From general research questions to specific answers: Underspecificity as a source of uncertainty in biological conservation

Susana Suárez-Seoane[1,a,*]

s.seoane@unileon.es

Jose Manuel Álvarez-Martínez[2,b]

Carlos Palacín[3,c]

Juan Carlos Alonso[3,c]

[1,a]Área de Ecología, Departamento de Biodiversidad y Gestión Ambiental, Facultad de Ciencias Biológicas y Ambientales, Universidad de León, Campus de Vegazana s/n, 24071 León, Spain

[2,b]Environmental Hydraulics Institute "IH Cantabria", Universidad de Cantabria, Parque Científico y Tecnológico de Cantabria, C/Isabel Torres, 15, 39011 Santander, Spain

[3,c]Museo Nacional de Ciencias Naturales (CSIC), José Gutiérrez Abascal 2, 28006 Madrid, Spain

*Corresponding author.

Abstract

Species distribution modelling may support ecologists in conservation decision-making. However, the applicability of management recommendations depends on the uncertainty associated to the modelling process. A key source of uncertainty is the underspecificity of the research question. Modelling specific questions is straightforward since they drive clearly the methodological choices about input data and model building. Nevertheless, when the research questions remain underspecific, modellers must choose among a wide spectrum of choices, with each decision sequence driving to a different outcome that explain partially the target question. We show how the underspecificity associated to a general research question about Great Bustard breeding success at geographic scale drives to multiple decision choices, leads to a variety of model outcomes and hampers the identification of specific conservation actions. We ran generalised linear models using multi-model inference on a set of databases built according to specific sequences of methodological choices. Then, we evaluated variations in model performance, complexity (parsimony) and nature of predictors, as well as averaged model predictions and spatial congruence among model outputs. Deviance and parsimony varied widely (11.46% to 83.33% and 7 to 18, respectively), as did model averaged mean predictions in occupied areas, contributing predictors and spatial congruence among outputs ($r_{Pearson} = 0.44 \pm 0.23$ for models calibrated in occupied areas; 0.48 ± 0.06 for models calibrated in potential/accessible areas). We recommend to carefully fix research questions and associated methodological options through collaborative working frameworks to conceptualize modelling approaches and, thus, to mitigate problems arising from underspecificity and other forms of uncertainty in conservation applications.

Keywords: Breeding success; Breeding site selection; Great Bustard; Species distribution modelling

1 Introduction

Species distribution models (SDM) are correlative approaches that allow for the estimation of species' ecological requirements at multiple scales in the framework of the ecological niche theory (Austin, 2002, 2007; Kearney, 2006; Peterson, 2006; Jiménez-Valverde et al., 2008; Warren, 2012). When SDM are based on static limiting conditions (scenopoetic variables) that may control species' ecophysiology and drive its occurrence at large scale, these models can be interpreted as describing the Grinnellian niche. When SDM are based on biotic interactions and resource consumer dynamics (bionomic variables) that determine species abundance and reproductive rates operating at more detailed scales, they can be interpreted as describing the Eltonian niche (Soberon, 2007; Tingley et al., 2009; Peterson et al., 2012; Alvarez-Martínez et al., 2015).

SDM are, therefore, potentially useful tools for ecologists and land managers dealing with processes of decision-making addressed to biological conservation actions (Suárez-Seoane et al., 2002; Franklin, 2010a). However, their applicability decreases if they fail to describe the ecological system under study across different scenarios (Elith et al., 2002; Dawson et al., 2011; Guisan et al., 2013). This problem is inherently linked to the uncertain nature of ecological processes, but it might also depend on the methods and decision-making applied by researchers. In this sense, scientists must cope with the uncertainties derived from incompatibilities between ecological background, input data and statistical methods (Austin, 2007) that can permeate model results and management recommendations. Authors such as Harwood and Stokes (2003) and Ascough II et al. (2008) have argued that the failure of ecologists to

evaluate accurately the uncertainties associated with their advice (e.g., integrated assessment models, optimization algorithms and multi-criteria decision analysis tools) diminishes their influence in decision-making. Therefore, it is compulsory to understand the practical consequences of these uncertainties, which may be exacerbated when the system under study is complex and changing (Álvarez-Martínez et al., 2010; Polasky et al., 2011).

Random uncertainty refers to the inherent variability of a given system and is typically named as variability, irreducible uncertainty, inherent uncertainty or stochastic uncertainty. Besides this uncertainty due to chance, three main human dependant sources of uncertainty can affect soundly SDM predictions: epistemic, linguistic and human decision-making (Regan et al., 2002; Elith et al., 2002; Kujala et al., 2013). Epistemic uncertainty is due to a lack of knowledge about the state of the system that is being modelled. It is associated to technical decisions, as well as systematic and measurement errors in data collection, such as those related to location (e.g., spatial accuracy and representativeness of the species' range requirements; Thuiller et al., 2004; Randin et al., 2006; Menke et al., 2009), shape (e.g., areal unit for which data are collected, which may lead to the modifiable areal unit problem MAUP; Openshaw and Taylor, 1981; Wong, 2004) or nature of input data (e.g., species data can be continuous, counts or binary and predictor values can be collected from different sources; Synes and Osborne, 2011). Linguistic uncertainty originates when language used for stating research questions or defining terms is underspecific, vague, ambiguous or context-dependent. Underspecificity arises when the research question is excessively general and portraits a lack of clarity. For example, the question "which environmental factors determine temporal changes in biological fitness?" is underspecific because it does not clarify whether we are interested in the assessment of factors behind inter or interannual variability in fitness, neither if we are looking for the factors that determine the mean or the variation of fitness values for a particular time span, the biotic or abiotic factors driving the temporal changes and so on. On the opposite, "is annual maximum temperature behind interannual variation of mean breeding success?" or "does spring cumulative rainfall determine weekly variations in breeding success during the reproductive season?" are both specific research questions that can be nested within the former underspecific question. Vagueness refers to the possibility of borderline cases due to categorical classifications of data (Regan et al., 2002). It arises when terms are defined using arbitrarily sharp boundaries (e.g., "high fitness", "optimal habitat" or "viable population") that may alter drastically the output (Bull et al., 2016). Ambiguity originates because some words may have more than one meaning (e.g., "fitness" can be defined either as the variation in reproductive success or as the genotype's rate of increase in future generations; Michod, 2000). Context-dependence is problematic when the framework of the question at hand has not been correctly specified, which may affect its interpretation. The lack of a coherent understanding of the context under which to answer ecological questions may be associated with large variability in the research predictions (Lajoie and Vellend, 2015). For instance, different answers should be expected if we explore the fitness variations in Great Bustard populations in Spain or Russia due to different biotic (e.g. interspecific relationships) and abiotic (e.g. climate constraints) contexts. Finally, uncertainty derived from human decision-making may arises from human beliefs, values, preferences, choices and actions, as it is the case of scale preferences (e.g. spatial resolution) or subjective choice of predictors to be included in a model. However, the best recognized type is subjective judgement, which is frequently associated to scenario planning or interpretation of model results. This problem is especially relevant when data are scarce and error prone. The standard way of dealing with it is to assign a level of certainty to the target event in the form of a subjective probability (Regan et al., 2002). For instance, we might assign a probability of 0.8 to the event "the mean breeding success of Great Bustard in Spain during the last five years was 15%". There are different techniques within decision science that can help to address this area of uncertainty. Among them, structured decision making (Gregory and Keeney, 2002) and adaptive management (Runge, 2011) are particularly relevant for applying formal decision-analysis tools in natural resource management decisions.

In this context, the underspecificity associated to the statement of the research question emerges as a critical and under evaluated issue that can be particularly relevant in conservation applications. Despite wide assessment of epistemic uncertainty in SDM approaches (e.g., Elith et al., 2002; Thuiller, 2003; Pearson et al., 2006; Convertino et al., 2012), the role of underspecificity still remains underexplored, as it occurs with other forms of linguistic and human decision-making uncertainty. In fact, according to Kujala et al. (2013), from the set of papers considering uncertainty that were published during the period 1945-2011 with an ecological scope, only 1.5% mentioned specifically the underspecificity, as did the 1% of the papers with a conservation scope. Therefore, underspecificity should be deeper explored, as it may have serious implications in model interpretation and subsequent applicability. The specific definition of the research question is crucial since it frames the problem and drives the methodological choices to be done, structuring the modelling approach and the nature of the requested data (Kuhnert et al., 2010). Modelling specific questions, with a low level of underspecificity, is straightforward since they drive clearly the methodological choices to be done through the development of the modelling approach. However, research questions might be underspecific due to a lack of theoretical knowledge or because scientists do not consider explicitly the complexity and dynamics of the ecological systems under evaluation. General questions, with a high level of underspecificity, are associated to broader environmental responses and result much more difficult to capture in a single model. This is problematic since scientists are required to make specific decisions about input data and methods, which leaves open a wide spectrum of possibilities and, therefore, underspecificity issues. Each sequence of decisions will determine the nature of the outputs, the model performance and the chance of error given the different types of data and approaches, as well as the different levels of risk aversion (Barry and Elith, 2006). Practitioners must be aware of the implications of their decisions since each possible model outcome would explain only partially the target question, with a wide array of different outputs.

In this work, we aim to show how the underspecificity associated to a general research question drives to multiple modelling decisions, leads to a variety of model outcomes and, therefore, hampers the identification of specific management choices. We formulate an underspecific research question about Great Bustard (*Otis tarda*) breeding success in Spain during the last two decades (period 1987-2010) and evaluate differences in explained deviance, averaged mean predicted values across distributional areas, model complexity, nature of predictors (scenopoetic and bionomic) and spatial prediction patterns obtained when applying different sequences of methodological choices (with different ecological meaning) in a SDM framework. The evaluation of this ecological trait remains a challenge because of complex trade-offs between individual life traits of Great Bustards (quality, age or experience; Lescroël et

al., 2009), social relationships (e.g., reproductive skew in social species; Johnstone, 2000; Ryder et al., 2009) and environmental constraints of distribution (widely changing across space and time; Barbraud and Weimerskirch, 2005). We intend to draw attention to the risks of generalising the outcomes obtained by applying specific methodological choices when modelling underspecific questions, highlighting the importance of carefully specifying the ecological question that one aims to disentangle.

~~2.~~2 Methods

~~2.1.~~2.1 Data on Great Bustard breeding success in Spain

Great Bustard is a globally threatened lekking bird species with a population severely fragmented throughout the Paleartic (BirdLife International, 2017). Almost 60–70% of the world population of the species is located in the Iberian Peninsula (Alonso and Palacín, 2010). To calculate Great Bustard breeding success in Spain, we used a 24-year series (1987–2010) of population censuses carried out across its breeding range. Censuses were conducted in September, when families can easily be detected (Martin et al., 2007). The geographic position of all located flocks of females with or without chicks was incorporated into a GIS database. Flocks ~~where~~were assigned to the nearest lek (location of the largest male flock in late winter and early spring) identified in Spain (Alonso et al., 2011)~~, by~~ using a minimum distance criterion refined by considering discontinuities and physical barriers (roads, rivers, high elevations), as well as information derived from long-term radio-tracking studies (Palacín et al. 2011). Leks are spatial and functional units to which adult males and females show a marked site-fidelity throughout the reproductive season (Alonso et al., 2000; Magaña et al., 2011). All flocks assigned to the same lek and year formed a reproductive group (RG). We accounted for 208 RGs. Breeding success was estimated annually for each RG as young productivity (i.e., ratio of the number of chicks to females). Productivity ranged from 0 to 100%. RGs showing values of productivity higher than 100% (small groups sampled only once or twice ~~and~~ consisting of one or two females and their chicks mainly located in peripheral areas) or with a low detectability (the number of females in autumn was lower than 30% of those counted in spring) were discarded. These values were incorporated in a multi-temporal database. We used annual productivity values for each of the 208 RG to define temporally averaged productivity patterns across Spain for the study period (see Section 2.3). Alvarez-Martínez et al. (2015) and Suárez-Seoane et al. (2017) provide more technical details on the GIS database building.

~~2.2.~~2.2 Environmental variables

We created a comprehensive pool of environmental GIS predictors related to Great Bustard breeding success (Morales et al., 2002; Alonso et al., 2004; Pinto et al., 2005; Martínez, 2008; Alvarez-Martínez et al., 2015; Suárez-Seoane et al., 2017) (Table 1). We chose a wide spectrum of potentially explanatory variables, both scenopoetic and bionomic, to allow for flexibility when fitting the models characterizing the ecological niche of the species (Alvarez-Martínez et al., 2015). The risk of over-fitting when using large pools of variables is arguable (Knape and de Valpine, 2011), but increasing the number of covariates improves the chance of including the most relevant predictors. Variables described topography and ~~geographic~~ position, climate, primary production, landscape structure and human disturbances. The scale of original data ranged from 1:25000 to 1:200000 and the pixel size from 200 m to 1 km. ~~Therefore, all~~All data were rescaled to the same spatial resolution, matching the broader pixel size of 1 km.

Table 1 List of environmental predictors. Spring (SP) includes March, April and May; summer (SU): June, July and August; and autumn-winter (WI): September to March. CV is the coefficient of variation $= \left[(SD/\overline{x}) * 100 \right]$, where SD is the standard deviation.

alt-text: Table 1

	Variable	Code	Units	Source	Pixel (m)	Scale
Topography, geographic position	Southness (south-north downslope, ranging from +1 to –1)	SOUTH ~~EAST~~	Adimensional	DEM 25 m (CNIG, 2012)	200	1:200000
	Eastness (east-west downslope, ranging from +1 to –1)	EAST	Adimensional			
	Curvature (second derivative of the surface; –1 in valleys, +1 in ridges)	CURV	Adimensional			
	Roughness (standard deviation of the slope)	RUG	Degrees			
	Mean slope	SLOSP	0–255	GIS database of agricultural plots (SIGPAC, 2012)	1000	1:10000
	CV of slope	SLOSPcv	0–255			
	Longitude	LONG	m	ArcGIS desktop 10.5 (ESRI, 2016)	100	1:50000

Climate	Mean temperature: summer (CV)	TMEvSU	Dimensionless			
	Mean temperature: autumn-winter (CV)	TMEvWI	Dimensionless			
	Mean temperature: thermal amplitude of spring (range)	TMErSP	°C			
	Minimum temperature: summer (mean)	TMImSU	°C			
	Minimum temperature: autumn-winter (CV)	TMIvWI	Dimensionless			
	Minimum temperature: annual thermal amplitude (range)	TMIrAN	°C			
	Minimum temperature: thermal amplitude of summer (range)	TMIrSU	°C	Climatic Map of the Iberian Peninsula (Ninyerola et al., 2005, 2007) (Original data: monthly, period 1950–99)	200	1:200000
	Maximum temperature: summer (CV)	TMAvSU	Dimensionless			
	Maximum temperature: autumn-winter (CV)	TMAvWI	Dimensionless			
	Maximum temperature: thermal amplitude of summer (range)	TMArSU	°C			
	Rainfall: spring (mean)	PPmSP	mm			
	Rainfall: spring (CV)	PPvSP	Dimensionless			
	Rainfall: autumn-winter (CV)	PPvWI	Dimensionless			
	Rainfall amplitude: annual (range)	PPrAN	mm			
	Rainfall amplitude: spring (range)	PPrSP	mm			
	Absolute evapotranspiration (AE): autumn-winter (mean)	AEVPmWI	[(mm)/day]			
	Absolute evapotranspiration: summer CV	AEVPvSU	Dimensionless			
	Absolute evapotranspiration: autumn-winter CV	AEVPvWI	Dimensionless			
	Relative evapotranspiration (soil moisture): autumn-winter (mean)	REVPmWI	[%]			
	Relative evapotranspiration (soil moisture): annual (CV)	REVPvAN	Dimensionless			
	Relative evapotranspiration (soil moisture): autumn-winter (CV)	REVPvWI	Dimensionless	METEOSAT (Original data: 10-daily basis, period 1988–2010)	5000	
	Net Radiation: spring (CV)	NRADvSP	Dimensionless			
	Net Radiation: summer (CV)	NRADvSU	Dimensionless			
	Maximum difference in annual mean absolute evapotranspiration	MD_EmAN	[(mm)/day]			
	Maximum difference in annual mean net radiation	MD_NmAN	W/m^2			
	Maximum difference in annual CV relative evapotranspiration	MD_RvAN	Dimensionless			
	Maximum difference in annual CV net radiation	MD_NvAN	Dimensionless			
Primary production	NDVI: spring (mean)	NDVImSP	Dimensionless			
	NDVI: summer (mean)	NDVImSU	Dimensionless			
	NDVI: annual (CV)	NDVIvAN	Dimensionless	NOAA-AVHRR (monthly data, period 1987–2010)	1000	

	NDVI: spring (CV)	NDVIvSP	Dimensionless			
	NDVI: summer (CV)	NDVIvSU	Dimensionless			
	NDVI: autumn-winter (CV)	NDVIvWI	Dimensionless			
Landscape structure	Plot Fragmentation Index (no. plots divided by their average area)	PFI	%	GIS database of agricultural plots (SIGPAC, 2012)	1000-1	1:10000
	CV of plot perimeter	PERIMv	Dimensionless			
	Sum of plot perimeter	PERIMsum	km			
	Maximum perimeter of plots	PERIMmax	km			
	Arable land (% occupation)	ARLAND	%	Land Cover Information System of Spain (SIOSE, 2012)	200	1:25000
	Other crops (% occupation): olive and vineyards	OLIVIN2	%			
	Irrigated lands (% occupation)	IRRIG	%			
	Dry cereal crops (% occupation)	CEREAL	%			
	Dry olives and vineyards (% occupation)	OLIVIN	%			
Human disturbances	Distance to paved roads of communication (roads and highways)	DISTRO	m	BCN200 (CNIG, 2012)	200	1:200000
	Distance to human infrastructures (villages, cities, urban sprawl, industries)	DISTED	m	SIOSE (2012)	200	1:25000
	Human footprint	HUMFP	Dimensionless	NASA (2012)	1000	1:200000
	Distance to SCI (Sites of Community Importance; "Natura 2000" Network)	DISTSCI	m	"Natura 2000" Ecological Network (MAGRAMA, 2012)	200	1:200000
	Distance to SPAs (Special Protection Areas; "Natura 2000" Network)	DISTSPA	m			
	Distance to IBAs (Important Bird Areas; BirdLife International)	DISTIBA	m			

To evaluate the effect of climate, we developed two datasets: (i) Temperature (maximum, mean and minimum) and rainfall, extracted from the Climatic Map of the Iberian Peninsula (Ninyerola et al., 2005, 2007). These measures were achieved by interpolating (multiple regression in combination with a residual correction method) the observed monthly averaged ground-data collected from meteorological stations distributed all across the Iberian Peninsula for the period 1950-1999 at 200 m resolution. (ii) Net radiation and both absolute and relative evapotranspiration derived from a combination of METEOSAT satellite data (period 1988-2010, 10-day products) at ca. 5 km resolution (visible and infrared channels) with ground-truth information. Data came from the European Energy and Water Balance Monitoring System (EWBMS; Rosema et al., 2001). See Suárez-Seoane et al. (2004) for a more complete explanation of these variables. The above-ground net primary production was estimated using the Normalized Difference Vegetation Index (NDVI; Pettorelli et al., 2011; Bro-Jørgensen et al., 2008) obtained from a temporal series of NOAA-AVHRR satellite imagery (period 1987-2010) acquired from the Spanish National Research Council (CSIC) databases. Data on climate and primary production were annually and seasonally averaged (mean, variation coefficient and range) to assess their effect on breeding success during critical periods of the year. Landscape structure was quantified from the Spanish Geographic Information System for Agricultural Plots (SIGPAC, 2012) and the Land Cover Information System of Spain (SIOSE, 2012). SIGPAC was used to create spatial indices of landscape structure on the basis of the property limits in 1x1km grids (authors' unpublished data). SIOSE was handled to obtain single land covers types of interest. Independent layers for each land cover were achieved by merging those polygons where target land covers were dominant (over than 70% of coverage). Human disturbances were quantified as the Euclidean-distance from each pixel to the nearest town, road or other infrastructure; human footprint (NASA, 2012); and, land protection status. All GIS analyses were done in ArcGIS 10.5 (ESRI, 2016).

2.3.2.3 Methodological choices and databases

Six methodological dilemmas emerged when building both the response variables and the environmental predictors from the original multi-temporal dataset. Each dilemma implied alternative methodological choices with different ecological meaning (Figure 1). (1) Dilemma 1: Data could be collected across different spatial extents, i.e. within occupied nesting areas or within potential/accessible breeding areas (Suárez-Seoane et al., 2002; Alvarez-Martínez

et al., 2015). (2) Dilemma 2: Dependent variables could be made of continuous or binary values. (3) Dilemma 3: In the case of continuous dependent variables, annual productivity data could be averaged across the temporal series using either the mean or the range of the values. (4) Dilemma 4: Regarding binary dependent variables, productivity values could be split into "high" or "positive" vs. "low" or "null". "High productivity" corresponds to those RGs showing consistently the best values of fitness (mean productivity higher than the averaged mean value for all RGs across the study period (i.e., 15%). "Positive productivity" allocates to RGs with a productivity value greater than "0" (i.e., birds successfully bred, independent of the number of chicks raised). (5) Dilemma 5: On their hand, "low productivity" corresponds to RGs where productivity is positive, but lower than the averaged mean value of all RGs across the study period (i.e., marginal RGs in terms of fitness). "Null productivity" corresponds to random points located in non-occupied sites within the potential/accessible breeding distribution area where it is known that birds were not able to breed (Suárez-Seoane et al., 2002). (6) Dilemma 6: Environmental predictors could be gathered using points or patches. Points were spatially assigned to the point location of the female flock with chicks (isolated family or flock of females including at least one family) closest to the "centroid" of all female flocks, with or without chicks, in a RG (FamCRG; Alvarez-Martínez et al., 2014). The 1 km-pixels holding FamCRG locations were overlapped with the explanatory variables to extract their values. Patches consisted of reproductive areas (RA) defined by applying 1000-m buffers around all flocks included in the same RG to simulate family home ranges during the early rearing period in summer (unpublished data from intensive radio-tracking, Martín et al., 2007), around all flocks included in the same RG. Patches could be continuous or discontinuous, depending on whether between-flock distances were smaller or higher than 1 km. Explanatory variables' values were extracted for all pixels conforming each patch and their values averaged (mean) for the whole patch. Both points and buffers were recalculated for each year during the study period. The time-averaged productivity values were associated with the family flock closest to the centroid of all "centroids" defined for successive years, or to the corresponding RAs. For "null productivity" cases in potential/accessible sites, we applied a 2000-m buffers around each point, since this is approximately in correspondence with the mean area of all reproductive areas defined within occupied nesting sites (1296 and 1361 ha, respectively). After applying the cascade of methodological dilemmas and choices, we obtained 10 databases (Figure. 2) that were used for model calibration.
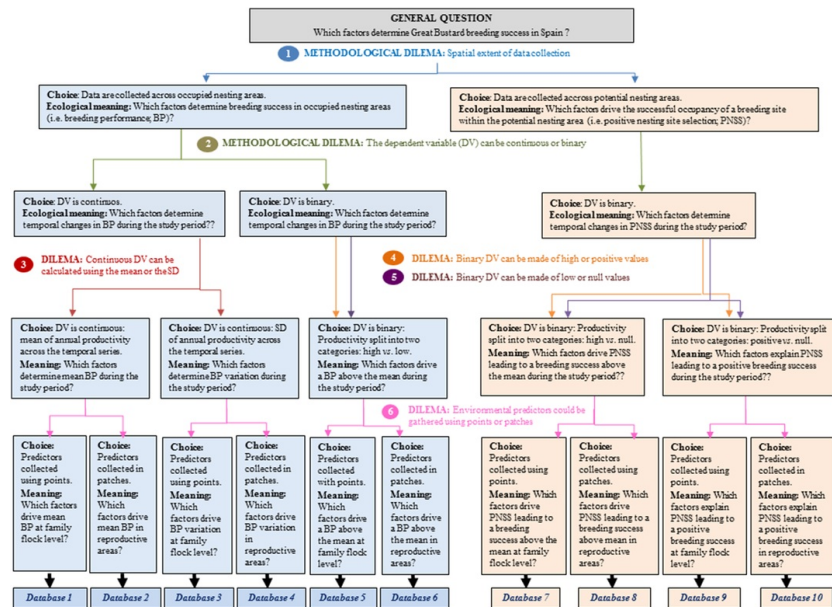

Figure 1.Fig. 1 Methodological dilemmas and specific choices (each with a different ecological meaning) emerging when modelling the underspecific research question related to Great Bustard breeding success.
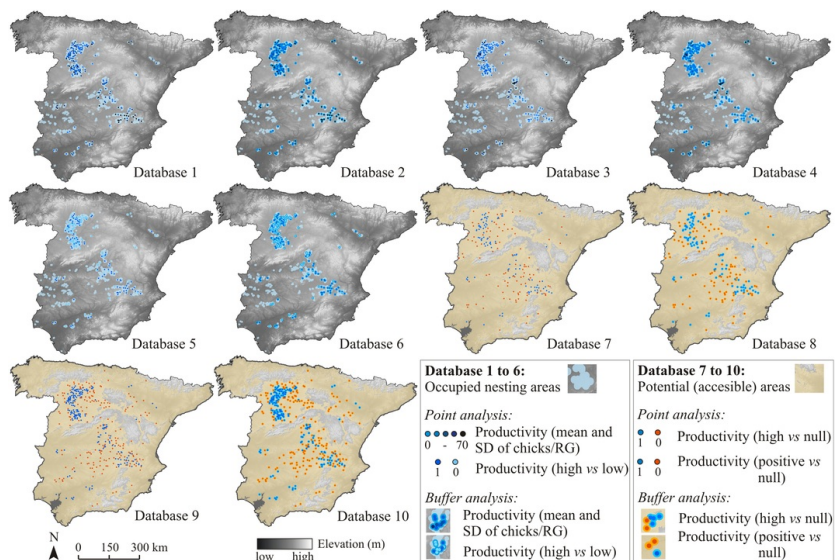
alt-text: Fig. 1

alt-text: Fig. 2

2.4.2.4 SDM building: multicollinearity, step-AIC, multi-model inference and spatial outputs

We independently modelled each of the 10 databases (Figuress. 1 and 2). To reduce multicollinearity problems that may lead to variance inflation and parameter bias (Freckleton, 2011), we checked Pearson's bivariate correlations among all predictors for each database. Each variable from a highly correlated pair ($r_{Pearson} > 0.7$; Randin et al., 2006) was retained/rejected according to the ratio model deviance/null deviance estimated for that variable in univariate generalised linear models (GLM). Additionally, we eliminated predictors showing a variance inflation factor (VIF) higher than 4. There is no formal VIF threshold, but a value of 10 is commonly used as an indicator of severe multicolinearity (Neter et al., 1990; Graham, 2003; Zuur et al., 2010).

After filtering the effect of correlation, we performed stepwise selection on the basis of AIC (Akaike's Information Criterion) to reduce the number of predictors in each database (Venables and Ripley, 2002). Some statisticians strongly discourage automatic stepwise methods with many potential predictors since they may increase the chance of spurious correlations (Mundry and Numm, 2009). However, its disciplined application is considered appropriate when data exploration is clearly separated from hypothesis testing (Pinheiro and Bates, 2000; Zuur et al., 2010).

Then, we implemented a multi-model inference analysis on the remaining variables for each dataset to select the best subset (i.e., that with the smallest AIC value indicating the most parsimonious candidate model), among all possible ones (Burnham et al., 2011). Models with Δ-AIC ≤ 2 were considered substantially supported by the data and similar to the best model in their empirical support (Burnham and Anderson, 2002; Heinze et al., 2018). Combining all possible candidate models, we estimated the averaged standardized coefficients, significance and relative importance of each predictor. The predictor relative importance is the sum of the Akaike weights of all models in the subset where the predictor is present. The value of the summed Akaike weight ranges from 0 (the predictor appears only in the most unlikely models) to 1 (it appears in all the best models) (Burnham and Anderson, 2002; Symonds and Moussalli, 2011). GLMs were built using either a Gamma distribution with log link, when the response variable was continuous, or a Binomial distribution with logit link, when it was binomial. Models were spatially projected within a GIS framework to generate maps of breeding habitat suitability.

Differences across the 10 model outputs were evaluated in terms of: i) Deviance explained by the best subset of variables against a null model. ii) 10 cross-folder validation within each of the 10 datasets, which allowed for calculating a value of dDelta (raw cross-validation estimate of prediction error) and delta adjusted (adjusted cross-validation estimate, designed to compensate for the bias introduced by not using leave-one-out cross-validation) for each case values derived from 10 cross-folder validation analysis (Davison and Hinkley, 1997). (iii) Averaged predicted values (mean) in each distributional range of Great bustard in Spain (occupied, potential and accessible areas; Alvarez-Martínez et al., 2015), with predicted values extracted using a sample of 10,000 independent points randomly distributed. (iv) Model complexity (i.e., parsimony). (v) Nature of the predictors included in the model. vi) Broad spatial patterns of model predictions (visual inspection). Additionally, we assessed the spatial agreement between the predictions achieved by models calibrated in occupied (databases 1–6) and potential/accessible (databases 7–10) areas in two ways. First, we

combined the spatial predictions obtained for each of these areas for visual inspection. Each continuous output was thresholded by applying 75th percentile on the prediction values to produce Boolean maps of breeding suitability. These Boolean images were overlaid and the number of times each pixel was defined as suitable recorded to produce integrated maps of suitability for occupied and potential/accessible areas. Second, we calculated bivariate Pearson correlations among the spatial predictions achieved by models calibrated in both each areas. Correlation analyses were carried out using two separate samples of 10,000 random points with a minimum allowed distance of 300 m, each distributed across either occupied or potential/accessible area.

Analyses were done with the packages MASS (Venables and Ripley, 2002) and MuMIn (Barton, 2018) from R 3.4.3 statistical software (R Development Core Team, 2017) and ArcGIS 10.5 (ESRI, 2016). See Appendix 1 for checking the R script.

3.3 Results

Methodological decisions taken when building the input databases (Figure 1) had clear implications in model outcomes (Table 2). The deviance explained by the best subset of variables varied substantially across the ten models, ranging from 11.46% to 83.33% (46.01 ± 26.12). Delta adjusted strongly differed across databases, with values ranging from 198.82 to 281.56 in Gamma models (databases 1–4) and from 0.05 to 0.23 in binary models (databases 5 to 10). The averaged mean of the predicted values also changed across databases and distributional ranges (see Appendix 2). For the case of models calibrated in occupied breeding areas with continuous dependent variables (databases 1 to 4), model predictions were the most consistent for occupied areas, with values narrowly ranging from 0.05–0.06 (i.e., averaged mean value of 5–6 chicks per 100 females). On their hand, models based on binary dependent variables (databases 5 to 10), calibrated in either occupied or potential/accessible areas, achieved an averaged mean predicted value ranging from 0.48 to 0.72 in occupied areas, with values significantly much lower in both potential (0.24–0.30) and accessible (0.10–0.11) areas. The complexity (parsimony) and nature of the most relevant predictors changed markedly across the ten outputs. Parsimony ranged from 7 to 18 (12 ± 4.55), being summer mean NDVI, distance to IBAs, mean relative evapotranspiration in autumn-winter and distance to roads the variables most regularly included in the models (in 50% or even more). The visual inspection of the ten spatial outputs highlighted important differences regarding breeding suitable areas, as shown in Appendix 3. For the case of the spatial agreement analysis, we found a $r_{Pearson}$ value (mean and standard deviation) of 0.44 ± 0.23 for models calibrated in occupied areas and 0.48 ± 0.06 for models calibrated in potential/accessible areas (Figure 3 and Appendix 4).

Table 2 Results of multi-model averaging (see Table 1 for variable codes). Each cell shows the sign, model-averaged standardised standardized coefficients (β) multiplied by 100, significance (***$p < 0.001$, **$p < 0.01$; *$p < 0.05$) and, in parenthesis, relative importance of each variable included in n subsets of models having Δ_i (AICbest-AIC$_i$) ≤ 2. Deviance explained in relation to the null model was calculated for the best subset of variables, as well as delta and delta adjusted for 10 cross-folder validation models (10-cv).

| Variables		Database1	Database2	Database3	Database4	Database5	Database6	Database7	Database8	Database9	Database10
Topography, geographic location	SOUTH					-92.32**(1)					
	EAST										+184.38*(1)
	CURV					+51.71 (0.50)					
	RUG								-413.60*(1)		-645.78*(1)
	SLOSP			-0.68 (0.06)				+381.63*(1)		-170.20 (0.72)	
	SLOSPcv									-122.08** (1)	
	LONG					93.62*(0.83)				-29.83 (0.04)	444.74***(1)
Climate	TMEvSU									-73.37 (0.35)	-386.45**(1)
	TMEvWI						+69.99 (0.37)				
	TMErSP										-99.51 (0.07)
	TMImSU		+2.48***(1)	+0.76 (0.56)							
	TMIvWI										
	TMIrAN	+1.39**(1)					+194.18**(1)				
	TMIrSU							-295.58**(1)	-358.84**(1)		
	TMAvSU								-153.02 (0.52)		
	TMAvWI		+1.41*(1)								
	TMArSU					-115.84 (0.83)				-70.33 (0.27)	
	PPmSP									+102.34 (0.18)	+217.80 (0.59)
	PPvSP								+208.68 (0.63)		
	PPvWI			+0.80*(0.82)				-199.08*(1)	-244.27 (1)		
	PPrAN							402.72**(1)	+520.91**(1)		
	PPrSP	-0.72 (1)									
	AEVPmWI	-							-530.67*(1)		
	AEVPvSU	0.85*(0.96)							-323.75*(1)		
	AEVPvWI							+147.51 (0.67)			-130.17 (0.34)
	REVPmWI	+0.71*(1)	+1.00**(1)	-0.75 (0.54)						-110.58*(1)	-541.29***(1)
	REVPvAN							+88.43 (0.03)			
	REVPvWI					-95.42*(1)	-64.43 (0.47)				
	NRADvSP				-1.41*(1)						
	NRADvSU				-0.10* (1)						+170.52 (0.14)
	MD_EmAN				-1.26*(0.91)						
	MD_NmAN			+0.92*(0.91)	+1.05 (0.47)			+387.48**(1)		+305.19***(1)	
	MD_RvAN						+79.45 (0.86)				
	MD_NvAN					-116.41*(1)					
Primary production	NDVImSP	-0.63 (0.76)	-0.91**(1)								
	NDVImSU	-0.55 (0.45)		-2.01***(1)	-2.16***(1)	-81.57*(0.98)		-519.40***(1)	-512.74**(1)		
	NDVIvAN					-57.27 (0.61)					
	NDVIvSP								-140.35 (0.59)		-246.93*(1)
	NDVIvSU									-63.67 (0.08)	
	NDVIvWI									+232.56***(1)	+446.27*(1)
	PFI									-2624.61**(1)	
Landscape structure	PERIMv			+1.07**(0.75)							
	PERIMsum			-1.90***(1)							
	PERIMmax			+1.06*(0.29)				+143.57 (0.25)	-364.07**(1) / +378.04***(1)		+316.69*(1)
	ARLAND							+395.17**(1)	+624.84***(1)	+382.93***(1)	+1134.13***(1)
	OLIVIN2			+0.61 (0.36)						+97.69*(0.96)	+173.90 (0.87)
	IRRIG			+1.06*(1)	+0.65 (0.49)			+309.92***(1)			
	CEREAL							+187.43*(1)		+82.37 (0.74)	+1242.23***(1)
	OLIVIN	-0.89**(1)				-103.52 (0.70)		-431.99 (0.3)			
Human disturbances	DISTRO	+0.44 (0.45)	+0.53 (0.60)			+46.65 (0.47)		+170.21*(1)		+94.80*(1)	
	DISTED		-0.52 (0.57)	-0.64 (0.55)				+112.73 (0.48)			
	HUMFP						-56.43 (0.67)			-98.29 (0.90)	-358.45**(1)
	DISTSCI			-2.10***(1)	-1.31***(1)					+70.24 (0.62)	-106.12 (0.24)
	DISTSPA			+1.59***(1)	+0.87(0.89)				-93.41 (0.36)		
	DISTIBA	+0.52 (0.59)	+0.80*(1)			+78.27 (0.82)	+75.43*(0.93)			-246.08***(1)	-590.19***(1)
Nb. predictors		9	7	15	7	10	7	16	13	18	18
Deviance (%)		28.52	24.72	46.92	35.92	16.90	11.46	72.06	71.23	69.07	83.33
Delta (10-cv)		205.91	237.80	294.52	203.51	0.25	0.23	0.10	0.08	0.08	0.05
Delta adjusted (10-cv)		203.05	234.05	281.56	198.83	0.22	0.23	0.10	0.08	0.08	0.05

alt-text: Table 2
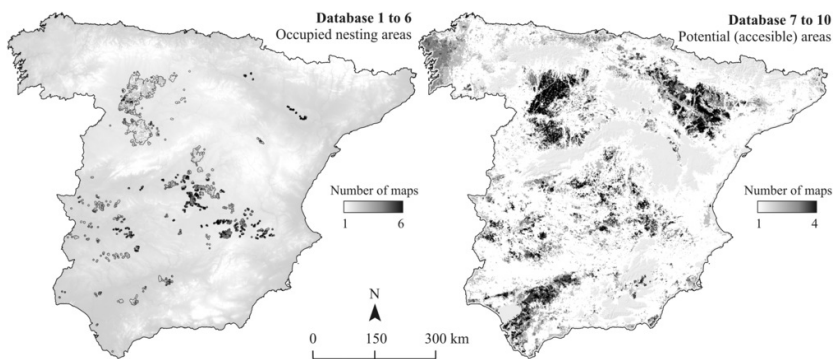
Figure 3. Fig. 3 Spatial agreement between Boolean model outputs in occupied and potential/accessible areas. The threshold for map reclassification was the 75th percentile of the prediction values. The maximum value was 6 and 4, respectively, for occupied and potential/accessible areas.

alt-text: Fig. 3

The first methodological dilemma (i.e., spatial extent of the area of data collection) had a major impact on model outputs. Models of breeding performance, that were calibrated in occupied areas (databases 1 to 6), were more

parsimonious and explained less deviance than models of nesting site selection, calibrated in potential/accessible nesting areas (databases 7 to 10). The last group of models was more consistently associated with typical scenopoetic variables, such as land cover, topography and climate than the former.

According to the second dilemma, models calibrated in occupied areas with continuous values of breeding success explained a higher deviance (24.72%–46.92%) with a lower parsimony than those calibrated with binary values (11.46%–16.90%). In potential/accessible areas, where all models were calibrated with binary dependent variables, explained deviance was the highest, with values over 69.07%.

For the third dilemma (i.e., continuous dependent variable calculated using the mean or the range of temporal breeding success data), models of mean breeding success (databases 1 and 2) explained less deviance with lower levels of complexity (higher parsimony) than those based on breeding success variation range (databases 3 and 4). Scenopoetic and bionomic variables informing on climate, primary production and human disturbances played an equivalent role when modelling the two pairs of databases, but topography and landscape structure were more relevant when modelling the variation of breeding success.

Concerning the fourth (i.e., binary dependent variable can be made of high or positive values of breeding success) and fifth (i.e., binary dependent variable can be made of low or null values) dilemmas, models calibrated with databases 5 and 6 (high vs. low breeding success; breeding performance) explained less deviance and were more parsimonious than those calibrated with databases 7 and 8 (high vs. null breeding success; nesting success). The last pair of models included more variables, with climate and landscape structure playing a more important role than the former. Models developed from databases 9 and 10 (positive vs. null breeding success) were the most complex of the whole set (lowest parsimony), showed a strong effect of all predictor families and explained the highest values of deviance.

Fort the sixth dilemma (i.e., environmental predictors can be collected using points or patches), we found that, when modelling reproductive performance (databases 1 to 6), models built at the point scale (family flock level) explained more deviance with a lower parsimony than those built at a patch scale (reproductive area level). However, when modelling positive nesting site selection (databases 7 to 10), we found the opposite trend. The role played by scenopoetic variables, as topography and landscape structure was more important at the point scale.

4.4 Discussion

The results achieved in this research highlight the importance of underspecificity as a source of linguistic uncertainty. Scientific approaches developed in SDM frameworks require the research question at hand to be explicitly defined and described (i.e., low levels of underspecificity) since technical choices about input data are very specific. As we have shown, when the question to be modelled is underspecific, multiple sequences of methodological choices emerge, driving to large differences in explained deviance, complexity and spatial congruence within the pool of results. Consequently, each model output can only be seen as a partial explanation of the general question under study, presenting limited inference and practical applicability for conservation and management applications. The combination of these outputs (multiple responses) into joined spatial products or meta-models (Borsuk et al., 2004) emphasized the need of considering explicitly the uncertainties associated to methodological decisions for a complete understanding of the process under study (Uusitalo et al., 2015).

Typically, different forms of uncertainty are present on every step of ecological analyses (objectives, choice of technical options, influence and constraints of these options, result interpretation), being most of them interrelated (Beven, 2005). In SDM approaches, environmental managers and decision-makers should be aware not only to the inherent randomness and natural variability of the topic under study (aleatory uncertainty), but also to the underspecific use of natural language (linguistic uncertainty) and the methodological decisions applied during the process of building spatial databases (epistemic uncertainty). The important message is that, while aleatory uncertainty is irreducible, epistemic and linguistic uncertainties can be controlled (Uusitalo et al., 2015). In our study, the methodological dilemmas that emerged during the modelling process because underspecificity were directly linked to other forms of linguistic and epistemic uncertainty.

The first dilemma (i.e., the extent of the area of data collection) was a source of epistemic uncertainty that can be linked to the observation and representation of the target system. This is in coincidence with McCarthy (2014), who concluded that some aspects of linguistic uncertainty are typically related to epistemic uncertainty, which is the most recognized form of uncertainty in conservation studies (Burgman et al., 1993). In fact, this dilemma had the highest impact on model outcomes. Despite controversies (e.g. Hirzel and Le Lay, 2008; Jiménez-Valverde et al., 2008), the large differences in model performance generated by this methodological choice might be understood in the framework of the fundamental and realized niche (Pearman et al., 2008). We may postulate that models calibrated in occupied areas (databases 1 to 6) could be interpreted in terms of realized niche ("when and why do species occur?"), so informing on the state of the environment allowing the species to exist indefinitely at its actual distribution in presence of other species. On its hand, models calibrated in potential/accessible areas (databases 7 to 10) would allow for evaluating the species potential distribution in absence of biotic interactions ("when and why do species potentially occur?"; fundamental niche). This connection between SDM and niche concept could only be made assuming the data choices we made in terms of sampling, type and scale of predictors, as well as the predictor selection (scenopoetic vs. bionomic variables) resulting from the applied modelling approach (Soberon and Peterson, 2005; Hirzel and Le Lay, 2008; Franklin, 2010b). Modelling the realized niche was much more challenging (explained deviance: 11.46%–46.92%) than modelling the fundamental niche (deviance: 69.07%–83.33%). A reason may be that models developed in occupied areas are narrowly calibrated across both spatial and environmental dimensions, which implies a high difficulty for identifying local differences in terms of both scenopoetic and bionomic variables. In occupied areas, local differences are not only related to environmental heterogeneity, but also to biotic interactions and dispersal, that tend to occur at short distances and fine-scale (Hirzel and Le Lay, 2008), remaining a challenge

their incorporation in SDM approaches at large scale (Alvarez-Martínez et al., 2015). Oppositely, models calibrated in potential/accessible areas are typically based on scenopoetic variables that vary at larger scales, being easier to find differences in terms of habitat suitability across space (see also Lobo et al., 2010; Suárez-Seoane et al., 2017).

The second dilemma (i.e., assignment of the dependent variable as continuous or binary) was associated with vagueness, a form of linguistic uncertainty. When continuous variables are converted into binary variables, their inherent variability (variable values are only restricted by measure accuracy) is summarised according to unique thresholds (Suárez-Seoane et al., 2017). Dichotomization may imply a huge loss of information about individual differences, as well as spurious statistical significance and overestimation of the size effect in bivariate relationships, loss in the potential to overlook nonlinear relationships and loss of measurement reliability (MacCallum et al., 2002). Therefore, there should not be benefits in this approach if the true outcomes can be observed and the model approach allows for describing the population at hand (Royston et al., 2006; Fedorov et al., 2009). Despite these circumstances, many researchers continue applying dichotomization, probably due to a lack of awareness about the statistical consequences, an absence of appropriate methods of analysis, a belief in the existence of types of individuals or a confidence in dichotomization as a tool to improve reliability (MacCallum et al., 2002). Some particular arguments in favour of dichotomization have been stated (Fedorov et al., 2009): (i) The dichotomized estimator may lead to better results when we intend to estimate a large cumulative distribution function and sample size is large. In that case, the biasedness of model-based estimators will overpower the improvement in variance. (ii) Dichotomization may drive to easiness and simplicity of the reporting results, with the success of such approach depending on the optimal choice of cut points (thresholds). As far as the threshold is deviated from the optimal point, more severe are the consequences of variable dichotomizing. In our study, we have chosen different thresholds for building binary variables (dilemmas 4 and 5) that had a great impact on model outcomes. We have found that models calibrated in occupied areas with continuous values of breeding success explained a higher deviance than those calibrated with binary values. This result can be justified by the difficulty of discriminating between occupied locations classified according to an exigent threshold (high vs. low breeding success) into two categories of breeding success, with all locations exhibiting close environmental features. In potential/accessible areas, where all models were calibrated with binary variables, the explained deviance was the highest. In these models, comparisons are made against potential breeding locations with null productivity (positive/high vs. null values of breeding success), which implies sharp environmental differences among locations that are easier to capture in the models.

For the case of dilemma 3, we found that temporal variation (range values) through the study period was better predicted than the temporal general pattern (mean values) of breeding success. However, even if the former models explained larger amounts of deviance, their parsimony was the lowest. A reason could be that, when birds are looking for a breeding area, their decision is based on many interacting environmental and social "proximate cues", reflecting environmental conditions and resource availability. Some of them are scenopoetic predictors (as topography or landscape structure -land property-) that change slowly in human-dominated landscapes, such as the cereal pseudo-steppes of the Iberian Peninsula where Great Bustards reside; while others are highly dynamic bionomic predictors, as vegetation structure and phenology, food availability or conspecific attraction (Parejo et al., 2006; Osborne et al., 2007; Rieucau, 2011).

In the case of dilemma 6 (effect of gathering predictors at different observation levels), we found that, in general terms, models built at point level (family flock locations) gave better results than those built at patch level (reproductive areas). This result could be related with the kind of selection made by birds of areas suitable for breeding. In this sense, lek-breeding species, as Great Bustards, evaluate their use of the space in the surroundings of potential/accessible breeding locations. However, breeding performance may essentially depend on other factors that are relevant at a more local scale, like differences in phenology or agricultural practices. This effect illustrates how understanding landscape patterns and processes may depend on the level at which observations are made (Suárez-Seoane and Baudry, 2002).

5.5 Recommendations and implications for conservation

Researchers should recognise underspecificity as a main problem affecting the applicability of model predictions in conservation and management (Burgman et al., 2005). The inaccurate statement of the research objectives generates a potential unreliability of the results, implying a loss of effective management and, also, of public trust and confidence (Ascough II et al., 2008). In this sense, we strongly recommend to fix carefully both the research questions and methodological choices through collaborative working frameworks, involving stakeholders and experts altogether from the earliest stages of the ecological modelling process (Redpath et al., 2004; Milner-Gulland and Shea, 2017). Such assisted frameworks allow for conceptualizing modelling approaches, achieving a full understanding of the target question and mitigating many of the subsequent problems that arise from underspecificity and other forms of uncertainty (Mostashari and Sussman, 2005; Jonsson et al., 2007; Martin et al., 2012). Some specific recommendations to reduce the impact of underspecificity in SDM approaches applied to conservation and management are the following:

(1) Research questions should be framed carefully, examining all possible interpretations in the context in which the conservation decisions are going to be made (i.e., nature of that decision and biological, legal and social context in which it occurs; Runge, 2011). By doing so, "the right questions" can be identified and the research objectives and hypothesis can be structured and clarified with low levels of uncertainty (Kuhnert et al., 2010).

(2) Research objectives should be organized hierarchically, from general to specific. General objectives are typically underspecific and can often conflict with others, so they should be clarified or removed according to their degree of relevance. This is a critical step because effective decisions must be based on the assessment of all relevant objectives. Once the general objective is defined, the specific ones should be fully described as free from linguistic uncertainty as possible. If ones

are in conflict with others, they can be either weighted depending on the preferences of the decision-maker or analyzed separately to reveal trade-offs between them (McCarthy, 2014).

(3) Research questions should be unambiguous, comprehensive, direct, operational and understandable (Keeney and Gregory, 2005). In SDM approaches, the consideration of these attributes is essential for taking adequate specific technical decisions that will enhance the value of subsequent analysis.

(4) Each methodological decision taken about input data should be conceptually supported and strongly justified (both in the context of data availability and methodological limitations) when describing the modelling approach, since different sequences of technical choices drive to large differences in model performance, complexity and spatial congruence across the model outputs.

(5) When building dependent variables, decisions about the extent of the area (occupied vs. potential/accessible) of data collection should be specifically pondered, since this methodological choice has a major impact on model performance and ecological interpretation of model outputs.

(6) We advise the use of continuous against binary dependent variables. Dichotomization of continuous variables is not a recommended practice in ecology, due to loss of information and statistical constraints. However, its application can be useful for explaining global ecological patterns when sample size is large and the threshold (cut-off point) is close to the optimal point (Fedorov et al., 2009).

(7) When evaluating temporal patterns of ecological parameters, not only general patterns based on averaged mean values should be explored, but also the variability patterns based on variation metrics of observed values in order to give a full overview of the process under study.

(8) The observation level (point or patch) at which predictors are gathered have a relevant effect on model outcomes, with models built at point level performing generally better. However, this question should be further explored, as we have found differences in model performance depending on the area (occupied vs. potential/accessible) where predictors are collected.

Acknowledgments

Appendix 1.Appendix 1

```r
##########
# stepAIC #
##########

# Libraries
library(MASS)
library(MuMIn)
library (boot)

#Importing data
setwd("X:/x")
tablaPROD    <-read.table("tablaPROD.txt",   header=TRUE,   sep="\t",   na.strings="NA",   dec=".",
strip.white=TRUE)
head(tablaPROD)
dim(tablaPROD)
colnames(tablaPROD)

# Fitting the multiple regression model with Gamma distribution (for continuous data). Introducing
variables sorted according the results of univariate regression models.
# FULL MODEL
fitPROD <- glm (PROD ~ TMIrAN +LONG +PPrSP +LAT +NDVIvAN +NDVImSP +DISTIBA
+AEVPvAN +PPrWI +PPvWI +PERIMmax +TMAvWI +NDVImSU +TMEvWI +PPrSU +DISTED
+DISTLIC +DISTRO +PAR +OLIVIN +DENSINF +AEVPvSU +SOUTH +MD_NmAN +EAST
+PERIMv +CEREAL +MD_RmAN +REVPmWI +HUMFP +SLOSPcv, family = Gamma (log), data =
tablaPROD) # Change family = Gamma (log) for family = Binomial for dichotomic variables
summary(fitPROD)
(1-(fitPROD$deviance/fitPROD$null.deviance))*100
vif(fitPROD)
plot(fitPROD)

# Stepwise Regression (orden GLMvar)
stepPROD <- stepAIC(fitPROD, direction="both")

# Display results
stepPROD$anova
stepPROD$formula

fit.stepPROD <- glm (stepPROD$formula, family = Gamma(log), data = tablaPROD) # Change family =
Gamma (log) for family = Binomial for dichotomic variables
summary(fit.stepPROD)
(1-(fit.stepPROD$deviance/fit.stepPROD$null.deviance))*100
r.squaredLR (fit.stepPROD)
vif(fit.stepPROD)
plot(fit.stepPROD)

#########
# MuMIn #
#########

# All models:
all.fit.stepPROD <- dredge(fit.stepPROD)

# Model averaging:
avg.all.fit.stepPROD <- model.avg(all.fit.stepPROD, beta = TRUE, subset=delta<2)
summary(avg.all.fit.stepPROD)
avg.all.fit.stepPROD$formula

# Final model                                                                          46
def.avg.all.fit.stepPROD <- glm(avg.all.fit.stepPROD$formula, family = Gamma, data = tablaPROD) #
Change family = Gamma (log) for family = Binomial for dichotomic variables
summary(def.avg.all.fit.stepPROD)
(1-(def.avg.all.fit.stepPROD$deviance/def.avg.all.fit.stepPROD$null.deviance))*100
r.squaredLR (def.avg.all.fit.stepPROD)
vif(def.avg.all.fit.stepPROD)


#########
# boot #
#########
# final model after stepAIC and model averaging of significant variables
k    <- 10
kfCV <- cv.glm(data=tablaPROD, glmfit=def.avg.all.fit.stepPROD.1, K=k)
kfCV$delta
```
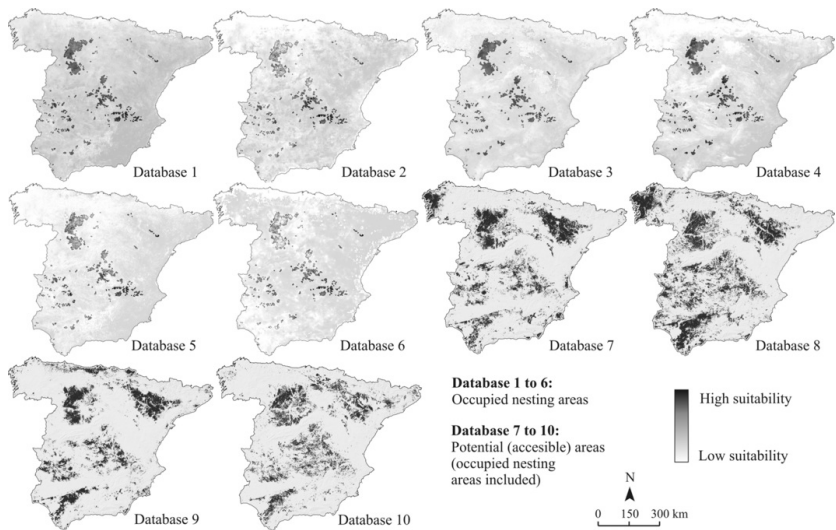
alt-text: Unlabelled Image

Appendix 2.Appendix 2. Averaged mean predicted values achieved for each database and distributional area of Great Bustard in Spain. Predicted values Predictions obtained from databases 1 to 4 (Gamma GLM) can be interpreted as the breeding success value predicted as a function of the considered covariates, with values

productivity values, ranging from 0 (0 chicks per 100 females) to 1 (100 chicks per 100 females). Predicted values achieved from databases 5 to 10 (binomial GLM) are the probability of having either a positive or a high breeding success, with probability values ranging from 0 to 1. Databases 1 to 6 only apply to occupied areas, being not possible to extrapolate model predictions to potential and accessible ranges. Predicted values were extracted using an independent sample of 10,000 points randomly distributed across the study area

alt-text: Unlabelled Table

	Occupied area	Potential area	Accessible area
(a) Occupied area			
Database 1	0.06		
Database 2	0.06		
Database 3	0.05		
Database 4	0.05		
Database 5	0.49		
Database 6	0.48		
(b) Potential/accessible area			
Database 7	0.66	0.25	0.10
Database 8	0.71	0.30	0.11
Database 9	0.67	0.24	0.10
Database 10	0.72	0.29	0.11

Appendix 3. Spatial outputs achieved for each of the ten models developed to answer the target specific question

alt-text: Unlabelled Image

Appendix 4.Appendix 4. Pearson bivariate correlations among predicted continuous prediction values achieved from models calibrated in: (a) occupied area (models 1 to 6) and (b) potential/accessible area (models 7 to 10). Analyses were carried out using two separate samples of 10,000 random points, separated at least 300 m, being each sample distributed across either occupied or potential/accessible areas.

alt-text: Unlabelled Table

	(a) Occupied area					
$r_{Pearson}$	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Model 1	1					
Model 2	0.69	1				
Model 3	0.11	0.22	1			
Model 4	0.28	0.30	0.63	1		
Model 5	0.81	0.59	0.18	0.35	1	
Model 6	0.59	0.75	0.22	0.36	0.56	1
Mean	**0.50**	**0.51**	**0.27**	**0.38**	**0.50**	**0.49**
SD	**0.29**	**0.23**	**0.20**	**0.14**	**0.24**	**0.49**

	(b) Potential/accessible area			
$r_{Pearson}$	Model 7	Model 8	Model 9	Model 10
Model 7	1			
Model 8	0.53	1		

Model 9	0.57	0.48	1	
Model 10	0.43	0.40	0.49	1
Mean	**0.51**	**0.48**	**0.51**	**0.44**
SD	**0.08**	**0.09**	**0.05**	**0.05**

References

Alonso J.C. and Palacín C., The world status and population trends of the Great Bustard: 2010 update, *Chinese Birds* **1**, 2010, 141–147.

Alonso J.C., Morales M.B. and Alonso J.A., Partial migration, and lek and nesting area fidelity in female great bustards, ~~The Condor~~*Condor* **102**, 2000, 127–136.

Alonso J.C., Martín C.A., Alonso J.A., Palacín C., Magaña M. and Lane S.J., Distribution dynamics of a great bustard metapopulation throughout a decade: influence of conspecific attraction and recruitment, ~~Biodiversity and Conservation~~*Biodivers. Conserv.* **13**, 2004, 1659–~~16~~74.

Alonso J.C., Álvarez-Martínez J.M. and Palacín C., Leks in ground-displaying birds: hotspots or safe places?, ~~Behavioral Ecology~~*Behav. Ecol.* **23**, 2011, 491–501.

Álvarez-Martínez J.M., Stoorvogel J.J., Suárez-Seoane S. and de Luis E., Uncertainty analysis as a tool for refining land dynamics modelling on changing landscapes: a case study in a Spanish Natural Park, ~~Landscape Ecology~~*Landsc. Ecol.* **25**, 2010, 1385–1404.

Alvarez-Martínez J.M., Suárez-Seoane S., Palacín C., Sanz J. and Alonso J.C., Can Eltonian processes explain species distributions at large scale? A case study with Great Bustard (*Otis tarda*), ~~Diversity and Distributions~~*Divers. Distrib.* **21**, 2015, 123–138.

Ascough J.C., II, Maier J.H.R., Ravalico K. and Strudley M.W., Future research challenges for incorporation of uncertainty in environmental and ecological decision-making, *Ecol. Model.* **219**, 2008, 383–399.

Austin M.P., Spatial prediction of species distribution: an interface between ecological theory and statistical modelling, ~~Ecological Modelling~~*Ecol. Model.* **157**, 2002, 101–118.

Austin M., Species distribution models and ecological theory: ~~A~~a critical assessment and some possible new approaches, ~~Ecological Modelling~~*Ecol. Model.* **200**, 2007, 1–19.

Barbraud C. and Weimerskirch H., Environmental conditions and breeding experience affect costs of reproduction in blue petrels, *Ecology* **86**, 2005, 682–692.

Barry S. and Elith J., Error and uncertainty in habitat models, ~~Journal of Applied Ecology~~*J. Appl. Ecol.* **43**, 2006, 413–423.

Barton K., MuMIn: Multi-model Inference. R package Version 1.40.4, https://CRAN.R-project.org/package=MuMIn, 2018.

Beven K.J., On the concept of model structural error, ~~Water Science and Technology~~*Water Sci. Technol.* **52**, 2005, 165–175.

BirdLife International, *Otis tarda*. The IUCN red list of threatened species 2017: e.T22691900A119044104, https://doi.org/10.2305/IUCN.UK.2017-3.RLTS.T22691900A119044104.en, 2017, Accessed 10 April 2018.

Borsuk M.E., Stow C.A. and Reckhow K.H., A Bayesian network of eutrophication models for synthesis, prediction, and uncertainty analysis, ~~Ecological Modelling~~*Ecol. Model.* **173**, 2004, 219–239.

Bro-Jørgensen J., Brown M.E. and Pettorelli N., Using the satellite-derived normalized difference vegetation index (NDVI) to explain ranging patterns in a lek-breeding antelope: the importance of scale, *Oecologia* **158**, 2008, 177–182.

Bull J.W., Gordon A., Watson J.E. and Maron M., Seeking convergence on the key concepts in 'no net loss' policy, ~~Journal of Applied Ecology~~*J. Appl. Ecol.* **53**, 2016, 1686–1693.

Burgman M.A., Ferson S. and Akcakaya H.R., Risk Assessment in Conservation Biology, 1993, Chapman and Hall; London.

Burgman M.A., Lindenmayer D.B. and Elith J., Managing landscapes for conservation under uncertainty, *Ecology* **86**, 2005, 2007–2017.

Burnham K.P. and Anderson D.R., Model ~~s~~Selection and ~~m~~Multi-model ~~i~~Inference. A ~~p~~Practical ~~i~~Information-theoric ~~a~~Approach, 2002, Springer; New York.

Burnham K.P., Anderson D.R. and Huyvaert K.P., AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons, *Behav. Ecol. Sociobiol.* **65**, 2011, 23–35.

CNIG, Centro Nacional de Información Geográfica. Instituto Geográfico Nacional. Gobierno de España, http://www.cnig.es, 2012.

Convertino M., Welle P., Muñoz-Carpena R., Kiker G.A., Chu-Agor M.L., Fischer R.A. and Linkov I., Epistemic uncertainty in predicting shorebird biogeography affected by sea-level rise, *Ecol. Model.* **240**, 2012 1-15.

Davison A.C. and Hinkley D.V., Bootstrap Methods and Their Application, 1997, Cambridge University Press.

Dawson T.P., Jackson S.T., House J.I., Prentice I.C. and Mace G.M., Beyond predictions: biodiversity conservation in a changing climate, *Science* **332**, 2011, 53–58.

Elith J., Burgman M.A. and Regan H.M., Mapping epistemic uncertainties and vague concepts in predictions of species distribution, *Ecol. Model.* **157**, 2002, 313–329.

ESRI, ArcInfo desktop GIS 10.5, http://www.esri.com, 2016.

Fedorov V., Mannino F. and Zhang R., Consequences of dichotomization, *Pharm. Stat.* **8**, 2009, 50-61.

Franklin J., Moving beyond static species distribution models in support of conservation biogeography, *Divers. Distrib.* **16**, 2010a, 321–330.

Franklin J., Mapping Species Distributions: Spatial Inference and Prediction, 2010b, Cambridge University Press.

Freckleton R., Dealing with collinearity in behavioural and ecological data: model averaging and the problems of measurement error, *Behav. Ecol. Sociobiol.* **65**, 2011, 91–101.

Graham M.H., Confronting multicollinearity in ecological multiple regression, *Ecology* **84**, 2003, 2809–2815.

Gregory R.S. and Keeney R.L., Making smarter environmental management decisions, *J. Am. Water Resour. Assoc.* **38**, 2002, 1601–1612.

Guisan A., et al., Predicting species distributions for conservation decisions, *Ecol. Lett.* **16**, 2013, 1424–1435.

Harwood J. and Stokes K., Coping with uncertainty in ecological advice: lessons from fisheries, *Trends Ecol. Evol.* **18**, 2003, 617–622.

Heinze G., Wallisch C. and Dunkler D., Variable selection—a review and recommendations for the practicing statistician, *Biom. J.* **60**, 2018, 431–449.

Hirzel A.H. and Le Lay G., Habitat suitability modelling and niche theory, *J. Appl. Ecol.* **45**, 2008, 1372–1381.

Jiménez-Valverde A., Lobo J.M. and Hortal J., Not as good as they seem: the importance of concepts in species distribution modelling, *Divers. Distrib.* **14**, 2008, 885–890.

Johnstone R.A., Models of reproductive skew: a review and synthesis, *Ethology* **106**, 2000, 5–26.

Jonsson A., Andersson L., Alkan-Olsson J. and Arheimer B., How participatory can participatory modeling be? Degrees of influence of stakeholder and expert perspectives in six dimensions of participatory modeling, *Water Sci. Technol.* **56**, 2007, 207-214.

Kearney M., Habitat, environment and niche: what are we modelling?, *Oikos* **115**, 2006, 186–191.

Keeney R.L. and Gregory R.S., Selecting attributes to measure the achievement of objectives, *Oper. Res.* **53**, 2005, 1–11.

Knape J. and de Valpine P., Effects of weather and climate on the dynamics of animal population time series, *Proc. R. Soc. B Biol. Sci.* **278**, 2011, 985–992.

Kuhnert P.M., Martin T.G. and Griffiths S.P., A guide to eliciting and using expert knowledge in Bayesian ecological models, *Ecol. Lett.* **13**, 2010, 900-914.

Kujala H., Burgman M.A. and Moilanen A., Treatment of uncertainty in conservation under climate change, *Conserv. Lett.* **6**, 2013, 73-85.

Lajoie G. and Vellend M., Understanding context dependence in the contribution of intraspecific variation to community trait–environment matching, *Ecology* **96**, 2015, 2912–2922.

Lescroël A., Dugger K.M., Ballard G. and Ainley D.G., Effects of individual quality, reproductive success and environmental variability on survival of a long-lived seabird, *J. Anim. Ecol.* **78**, 2009, 798–806

Lobo J.M., Jiménez-Valverde A. and Hortal J., The uncertain nature of absences and their importance in species distribution modelling, *Ecography* **33**, 2010, 103–114.

MacCallum R.C., Zhang S., Preacher K.J. and Rucker D.D., On the practice of dichotomization of quantitative variables, *Psychol. Methods* **7**, 2002, 19–40.

Magaña M., Alonso J.C., Alonso J.A., Martín C.A., Martín B. and Palacín C., Great bustard (*Otis tarda*) nest locations in relation to leks, *J. Ornithol.* **152**, 2011, 541–548.

MAGRAMA, Natura 2000 Ecological Network. Birds Directive (Special Protection Areas) and Habitats Directive (Sites of Community Importance and Special Areas of Conservation), http://www.magrama.gob.es, 2012.

Martín C.A., Alonso J.C., Alonso J.A., Palacín C., Magaña M. and Martín B., Sex-biased juvenile survival in a bird with extreme size dimorphism, the Great Bustard, *J. Avian Biol.* **38**, 2007, 335–346.

Martin T.G., Burgman M.A., Fidler F., Kuhnert P.M., Low-Choy S., Mcbride M. and Mengersen K., Eliciting expert knowledge in conservation science, *Conserv. Biol.* **26**, 2012, 29–38.

Martínez C., Distribution, density and productivity of great bustards in northwestern Spain: a regional approach, *J. Ornithol.* **149**, 2008, 507–514.

McCarthy M.A., Contending with uncertainty in conservation management decisions, *Ann. N. Y. Acad. Sci.* **1322**, 2014, 77–91.

Menke S.B., Holway D.A., Fisher R.N. and Jetz W., Characterizing and predicting species distributions across environments and scales: argentine ant occurrences in the eye of the beholder, *Glob. Ecol. Biogeogr.* **18**, 2009, 50–63.

Michod R.E., Darwinian Dynamics: Evolutionary Transitions in Fitness and Individuality, 2000, Princeton University Press.

Milner-Gulland E.J. and Shea K., Embracing uncertainty in applied ecology, *J. Appl. Ecol.* **54**, 2017, 2063–2068.

Morales M.B., Alonso J.C. and Alonso J.A., Annual productivity and individual female reproductive success in a Great Bustard population, *Ibis* **144**, 2002, 293–300.

Mostashari A. and Sussman J., Stakeholder-assisted modelling and policy design process for environmental decision-making, *J. Environ. Assess. Policy Manag.* **7**, 2005, 355–386.

Mundry R. and Numm C.L., Stepwise model fitting and statistical inference: turning noise into signal pollution, *Am. Nat.* **173**, 2009, 119–123.

NASA, 2012. SEDAC (Socioeconomic Data and Applications Centre), DAACs (Distributed Active Archive Centres), EOSDIS (Earth Observing System Data and Information System) of the U.S. National Aeronautics and Space Administration. Government of the United States of America.

Neter J., Wasserman W. and Kutner M.H., Applied Linear Statistical Models: Regression, Analysis of Variance, and Experimental Designs, 1990, Irwin; Homewood, IL.

Ninyerola M., Pons X. and Roure J.M., Atlas Climático Digital de la Península Ibérica. Metodología y aplicaciones en bioclimatología y geobotánica, 2005, Universidad Autónoma de Barcelona; Bellaterra, (ISBN 932560-8-7).

Ninyerola M., Pons X. and Roure J.M., Monthly precipitation mapping of the Iberian Peninsula using spatial interpolation tools implemented in a Geographic Information System, *Theor. Appl. Climatol.* **89**, 2007, 195–209.

Openshaw S. and Taylor P.J., The modifiable areal unit problem, In: Wrigley and Bennett, (Eds.), *Quantitative Geography: A British View,* 1981, Routledge and Kegan Paul Ed; London, 60–69.

Osborne P.E., Suárez-Seoane S. and Alonso J.C., Behavioural mechanisms that undermine species envelope models: the causes of patchiness in the distribution of great bustards *Otis tarda* L. in Spain, *Ecography* **6**, 2007, 819–829.

Parejo D., Oro D. and Danchin E., Testing habitat copying in breeding habitat selection in a species adapted to variable environments, *Ibis* **148**, 2006, 146–154.

Pearman P.B., Guisan A., Broennimann O. and Randin C.F., Niche dynamics in space and time, *Trends Ecol. Evol.* **23**, 2008, 149–158.

Pearson R.G., Thuiller W., Araújo M.B., Martínez-Meyer E., Brotons L., McClean C., Miles L., Segurado P., Dawson T.P. and Lees D.C., Model-based uncertainty in species range prediction, *J. Biogeogr.* **33**, 2006 1704–1711.

Peterson A.T., Uses and requirements of ecological niche models and related distributional models, *Biodivers. Inform.* **3**, 2006, 59–72.

Peterson A.T., Soberon J., Pearson R.G., Anderson R.P., Martínez-Meyer E., Nakamura M. and Araújo M.B., Ecological Niches and Geographic Distributions, 2012, Princeton University Press; Princeton.

Pettorelli N., Ryan S., Mueller T., Bunnefeld N., Jedrzejewska B., Lima M. and Kausrud K., The Normalized Difference Vegetation Index (NDVI): unforeseen successes in animal ecology, *Clim. Res.* **46**, 2011, 15–27.

Pinheiro J.C. and Bates D.M., Mixed-effects Models in S and SPLUS, 2000, Springer.

Pinto M., Rocha P. and Moreira F., Long-term trends in great bustard populations in Portugal suggest concentration in single high quality area, *Biol. Conserv.* **124**, 2005, 415–423.

Polasky S., Carpenter S.R., Folke C. and Keeler B., Decision-making under great uncertainty: environmental management in an era of global change, *Trends Ecol. Evol.* **26**, 2011, 398–404.

R Development Core Team, R: A Language and Environment for Statistical Computing, 2017, R Foundation for Statistical Computing; Vienna (Austria) http://www.R-project.org.

Randin C.F., Dirnböck T., Dullinger S., Zimmermann N.E., Zappa M. and Guisan A., Are niche-based species distribution models transferable in space?, *J. Biogeogr.* **33**, 2006, 1689–1703.

Redpath S.M., Arroyo B.E., Leckie F.M., Bacon P., Bayfield N., Gutierrez R.J. and Thirgood S.J., Using decision modeling with stakeholders to reduce human–wildlife conflict: a raptor–grouse case study, *Conserv. Biol.* **18**, 2004, 350–359.

Regan H.M., Colyvan M. and Burgman M.A., A taxonomy and treatment of uncertainty for ecology and conservation biology, *Ecol. Appl.* **12**, 2002, 618–628.

Rieucau G., Exploring the costs and benefits of social information use: an appraisal of current experimental evidence, *Philos. Trans. R. Soc. B* **366**, 2011, 949–957.

Rosema A., Verhees L. and van Putten E., European Energy and Water Balance Monitoring System. Scientific Report. Contract ENV4-CT97-0478, 2001, EU 4th Framework Program.

Royston P., Altman D.G. and Sauerbrei W., Dichotomizing continuous predictors in multiple regression: a bad idea, *Stat. Med.* **25**, 2006, 127–141.

Runge M.C., An introduction to adaptive management for threatened and endangered species, *J. Fish Wildl. Manag.* **2**, 2011, 220–233.

Ryder T.B., Parker P.G., Blake J.G. and Loiselle B.A., It takes two to tango: reproductive skew and social correlates of male mating success in a lek-breeding bird, *Proc. R. Soc. B* **276**, 2009, 2377–2384.

SIGPAC, Database of Agricultural Plots Geographic Information System, 2012, Government of Spain http://www.magrama.gob.es/en/agricultura/temas/sistema-de-informacion-geografica-de-parcelas-agricolas-sigpac.

SIOSE, Land Cover Information System of Spain, 2012, Government of Spain http://www.siose.es/siose.

Soberon J., Grinnellian and Eltonian niches and geographic distributions of species, *Ecol. Lett.* **10**, 2007, 1115–1123.

Suárez-Seoane S. and Baudry J., Scale dependence of spatial patterns and cartography on the detection of landscape change: relationships with species' perception, *Ecography* **25**, 2002, 499–511.

Suárez-Seoane S., Osborne P.E. and Alonso J.C., Large-scale habitat selection by agricultural steppe birds in Spain: identifying species–habitat responses using generalized additive models, *J. Appl. Ecol.* **39**, 2002, 755–771.

Suárez-Seoane S., Osborne P.E. and Rosema A., Can climate data from METEOSAT improve wildlife distribution models?, *Ecography* **27**, 2004, 629–636.

Suárez-Seoane S., Álvarez-Martínez J.M., Wintle B.A., Palacín C. and Alonso J.C., Modelling the spatial variation of vital rates: an evaluation of the strengths and weaknesses of correlative species distribution models, *Divers. Distrib.* **23**, 2017, 841–853.

Symonds M.R. and Moussalli A., A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike's information criterion, *Behav. Ecol. Sociobio.* **65**, 2011, 13–21.

Synes N.W. and Osborne P.E., Choice of predictor variables as a source of uncertainty in continental-scale species distribution modelling under climate change, *Glob. Ecol. Biogeogr.* **20**, 2011,

904–914.

Thuiller W., BIOMOD-optimizing predictions of species distributions and projecting potential future shifts under global change, ~~Global Change Biology~~*Glob. Chang. Biol.* **9**, 2003, 1353-1362.

Thuiller W., Brotons L., Araújo M.B. and Lavorel S., Effects of restricting environmental range of data to project current and future species distributions, *Ecography* **27**, 2004, 165–172.

Tingley ~~M. W.~~M.W., Monahan ~~W. B.~~W.B., Beissinger ~~S. R.~~S.R. and Moritz C., Birds track their Grinnellian niche through a century of climate change, ~~Proceedings of the National Academy of Sciences~~*Proc. Natl. Acad. Sci.* **106**, 2009, 19637–19643.

Uusitalo L., Lehikoinen A., Helle I. and Myrberg K., An overview of methods to evaluate uncertainty of deterministic models in decision support, *Environ. Model. Softw.* **63**, 2015, 24–31.

Venables W.N. and Ripley B.D., Modern Applied Statistics With S, 4th ed., 2002, Springer.

Warren D.L., In defense of "niche modelling", *Trends Ecol. Evol.* **27**, 2012, 497–500.

Wong D., In: Warf and Hansen, (Eds.), *The Modifiable Areal Unit Problem (MAUP). WorldMinds: Geographical Perspectives on 100 Problems,* 2004, Springer; Berlin, 571–575.

Zuur A.F., Ieno E.N. and Elphick C.S., A protocol for data exploration to avoid common statistical problems, ~~Methods in Ecology and Evolution~~*Methods Ecol. Evol.* **1**, 2010, 3-14.

Highlights

- Underspecificity drives to multiple methodological choices.

- It hampers the definition of conservation actions.

- It can be reduced in assisted modelling frameworks.

Queries and Answers

Query:

Your article is registered as a regular item and is being processed for inclusion in a regular issue of the journal. If this is NOT correct and your article belongs to a Special Issue/Collection please contact s.ananthakrishnan@elsevier.com immediately prior to returning your corrections.

Answer: Yes

Query:

Please confirm that given names and surnames have been identified correctly and are presented in the desired order, and please carefully verify the spelling of all authors' names.

Answer: Yes

Query:

The author names have been tagged as given names and surnames (surnames are highlighted in teal color). Please confirm if they have been identified correctly.

Answer: Yes

Query:

Please check whether the designated corresponding author is correct, and amend if necessary.

Answer: It is correct, thank you.

Query:

Citation "Martin et al., 2007" has not been found in the reference list. Please supply full details for this reference.

Answer: Martín, C.A., Alonso, J.C., Alonso, J.A., Palacín, C., Magaña, M., Martín, B., 2007. Sex-biased juvenile survival in a bird with extreme size dimorphism, the Great Bustard. Journal of Avian Biology 38, 335–346

Query:

Citation "Alvarez-Martínez et al., 2014" has not been found in the reference list. Please supply full details for this reference.

Answer: Please change this citation by "Alvarez-Martínez et al., 2015", which it is included in the reference list.

Query:

Citation "Soberon and Peterson, 2005" has not been found in the reference list. Please supply full details for this reference.

Answer: Soberon, J., Peterson, A.T., 2005. Interpretation of models of fundamental ecological niches and species' distributional areas. Biodiversity Informatics 2, 1-10

Query:

The citation "Rieucau and Giraldeau, 2011" has been changed to "Rieucau, 2011" to match the author name/date in the reference list. Please check if the change is fine in this occurrence and modify the subsequent occurrences, if necessary.

Answer: This is the correct citation: Rieucau, G., Giraldeau, L.A., 2011. Exploring the costs and benefits of social information use: an appraisal of current experimental evidence. *Philosophical Transactions of the Royal Society of London B: Biological Sciences 366* (1567), 949-957.

Query:

The citation "Burgman et al., 2007" has been changed to "Burgman et al., 2005" to match the author name/date in the reference list. Please check if the change is fine in this occurrence and modify the subsequent occurrences, if necessary.

Answer: Ok

Query:

Have we correctly interpreted the following funding source(s) and country names you cited in your article: "Spanish Ministry of Science and Innovation, Spain".

Answer: Yes

Query:

Please check the layout and presentation of Appendix 2 if correct.

Answer: Yes, but we have made some changes on the legend

Query:

Please check the layout and presentation of Appendix 4 if correct.

Answer: Yes, but we have made some changes on the legend

Query:

Please provide a definition for the significance of bold in the Appendix 4.

Answer: Ok, done