a macroecological
37 level, species atlases are normally used as indicators of species distribution, but they
38 have the drawback that they can be biased depending on survey effort (Estrada et al.,
39 2008). Spatial models are normally used to attenuate this problem by providing
40 potential distribution of species, which do not depend so much on the survey effort
41 (Wilson et al., 2005; also in new generation atlases, e.g., Herrando et al., 2011).
42 Predicting distribution has thus become a fundamental component in disciplines such as
43 conservation or wildlife management, and in the assessment of the impact of climate
44 change (Carvalho et al., 2011; Estrada et al., 2010; Marini et al., 2010; Real et al.,
45 2010).

46 For example, it may be relevant to identify important areas for species of conservation
47 concern to concentrate conservation efforts there, or to assess how current distribution
48 of protected areas fits with the species distribution. These important areas (e.g. those
49 areas more suitable for the species, or more likely to hold high densities) are sometimes
50 identified based on existing information (i.e. local censuses), but other times they are
51 identified through modelling (Beresford et al., 2011; Coetzee et al., 2009; Estrada et al.,
52 2011; Greve et al., 2011; Marini et al., 2010). Many of those models use
53 presence/absence data rather than abundance data (e.g. Coetzee et al., 2009; Estrada et
54 al., 2011; Marini et al., 2010; Rodríguez-Estrella, 2007), partly because measuring
55 abundance is always more difficult than measuring presence and in a wide number of
56 species comparable data on spatial variation in abundance is lacking. Results of
57 presence/absence (P/A) models are sometimes considered as indicators of variations of

58 species abundance, as higher probability of occurrence may correlate with higher
59 number of individuals (Greaves et al., 2006; Sarà, 2008). The latter falls from the fact
60 that occurrence simply denotes that abundance at a site is greater than zero (Royle et al.,
61 2005), and thus from the hypotheses that P/A data follow an abundance-related
62 detection probability (Barbosa et al., 2009). Additionally, occurrence models reflect
63 habitat or environmental suitability (Jiménez-Valverde et al., 2009), and more
64 individuals are supposed to inhabit most suitable areas.

65 Being able to evaluate if P/A models can be reliably considered indicators of abundance
66 may thus be very useful. Overall, there are not numerous studies concerning the
67 presence/abundance relationship, and of those some have obtained a significant positive
68 correlation between suitability for the species and abundance (Coudun and Gégout,
69 2007; He and Gaston, 2003; Royle et al., 2005; Real et al., 2009) whereas others have
70 not found such association (Jiménez-Valverde et al., 2009; Nielsen et al., 2005; Pearce
71 and Ferrier, 2001). The general idea of these studies is that we can not assume that
72 occurrence models will always reflect species abundance. But the question is still
73 unresolved and it would be interesting to analyze in which cases this relationship is
74 valid, and where conclusions based on P/A models may be sufficient to identify
75 important areas for certain species.

76 In particular, bird species have a variety of breeding systems that strongly affect their
77 spatial distribution, thus spatial behaviour may modify the relationship between
78 occurrence and density. One could hypothesize that in species where nests are evenly or
79 randomly distributed (such as territorial species) both parameters are highly correlated,
80 whereas in semi-colonial species (where local breeding numbers may depend on social
81 parameters beyond the distribution of suitable habitat, Cornulier and Bretagnolle, 2006)
82 the relationship between occurrence and abundance may be weaker.

83 The Montagu's harrier *Circus pygargus* and the hen harrier *Circus cyaneus* are two
84 ground-nesting raptors listed on Annex 1 of the Birds Directive (Council Directive
85 2009/147/EC on the Conservation of Wild Birds). The Montagu's harrier is considered
86 vulnerable in the Iberian Peninsula, the stronghold of the Western European population
87 (Arroyo and García, 2008). Similarly, the hen harrier is considered vulnerable within
88 Spain, although the proportion of the Spanish population among European numbers is
89 smaller than for the other species (Arroyo and García, 2008). Both species are
90 extremely similar concerning their ecological requirements: they have similar diets,
91 both nest on the ground and use the same breeding habitats (Arroyo and García, 2008;
92 García and Arroyo, 2005; Millon et al., 2002). Regarding their breeding behaviour, the
93 Montagu's harrier is a semi-colonial species which presents clumped distribution of
94 nests (groups of 2 to > 40 pairs) (Arroyo, 1995; Cornulier and Bretagnolle, 2006). On
95 the other hand, the hen harrier is considered to be more territorial than the Montagu's
96 harrier (Arroyo and García, 2008; García and Arroyo, 2002), although breeding
97 densities may vary strongly spatially and some clumping of nests at larger scales have
98 been observed (Cornulier, 2005).

99 Our objective in this study was to test if occurrence models reflect predicted abundance
100 for the two harrier species, or if, in contrast, there are differences between the species in
101 relation to their level of nest aggregation. We discuss the relevance of the important
102 areas identified with one, the other or combined methods.

103

104 **2. Methods**

105 *2.1 Distribution models*

106 The most recent atlas for breeding birds in Spain was compiled in 2003 by the Spanish
107 Society of Ornithology (SEO/Birdlife) (Martí and del Moral, 2003) and it indicates
108 presence or absence of breeding pairs of each species in each 10 km x 10 km UTM cell.
109 Additionally, in 2006 SEO/Birdlife coordinated a national census of breeding
110 Montagu's and hen harriers in Spain (Arroyo and García, 2008), sampling random 100
111 km² UTM cells within the area of presence according to the atlas, and estimating the
112 minimum and maximum number of pairs of both species in each UTM cell. The
113 minimum number of pairs consisted on birds observed at the nest or with breeding
114 behaviour (feeding passes between male and female, nest material transport, etc.), while
115 the maximum number considered also pairs with possible breeding (males or females
116 observed in the area but no specific breeding behaviour). The protocol indicated a
117 minimum of three visits to each UTM cell at different times of the breeding cycles, each
118 lasting at least 3 hours. However, number of visits per cell varied among regions.
119 Maximum and minimum numbers calculated per cell were corrected for the number of
120 visits, which were found to influence total number of nests detected in a cell (Arroyo
121 and García, 2008). Therefore, the variable used for abundance analyses includes some
122 degree of correction for imperfect detection (Kéry et al., 2010), but it is likely that the
123 atlas data includes some non-quantified detection errors. The atlas presents combined
124 information collected (non systematically) over several years (with some repeated
125 observations of different areas), and a "presence" indicates a cell where breeding was
126 observed at least once, whereas "absence" indicates a cell where breeding has never
127 been observed.

128 We modelled the distributions of both species in Spain in relation to large-scale
129 environmental conditions, mainly related to topography, climate, land use and human
130 presence (Table 1). The digital procedure performed to obtain the variables was carried

131 out by the Biogeography, Diversity and Conservation Research Team of the University
132 of Málaga (Spain) and are detailed in Estrada et al. (2008). We fitted two models for
133 each species, i.e., a breeding P/A model with data from the atlas and a breeding density
134 model with data from the census. The P/A model was carried out using data from 5167
135 cells (covering the whole of peninsular Spain); for the breeding density model, data
136 used were 1220 cells (the extension of the census), but then this model was evaluated on
137 the complete set of Spanish UTM cells (n=5167). In order to be conservative, we
138 selected the minimum number of pairs of each species as the response variable in the
139 breeding density models. Both models were calibrated on a 75% random sample of the
140 original data and predictive accuracy was evaluated on the remaining 25% of the data
141 (Araújo et al., 2005).

142 The P/A model was calculated as follows. By performing logistic regression of each
143 species presence/absence on each variable separately we selected a subset of variables
144 significantly related to each species' distribution. To control for the increase in type I
145 error due to multiple testing (Benjamini and Hochberg, 1995; García, 2003), we
146 controlled the false discovery rate (FDR) using the procedure proposed by Benjamini
147 and Hochberg (1995), only accepting the variables that were significant under a FDR of
148 $q < 0.05$. Then we performed forward-backward stepwise multiple logistic regression of
149 presence/absence data on the subset of significant predictor variables. Finally we
150 applied the favourability function proposed by Real et al. (2006) on the predicted values
151 of logistic regression, thus converting logistic probabilities (P) into favourability (F)
152 values.

153 The logistic regression model has the form

$$154 \quad P = \frac{e^y}{1 + e^y}$$

155 In the favourability function the effect of prevalence is removed from the logit function
156 y . F can be obtained directly from probabilities of logistic regression (P) in the
157 following way (Real et al., 2006):

$$F = \frac{e^y}{\frac{n_1}{n_0} + e^y} = \frac{\frac{P}{(1-P)}}{\frac{n_1}{n_0} + \frac{P}{(1-P)}}$$

159 F is a logistic function with values between 0 and 1. Whereas P -values for different
160 species are not comparable because of the different prevalence of each species, F -values
161 are directly equivalent, and the models are then levelled to the same threshold of
162 favourability and can be compared and combined directly. The output value of $F=0.5$
163 will always correspond to the same neutral environmental threshold for all species,
164 whatever the proportion of presences in the sample. Favourability values higher than 0.5
165 correspond to areas where the probability of presence is higher than that expected
166 according to prevalence (i.e., higher than that expected by chance), and the opposite
167 occurs in areas with favourability below 0.5. Furthermore, the favourability value for a
168 species in a cell can be interpreted as the grade of membership of the cell to the fuzzy
169 set of cells that are favourable for the species. This fact allows the use of concepts and
170 operations of fuzzy logic to the resulting spatial analysis of the species. Additional
171 details on the favourability function can be found in Real et al. (2006).

172 We assessed the discrimination power of the P/A models with the validation dataset
173 (25% of data) by estimating the sensitivity, specificity, and their Correct Classification
174 Rate (CCR), using the favourability value of $F=0.5$ as classification threshold, and the
175 Area Under the Curve (AUC) of the Receiver Operating Characteristic, which is
176 independent of any favourability threshold (Hosmer and Lemeshow, 2000).

177 For the breeding density model, we fitted a Poisson GLM with log link (Yen et al.,
178 2004) to the minimum numbers of pairs on each variable separately, and then proceeded
179 analogously as in the P/A model to control for the FDR. We then performed forward–
180 backward stepwise GLM (according to AIC) fitted to a Poisson distribution of the
181 minimum number of pairs on the subset of significant predictor variables, to identify
182 factors affecting species abundance. In case any non-significant variables remained in a
183 model after this procedure, the model was further updated by removing them step by
184 step, starting with the least significant variable (Barbosa and Real, 2010; Crawley,
185 2007); we finally accepted the parsimonious model thus calculated provided that the
186 difference between the AIC score of the resultant model and the first model was less
187 than two units (Burnham and Anderson, 2002).

188 We assessed if Poisson models were affected by overdispersion. The dispersion
189 parameter is the quotient between the residual deviance and the degrees of freedom. It is
190 supposed to be equal to 1 under the Poisson model (i.e. the mean is equal to the
191 variance) but dispersion levels between 0.5 and 1.5 are frequently accepted (Zuur et al.,
192 2009). When a Poisson distribution was fitted to the observed counts of Montagu's
193 harrier we obtained overdispersion (3.7). When count data are not Poisson-distributed, a
194 good candidate distribution where the variance-mean ratio is greater than 1 is the
195 negative binomial distribution (Crawley, 2007, p.538). The meaning of the variance
196 being greater than the mean is that the population is more clumped or aggregated than
197 random, as Poisson distribution assumes that occurrences are randomly distributed
198 (Dytham, 1999). The Montagu's harrier is a semi-colonial species with clumped
199 distribution of nests, so it is logical that the number of pairs of this species had to be
200 modelled with a negative binomial distribution.

201 As with the P/A models, density models were calibrated using 75% of the data, and
202 were validated with the remaining 25% of the data by performing Spearman correlations
203 between results from the model and the raw abundance data from the validating data set.
204 We used Spearman correlations because model outputs had non-normal frequency
205 distributions. We also present the explained proportion of the deviance of the models.
206 Models were performed in SPSS and R.

207

208 *2.2 Combination of the models*

209 We computed Spearman correlations between the P/A model and the breeding density
210 model. With the intention of knowing how each model performed in predicting the
211 other dataset, we calculated the AUC and Spearman correlations between the abundance
212 model results (calculated with the training data set) and the raw presence/absence data
213 from the remaining 25% data. Analogously, we performed correlations between the P/A
214 model results (calculated with the training data set) and the raw abundance data from
215 the remaining 25% of data.

216 In order to know the degree with which the P/A model was spatially coincident with the
217 breeding density model, we performed and mapped different combinations between the
218 predicted values of both models for each species. These combinations were the
219 intersection and the subtraction. The favourability model (presence/absence) values
220 range between 0 and 1. We rescaled the breeding density model results dividing the
221 values in each cell by the maximum value predicted by the model, thus obtaining
222 predicted abundance also scaled between 0 and 1. The intersection represents the areas
223 that are simultaneously favourable for a set of criteria, and is computed in each locality
224 as the minimum value of both models. In our case, the biological meaning of the

225 intersection between the favourability model and the breeding density model represents
226 the areas that are simultaneously important for the presence of the species as a breeder
227 and for holding high breeding numbers. The subtraction of both models highlights the
228 areas where there are higher discrepancies between both models, and is calculated by
229 subtracting favourability values minus scaled predicted densities.

230

231 **3. Results**

232 *3.1 Favourability and breeding density models for both species*

233 We obtained the most favourable areas (Fig. 1a and Fig. 2a) and the estimated
234 abundance (Fig. 1b and Fig. 2b) of each species in peninsular Spain. Tables 2 and 3
235 show the environmental variables that affect their distributions.

236 Most favourable areas for Montagu's harrier were principally located in the Western
237 part of the country (Fig. 1a), with variables related to land use and spatial situation
238 among the first to form part of the model and the ones which define the distribution of
239 the species at a wide scale (Table 2). The favourability model presented an excellent
240 discrimination capacity according to Hosmer and Lemeshow (2000) (AUC: 0.835,
241 sensitivity: 0.79, specificity: 0.757, CCR: 0.773).

242 Highest predicted densities, after applying a negative binomial distribution, occurred in
243 the Southwestern part of the country and in some places of the Mediterranean coast
244 (Fig. 1b); the most important variables associated with higher densities at a wide scale
245 were climatic instability, low permeability and lowlands (Table 2). The correlation
246 between predicted values and the raw abundance data of the validation dataset was
247 significant ($\rho=0.445$, $p<0.001$) and the explained proportion of the deviance of the
248 model was 31.2%.

249 Hen harrier's favourability areas were concentrated in the Northern half of Spain (Fig.
250 2a), being primarily influenced at a macroecological scale by the spatial location, the
251 existence of broad valleys and water availability (Table 3). The favourability model for
252 this species presented also an outstanding discrimination capacity according to Hosmer
253 and Lemeshow (2000) (AUC: 0.901, sensitivity: 0.831, specificity: 0.794, CCR: 0.799).
254 Breeding density of hen harrier fitted a Poisson distribution (dispersion parameter=0.6).
255 Higher predicted density for this species occurred in Northern Spain (Fig. 2b), in areas
256 with low insolation and low climatic stability (Table 3). The correlation between
257 predicted values of the Poisson GLM and the raw abundance data of the validation
258 dataset was significant ($\rho=0.444$, $p<0.001$) and the explained proportion of the
259 deviance of the model was 54.1%.

260

261 *3.2 Coincidences between the P/A models and the breeding density models*

262 Fig. 1c and Fig. 2c show the intersection between the favourability model and the
263 breeding density model for the two species. Intersection values for Montagu's harrier
264 were lower than for the hen harrier, but were more widely distributed.

265 Results of the subtraction of the P/A model and the abundance model are shown in Fig.
266 1d and Fig. 2d. Dark colours highlight areas where favourability is higher than density,
267 and the opposite occurs with light colours. Neutral grey identifies coincidence between
268 both models. In the case of hen harrier, areas with high density were located within high
269 favourability areas (Fig. 2d). In the case of Montagu's harrier, there were large areas
270 where scaled density was much lower than favourability (in the North-West part of the
271 country, Fig. 1d), and some local areas where scaled density was higher than

272 favourability (along the Mediterranean coast and in the South-West part of the country,
273 Fig. 1d).

274 As can be deduced from Fig. 1 and Fig. 2, Spearman correlations between the P/A
275 model and the breeding density model were low for the Montagu's harrier ($\rho=0.219$,
276 $p<0.001$) and high for the hen harrier ($\rho=0.848$, $p<0.001$). Analogously, the
277 relationship between the abundance model results and the raw presence/absence data in
278 the 25% validating data set was low for the Montagu's harrier (AUC=0.611; Spearman
279 $\rho=0.192$, $p<0.001$) and high for the hen harrier (AUC=0.866; Spearman $\rho=0.437$,
280 $p<0.001$). Finally, the correlation between the P/A model results and the raw abundance
281 data in the 25% validating data set was 0.207 ($p<0.001$) for the former and 0.444
282 ($p<0.001$) for the latter.

283

284 **4. Discussion**

285 *4.1 Variables affecting distribution and abundance of harriers in Spain*

286 Our results highlight the environmental variables that are more likely to have a
287 relationship with both harrier species in Spain at a large scale. It is important to
288 remember that since GLM models assume perfect detection, and that it is likely not to
289 be the case, results obtained may be biased (Kéry et al., 2010). Additionally, our models
290 do not include food abundance, which is likely to influence breeding density (Millon et
291 al., 2008) and thus potentially occurrence. Inclusion of this variable would reduce the
292 percentage of unexplained deviance, but would probably change only slightly the shape
293 of the spatial predictions obtained in this study, because food abundance is likely to
294 have a strong significance at more local scales, whereas the environmental variables

295 included in this study shape macroecological distributions in these species (Anderson et
296 al. 2009; García and Arroyo, 2001; Redpath et al., 2002).

297 Montagu's harrier presence was mainly associated at a wide scale with land use and
298 location, secondarily to topographic or weather variables. This species occurred more
299 frequently in flat areas of continental Spain with large proportions of arable land (Table
300 2). This species is traditionally characteristic of grassland areas, and more recently has
301 taken to occupy agricultural areas, nesting mainly within cereal crops (Arroyo et al.,
302 2004), which concurs with our results. On the other hand, spatial variation in Montagu's
303 harrier breeding density at a wide scale was affected by location (being also higher in
304 the West), but also by climatic instability and soil permeability (density being higher in
305 areas with high pluviometric irregularity and impermeable soils). These latter variables
306 may be related to nesting habitat suitability. However, and against what could be
307 expected, breeding density is not affected at this scale by land use variables.

308 Hen harrier favourability areas were concentrated in the Northern half of Spain, and
309 both presence and density were higher in relatively high altitude areas and with low
310 insolation (Table 3). García and Arroyo (2001) have shown that high temperatures are
311 unfavourable for this species' breeding, so their absence from hotter Southern Spain and
312 their association with higher altitudes and latitudes is not surprising. Additionally, this
313 species nests more frequently in scrub and bushy areas (Arroyo and García, 2008),
314 which are more common in mountainous areas; shrublands, in fact, was the only land
315 use variable which formed part of both models (although availability of arable land,
316 another nesting habitat, also increased breeding density).

317

318 *4.2 Distribution and favourability models as indicators of abundance*

319 Regarding the coincidence between occurrence models and predicted abundance, we
320 obtained significant positive correlations between predictions of both models, but
321 results differed for the semi-colonial and the not colonial species, as the correlation was
322 much lower for the Montagu's harrier than for the hen harrier. In other words, in the hen
323 harrier, spatial predictions from the favourability model and the breeding density model
324 were largely concordant. Concurrently, 11 variables coincide in describing both models
325 (Table 3). On the other hand, there were large discrepancies between the favourability
326 model and the breeding density model in the Montagu's harrier, both in terms of
327 variables included in the models, or their spatial results.

328 The relationship between occurrence and abundance is directly influenced by mean
329 abundance (Royle et al., 2005), as abundance itself influences the probability of
330 detecting organisms. In general, stronger correlations between abundance and
331 occurrence are expected when abundance is lower. Therefore, the different relationship
332 between hen harrier and Montagu's harrier P/A and abundance predictions may simply
333 arise if mean abundance was lower in the former than the latter species. Mean observed
334 abundance (within their distribution range) was indeed higher in the Montagu's than the
335 hen harrier (2.5 vs 0.9 pairs/km² respectively), although differences were not large
336 enough to justify the strong difference in correlation values between both species (0.2
337 vs 0.8). Additionally, in our opinion, differences in abundance alone cannot explain the
338 considerable spatial discrepancies between models for only one of the species.

339 On the other hand, imperfect detection may also influence results (Kéry et al., 2010).
340 We could not estimate or integrate detection errors in our models. However, both
341 species are likely to have the same biases relating to presence data (the same protocols
342 are used, both species show very visible flights during pre-laying, used to identify
343 potential breeding areas, etc.), and errors associated to detection for the estimation of

344 abundance data have already been (at least partially) taken into account (Arroyo and
345 García, 2008, see methods). Therefore, we do not believe that differences in detection
346 probability alone can explain discrepancies observed.

347 Our results, in our opinion, are better explained by differences in territorial and spatial
348 behaviour between both species. In more territorial species, nest site occupancy is likely
349 to be strongly dependent on environmental quality. Concurrently, spatial predictions
350 from the favourability model and the breeding density model are likely to be largely
351 concordant, as found for the hen harrier. This means that areas identified as optimal
352 with P/A models are also likely to contain the highest densities. This has implications
353 for population monitoring and for conservation purposes: here, variations in
354 presence/absence (which are sometimes the only information available) are highly likely
355 to indicate population trends, and P/A models are adequate for identifying important
356 areas for the species.

357 In contrast, there was a large discrepancy between the favourability model and the
358 breeding density model in the semi-colonial Montagu's harrier. There were large
359 apparently favourable areas where predicted breeding density was low, and in contrast
360 zones with high density that fell outside most favourable areas (Fig. 1). This suggests
361 that other factors beyond those considered here strongly affect variation in abundance in
362 this species. Conspecific attraction in semi-colonial or lekking species can lead
363 individuals to leave unoccupied apparently suitable habitats (Heinänen et al., 2008;
364 Osborne et al., 2007). In our case, it is known that local abundance of Montagu's harrier
365 is influenced by both social attraction and spatial heterogeneity in environmental quality
366 (Cornulier, 2005), and is significantly related to the previous presence and productivity
367 of the species (Arroyo et al., 2002). Thus, the reduced strength of coincidence between

368 occurrence models and abundance models may be related to the past demographic
369 history of the species in different areas.

370 This means that using results from P/A models for estimating variations in abundance
371 may render spurious results. In contrast, comparisons between both estimates may be
372 particularly relevant to identify population processes. In particular, it may be insightful
373 to evaluate what happens in areas with highest intersection values (highly favourable,
374 with high densities) in comparison with others with high favourability but low densities,
375 as these differences may be indicative of a problem (e.g. lack of prey in an area, or
376 unusually low reproductive values). For example, there are no conservation programs
377 for protecting Montagu's harrier nests at harvest time in Castilla y León (with high
378 favourability but low breeding density), so it is possible that productivity is frequently
379 very low and this reduces the attractiveness of the area (Arroyo et al., 2002).

380 Additionally, it may be relevant to maintain monitoring programs and identify processes
381 in areas with high density but relatively low favourability, such as those in the
382 Mediterranean coast.

383 Several authors have suggested that relating P/A models to abundance could be a way
384 of validating the former (Lobo et al., 2008). But normally a positive and significant
385 association is considered enough to establish this relationship (Jiménez-Valverde et al.,
386 2009; Real et al., 2009). In the present study we have obtained positive and significant
387 associations for both species but we consider that it is important to take into account not
388 only the significance but also the rho values of the correlation. This is a restrictive
389 approach and in our case it is more realistic because it would not be consistent to
390 consider the occurrence/abundance relationship similar for both species when there is a
391 difference higher than 0.5 between the rho values for the two harrier species analyzed.

392 Overall, our results indicate that breeding system is likely to shape the relationship
393 between presence/absence vs density models. In species that are relatively evenly-
394 spaced, even if spatial variations in density occur, using the results of P/A models is
395 likely to be adequate for population monitoring. In contrast, in other species where the
396 social component is important in habitat selection, population monitoring needs to
397 specifically take local abundance into account. These results have general implications
398 because there are a considerable number of birds which are semi-colonial or aggregated
399 species, and it would be necessary to combine information from both favourability
400 models and abundance models to identify areas of conservation importance or concern
401 for them.

402

403 **Acknowledgments**

404 A. Estrada had a postdoctoral contract jointly financed by the European Social Fund and
405 by the Junta de Comunidades de Castilla-La Mancha (Spain), in the framework of the
406 Operational Programme FSE 2007-2013. We thank A. Márcia Barbosa for her valuable
407 R tips and Pelayo Acevedo for his help with the modelling procedure. Marc Kéry and
408 two anonymous referees improved the manuscript with their comments. Data from the
409 censuses were made available by J.C. del Moral from the Spanish Society of
410 Ornithology (SEO-Birdlife). We also thank all the volunteers that participated in the
411 censuses.

412

413 **References**

414 Anderson, B.J., Arroyo, B., Collingham, Y.C., Etheridge, B., Fernandez-de-Simon, J.,
415 Gillings, S., Gregory, R., Leckie, F., Thomas, C.D., Travis, J., Redpath, S.M.,

416 2009. Using distribution models to test alternative hypotheses about a species'
417 environmental limits and recovery prospects. *Biological Conservation* 142, 488-
418 499.

419 Araújo, M.B., Pearson, R.G., Thuiller, W., Erhard, M., 2005. Validation of species–
420 climate impact models under climate change. *Global Change Biology* 11, 1504-
421 1513.

422 Arroyo, B., 1995. Breeding ecology and nest dispersion of Montagu's Harrier *Circus*
423 *pygargus* in central Spain. PhD Thesis, University of Oxford.

424 Arroyo, B., García, J., 2008. El aguilucho cenizo y el aguilucho pálido en España.
425 Población en 2006 y método de censo. SEO/Birdlife, Madrid.

426 Arroyo, B., Garcia, J.T., Bretagnolle, V., 2002. Conservation of the Montagu's harrier
427 (*Circus pygargus*) in agricultural areas. *Animal Conservation* 5, 283-290.

428 Arroyo, B E., Garcia, J.T., Bretagnolle, V., 2004. Montagu's Harrier. BWP update 6,
429 41-55.

430 Barbosa A.M., Real R., 2010. Favourable areas for expansion and reintroduction of
431 Iberian lynx accounting for distribution trends and genetic diversity of the
432 European rabbit. *Wildlife Biology in Practice* 6, 34-47

433 Barbosa, A.M., Real, R., Vargas, J.M., 2009. Transferability of environmental
434 favourability models in geographic space: The case of the Iberian desman
435 (*Galemys pyrenaicus*) in Portugal and Spain. *Ecological Modelling* 220, 747-
436 754.

437 Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and
438 powerful approach to multiple testing. *Journal of the Royal Statistical Society:*
439 Series B 57, 289-300.

440 Beresford, A.E., Buchanan, G.M., Donald, P.F., Butchart, S.H.M., Fishpool, L.D.C.,
441 Rondinini, C., 2011. Poor overlap between the distribution of Protected Areas
442 and globally threatened birds in Africa. *Animal Conservation* 14, 99–107.

443 Brown, J. H., 1984. On the relationship between abundance and distribution of species.
444 *The American Naturalist* 124, 255-279.

445 Burnham, K.P., Anderson, D.R., 2002. Model selection and multimodel inference: a
446 practical information-theoretic approach, 2nd edn. Springer-Verlag, New York,
447 USA. 488 pp.

448 Carvalho, S.B., Brito, J.C., Crespo, E.G., Watts, M.E., Possingham, H.P., 2011.
449 Conservation planning under climate change: Toward accounting for uncertainty
450 in predicted species distributions to increase confidence in conservation
451 investments in space and time. *Biological Conservation* 144, 2020-2030.

452 Coetsee, B.W.T., Robertson, M.P., Erasmus, B.F.N., van Rensburg, B J., Thuiller, W.,
453 2009. Ensemble models predict Important Bird Areas in southern Africa will
454 become less effective for conserving endemic birds under climate change.
455 *Global Ecology and Biogeography* 18, 701-710.

456 Cornulier, T., 2005. Composantes de la distribution spatiale d'un prédateur: effets
457 respectifs de l'habitat, des ressources alimentaires et des interactions
458 comportementales. PhD thesis, Université Claude Bernard, Lyon 1.

459 Cornulier, T., Bretagnolle, V., 2006. Assessing the influence of environmental
460 heterogeneity on bird spacing patterns: a case study with two raptors. *Ecography*
461 29, 240-250.

462 Coudun, C., Gégout, J. C., 2007. Quantitative prediction of the distribution and
463 abundance of *Vaccinium myrtillus* with climatic and edaphic factors. *Journal of*
464 *Vegetation Science* 18, 517-524.

465 Crawley, M. J., 2007. *The R Book*. John Wiley & Sons, Chichester (UK).

466 Dytham, C., 1999. *Choosing and using statistics: a biologist's guide*. Blackwell Science
467 Ltd. Oxford.

468 Estrada, A., Arroyo, B., Márquez, A.L., 2010. Possible changes in favourability areas
469 for Montagu's and hen harriers in Spain according to climate change scenarios.
470 *Ardeola* 57, 123-128.

471 Estrada, A., Real, R., Vargas, J.M., 2008. Using crisp and fuzzy modelling to identify
472 favourability hotspots useful to perform gap analysis. *Biodiversity and*
473 *Conservation* 17, 857-871.

474 Estrada, A., Real, R., Vargas, J.M., 2011. Assessing coincidence between priority
475 conservation areas for vertebrate groups in a Mediterranean hotspot. *Biological*
476 *Conservation* 144, 1120-1129.

477 Farr, T.G., Kobrick, M., 2000. Shuttle Radar Topography Mission produces a wealth of
478 data. *Eos, Transactions, American Geophysical Union* 81, 583-585.

479 Font, I., 2000. *Climatología de España y Portugal*. Ediciones Universidad de
480 Salamanca.

481 García, L.V., 2003. Controlling the false discovery rate in ecological research. *Trends in*
482 *Ecology and Evolution* 18 (11), 553-554.

483 García, J.T., Arroyo, B.E., 2001. Effect of abiotic factors on reproduction in the centre
484 and periphery of breeding ranges: a comparative analysis in sympatric harriers.
485 *Ecography* 24, 393-402.

- 486 García, J.T., Arroyo, B.E., 2002. Intra and interspecific agonistic behaviour in
487 sympatric harriers during the breeding season. *Animal Behaviour* 64, 77-84.
- 488 García, J.T., Arroyo, B.E., 2005. Food-niche differentiation in sympatric Hen and
489 Montagu's harriers. *Ibis* 147, 144-154.
- 490 Greaves, R.K., Sanderson, R.A., Rushton, S.P., 2006. Predicting species occurrence
491 using information-theoretic approaches and significance testing: An example of
492 dormouse distribution in Cumbria, UK. *Biological Conservation* 130, 239-250.
- 493 Greve, M., Chown, S.L., van Rensburg, B.J., Dallimer, M., Gaston, K.J., 2011. The
494 ecological effectiveness of protected areas: a case study for South African birds.
495 *Animal Conservation* 14, 295-305.
- 496 He, F., Gaston, K.J., 2003. Occupancy, spatial variance, and the abundance of species.
497 *The American Naturalist* 162, 366-375.
- 498 Heinänen, S., Rönkä, M., von Numers, M., 2008. Modelling the occurrence and
499 abundance of a colonial species, the arctic tern *Sterna paradisaea* in the
500 archipelago of SW Finland. *Ecography* 31, 601-611.
- 501 Herrando, S., Brotons, L., Estrada, J., Guallar, S., Anton, M. (eds.), 2011. *Atles dels*
502 *ocells de Catalunya a l'hivern 2006-2009*. Institut Català d'Ornitologia (ICO) /
503 Lynx Edicions. Barcelona.
- 504 Hosmer, D.W., Lemeshow, S., 2000. *Applied Logistic Regression*, 2nd edition. John
505 Wiley and Sons, Inc. New York.
- 506 IGME., 1979. *Mapa hidrogeológico nacional. Explicación de los mapas de lluvia útil,*
507 *de reconocimiento hidrogeológico y de síntesis de los sistemas acuíferos. Vol.*
508 *81, 2ª ed.* Instituto Geológico y Minero de España. Ministerio de Industria y
509 Energía, Madrid.

510 IGN., 1999. Mapa de carreteras. Península Ibérica, Baleares y Canarias. Instituto
511 Geográfico Nacional, Ministerio de Fomento. Madrid.

512 Jiménez-Valverde, A., Diniz, F., de Azevedo, E.B., Borges, P.A.V., 2009. Species
513 distribution models do not account for abundance: the case of arthropods on
514 Terceira Island. *Ann. Zool. Fennici* 46, 451-464.

515 Kéry, M., Gardner, B., Monnerat, C., 2010. Predicting species distributions from
516 checklist data using site-occupancy models. *Journal of Biogeography* 37, 1851-
517 1862.

518 Lobo, J.M., Jiménez-Valverde, A., Real, R., 2008. AUC: a misleading measure of the
519 performance of predictive distribution models. *Global Ecology and*
520 *Biogeography* 17, 145-151.

521 Marini, M.A., Barbet-Massin, M., Martinez, J., Prestes, N.P., Jiguet, F., 2010. Applying
522 ecological niche modelling to plan conservation actions for the Red-spectacled
523 Amazon (*Amazona pretrei*). *Biological Conservation* 143, 102-112.

524 Martí, M., del Moral, J. C. (eds.), 2003. Atlas de las aves reproductoras de España.
525 Dirección General de Conservación de la Naturaleza-Sociedad Española de
526 Ornitología, Madrid, Spain.

527 MFE., 2007. Mapa Forestal de España. Ministerio de Medio Ambiente, Medio Rural y
528 Marino.

529 Millon, A., Bourrioux, J.L., Riols, C., Bretagnolle, V., 2002. Comparative breeding
530 biology of Hen Harrier and Montagu's Harrier: an 8-year study in north-eastern
531 France. *Ibis* 144, 94–105.

532 Millon, A., Arroyo, B.E., Bretagnolle, V., 2008. Variable but predictable prey
533 availability affects predator breeding success: natural versus experimental
534 evidence. *Journal of Zoology* 275, 349-358.

535 Montero de Burgos, J.L., González-Rebollar, J.L., 1974. Diagramas bioclimáticos.
536 ICONA, Madrid.

537 Muñoz, A.R., Real, R., 2006. Assessing the potential range expansion of the exotic
538 monk parakeet in Spain. *Diversity and Distributions* 12, 656-665.

539 Nielsen, S.E., Johnson, C.J., Heard, D.C., Boyce, M.S., 2005. Can models of presence-
540 absence be used to scale abundance? Two case studies considering extremes in
541 life history. *Ecography* 28,197-208.

542 ORNL., 2001. LandScan 2000 Global Population Database. Oak Ridge National
543 Laboratory (ORNL), Oak Ridge, TN.

544 Osborne, P.E., Suárez-Seoane, S., Alonso, J.C., 2007. Behavioural mechanisms that
545 undermine species envelope models: the causes of patchiness in the distribution
546 of great bustards *Otis tarda* in Spain. *Ecography* 30, 819-828.

547 Pearce, J., Ferrier, S., 2001. The practical value of modelling relative abundance of
548 species for regional conservation planning: a case study. *Biological*
549 *Conservation* 98, 33-43.

550 Real, R., Barbosa, A.M., Rodríguez, A., García, F.J., Vargas, J.M., Palomo, L.J.,
551 Delibes, M., 2009. Conservation biogeography of ecologically interacting
552 species: the case of the Iberian lynx and the European rabbit. *Diversity and*
553 *Distributions* 15, 390-400.

554 Real, R., Barbosa, A.M., Vargas, J.M., 2006. Obtaining environmental favourability
555 functions from logistic regression. *Environmental and Ecological Statistics* 13,
556 237-245.

557 Real, R., Márquez, A.L., Olivero, J., Estrada, A., 2010. Species distribution models in
558 climate change scenarios are still not useful for informing policy planning: an
559 uncertainty assessment using fuzzy logic. *Ecography* 33, 304-314.

560 Redpath, S.M., Arroyo, B.E., Etheridge, B., Leckie, F., Bouwman, K., Thirgood, S.J.,
561 2002. Temperature and hen harrier productivity: from local mechanisms to
562 geographical patterns. *Ecography* 25, 533-540.

563 Rodríguez-Estrella, R., 2007. Land use changes affect distributional patterns of desert
564 birds in the Baja California peninsula, Mexico. *Diversity and Distributions* 13,
565 877-889.

566 Royle, J.A., Nichols, J.D., Kéry, M., 2005. Modelling occurrence and abundance of
567 species when detection is imperfect. *Oikos* 110, 353-359.

568 Sarà, M., 2008. Breeding abundance of threatened raptors as estimated from occurrence
569 data. *Ibis* 150, 766-778.

570 SIGPAC., 2007. Sistema de Información Geográfica de Identificación de Parcelas
571 Agrícolas. Ministerio de Medio Ambiente, Medio Rural y Marino.

572 U. S. Geological Survey., 1996. GTOPO30. Land processes distributed active archive
573 center (<http://edcdaac.usgs.gov/gtopo30/gtopo30.asp>).

574 Wilson, K.A., Westphal, M.I., Possingham, H.P., Elith, J., 2005. Sensitivity of
575 conservation planning to different approaches to using predicted species
576 distribution data. *Biological Conservation* 122, 99-112.

577 Yen, P.P.W., Huettmann, F., Cooke, F., 2004. A large-scale model for the at-sea
578 distribution and abundance of Marbled Murrelets (*Brachyramphus marmoratus*)
579 during the breeding season in coastal British Columbia, Canada. Ecological
580 Modelling 171, 395-413.

581 Zuur, A.F., Ieno, E.N., Walker, N.J., Saveliev, A.A., Smith, G.M., 2009. Mixed Effects
582 Models and Extensions in Ecology with R. Springer Science+Business Media,
583 LLC, New York.

584

Table 1 Variables, and their associated factors, used to model harriers distribution in Spain.

Factor	Variable	Code
Topography	Mean altitude (m) ^(a)	<i>Alti</i>
	Slope (°) (calculated from <i>Alti</i>)	<i>Slop</i>
	Southward exposition degree ^(b)	<i>SE</i>
	Westward exposition degree ^(b)	<i>WE</i>
Water availability	Mean relative air humidity in January at 07:00 (%) ^(c)	<i>HJan</i>
	Mean relative air humidity in July at 07:00 (%) ^(c)	<i>HJul</i>
	Mean annual runoff (mm) ^(d)	<i>ROff</i>
	Mean annual number of snow days ^(c)	<i>DSno</i>
	Humidity index ^(c)	<i>HumI</i>
	Mean annual precipitation (mm) ^(c)	<i>Prec</i>
Energy availability	Mean annual insolation (h/year) ^(c)	<i>Inso</i>
	Mean annual solar radiation ((kWh/m ² /day)*100) ^(c)	<i>SRad</i>
	Mean temperature in January (°C) ^(c)	<i>TJan</i>
	Mean temperature in July (°C) ^(c)	<i>TJul</i>
	Mean annual temperature (°C) ^(c)	<i>Temp</i>
	Mean annual potential evapotranspiration (mm) ^(c)	<i>PET</i>
Productivity	Mean annual number of frost days (minimum temperature ≤ 0°C) ^(c)	<i>DFro</i>
	Mean annual actual evapotranspiration (mm) (= min [<i>Prec</i> , <i>PET</i>])	<i>AET</i>
Climatic stability	Annual relative air humidity range (%) (= <i>HuJan</i> - <i>HuJul</i>)	<i>HRan</i>
	Annual temperature range (°C) (= <i>TJul</i> - <i>TJan</i>)	<i>TRan</i>
	Mean annual number of days with precipitation ≥ 0.1 mm ^(c)	<i>DPre</i>
	Continental index ^(c)	<i>ConI</i>
Disturbances	Pluviometric irregularity ^(e)	<i>PIrr</i>
	Maximum precipitation in 24 h (mm) ^(c)	<i>MP24</i>
Spatial location	Relative maximum precipitation (= <i>MP24</i> / <i>Prec</i>)	<i>RMP</i>
	Latitude (°N) ^(f)	<i>Lati</i>
Human presence	Longitude (°E) ^(f)	<i>Long</i>
	Distance to the nearest highway (km) ^(f)	<i>DHi</i>
	Distance to the nearest urban center with more than 100 000 inhabitants (km) ^(f)	<i>U100</i>
	Distance to the nearest urban center with more than 500 000 inhabitants (km) ^(f)	<i>U500</i>
Lithology	Human population density in 2000 (number of inhabitants/km ²) ^(g)	<i>HPd</i>
	Soil permeability ^(d)	<i>Perm</i>
Land use	Arable land ^(h)	<i>Arab</i>
	Grasslands ⁽ⁱ⁾	<i>Gras</i>
	Shrublands ⁽ⁱ⁾	<i>Shru</i>

Sources: ^(a) U. S. Geological Survey (1996). ^(b) Shuttle Radar Topography Mission (SRTM), Farr and Kobrick (2000). ^(c) Font (2000). ^(d) IGME (1979). ^(e) Montero de Burgos and González-Rebollar (1974). ^(f) IGN (1999). ^(g) ORNL (2001). ^(h) SIGPAC (2007). ⁽ⁱ⁾ MFE (2007).

Table 2 Montagu's harrier models. Variables: variables included in the models; Rank: order of entrance of the variables in the models; β : coefficients; SE: standard errors; P: significance: ***<0.001, **<0.01, *<0.05. Variables codes as in Table 1. In bold coincident variables for both models.

Favourability model (presence/absence)					Breeding density model				
Variables	Rank	β	SE	P	Variables	Rank	β	SE	P
<i>Long</i>	2	-0.3	0.03	***	<i>Long</i>	6	-0.2	0.03	***
<i>Lati</i>	3	0.2	0.05	***	<i>Lati</i>	13	0.2	0.06	**
<i>Slop</i>	10	-0.2	0.03	***	<i>Slop</i>	3	-0.1	0.04	***
<i>RMP</i>	17	-5.8	1.02	***	<i>RMP</i>	4	-6.3	1.1	***
<i>MP24</i>	18	0.01	0.002	***	<i>MP24</i>	7	0.007	0.002	**
<i>PET</i>	12	0.004	0.001	*	<i>PET</i>	17	0.004	0.001	**
<i>DFro</i>	15	-0.01	0.004	*	<i>DFro</i>	15	-0.02	0.004	***
<i>Arab</i>	1	0.03	0.002	***	<i>PIrr</i>	1	0.07	0.01	***
<i>DHi</i>	4	-0.01	0.002	***	<i>Perm</i>	2	-0.1	0.06	**
<i>HJan</i>	5	0.1	0.02	***	<i>ROff</i>	5	-0.001	0.0004	**
<i>HRan</i>	6	-0.1	0.02	***	<i>Prec</i>	8	-0.001	0.0004	**
<i>AET</i>	7	-0.004	0.001	***	<i>ConI</i>	9	-0.03	0.02	*
<i>U500</i>	8	-0.004	0.001	***	<i>SRad</i>	10	0.009	0.003	***
<i>U100</i>	9	0.004	0.002	*	<i>Inso</i>	11	-0.001	0.0004	**
<i>Shru</i>	11	0.01	0.004	***	<i>DPre</i>	12	-0.01	0.003	**
<i>HPd</i>	13	-0.0004	0.0002	*	<i>DSno</i>	14	0.03	0.01	**
<i>TJan</i>	14	-0.3	0.06	***	<i>Alti</i>	16	0.001	0.0004	**
<i>HJul</i>	15	-0.05	0.02	*	Intercept		-9.3	3.8	*
<i>SE</i>	16	-0.007	0.003	*					
Intercept		-10.7	2.8	***					

Table 3 Hen harrier models. Variables: variables included in the models; Rank: order of entrance of the variables in the models; β : coefficients; SE: standard errors; P: significance: ***<0.001, **<0.01, *<0.05. Variables codes as in Table 1. In bold coincident variables for both models.

Favourability model (presence/absence)					Breeding density model				
Variables	Rank	β	SE	P	Variables	Rank	β	SE	P
<i>Inso</i>	13	-0.002	0.0005	***	<i>Inso</i>	1	-0.003	0.0005	***
<i>PIrr</i>	4	0.1	0.03	***	<i>PIrr</i>	2	0.2	0.04	***
<i>Lati</i>	1	2.03	0.2	***	<i>Lati</i>	3	0.9	0.2	***
<i>U500</i>	5	-0.01	0.002	***	<i>U500</i>	4	-0.01	0.002	***
<i>ROff</i>	7	-0.002	0.0005	***	<i>ROff</i>	14	0.001	0.0005	*
<i>Alti</i>	9	0.002	0.0005	***	<i>Alti</i>	7	0.002	0.0004	***
<i>U100</i>	20	-0.006	0.003	*	<i>U100</i>	8	-0.02	0.004	***
<i>Slop</i>	14	-0.2	0.04	***	<i>Slop</i>	9	-0.1	0.05	*
<i>Shru</i>	15	0.02	0.005	***	<i>Shru</i>	12	0.02	0.007	**
<i>AET</i>	17	-0.005	0.001	***	<i>AET</i>	13	0.003	0.001	*
<i>Prec</i>	19	-0.001	0.0005	*	<i>Prec</i>	15	-0.001	0.0005	*
<i>HJan</i>	2	0.05	0.02	*	<i>Arab</i>	5	0.02	0.005	***
<i>DSno</i>	3	0.06	0.01	***	<i>SE</i>	6	0.02	0.006	***
<i>Long</i>	6	-0.5	0.07	***	<i>TRan</i>	10	0.2	0.08	**
<i>MP24</i>	8	0.03	0.004	***	<i>Perm</i>	11	0.2	0.1	*
<i>RMP</i>	10	-12.2	2.8	***	Intercept		-43.2	8.9	***
<i>ConI</i>	11	0.1	0.03	**					
<i>WE</i>	12	-0.02	0.006	**					
<i>PET</i>	16	0.01	0.003	***					
<i>HumI</i>	18	1.97	0.6	**					
Intercept		-95.1	8.3	***					

Figure legends

Figure 1 Predicted maps for Montagu's harrier in Spain at 10 km x 10 km resolution. (a) Predicted favourability, (b) estimated abundance, (c) intersection, and (d) subtraction (favourability – scaled abundance).

Figure 2 Predicted maps for hen harrier in Spain at 10 km x 10 km resolution. (a) Predicted favourability, (b) estimated abundance, (c) intersection, and (d) subtraction (favourability – scaled abundance).

Figure 1

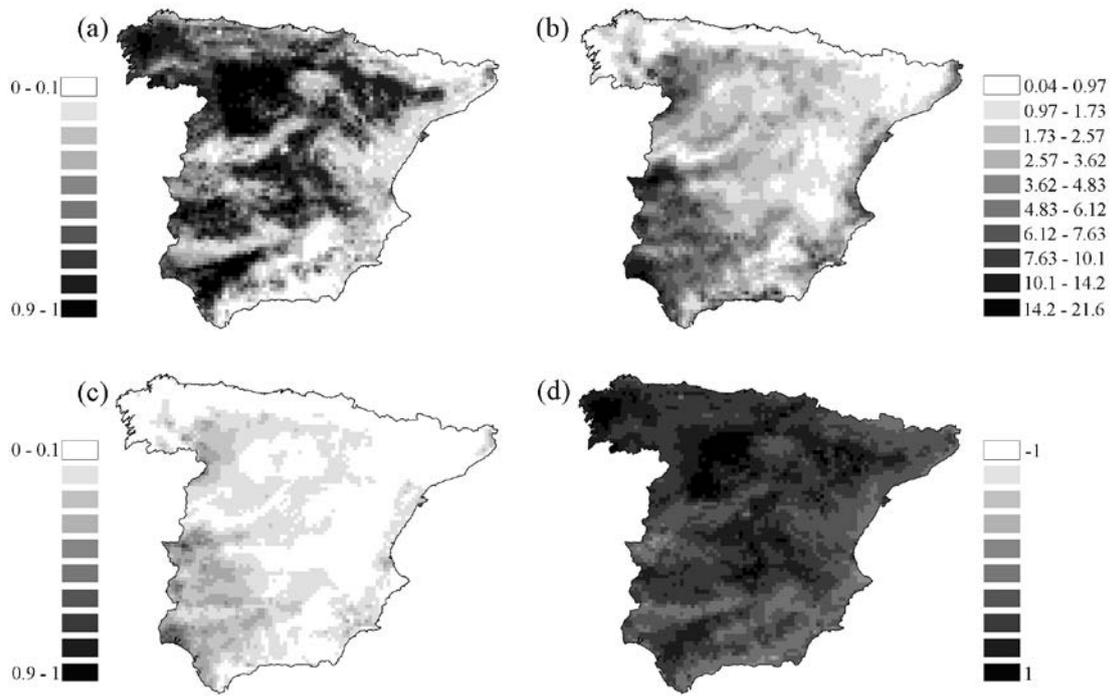


Figure 2

