

Research

Addressing common pitfalls does not provide more support to geographical and ecological abundant-centre hypotheses

L. Santini, S. Pironon, L. Maiorano and W. Thuiller

L. Santini (<http://orcid.org/0000-0002-5418-3688>) (luca.santini.eco@gmail.com), Dept of Environmental Science, Inst. for Water and Wetland Research, Radboud Univ., GL Nijmegen, the Netherlands. – S. Pironon (<http://orcid.org/0000-0002-8937-7626>), Royal Botanic Gardens, Kew, Richmond, UK. – L. Maiorano (<http://orcid.org/0000-0002-2957-8979>), Dept of Biology and Biotechnologies, Sapienza Univ. di Roma, Italy. – W. Thuiller, Univ. Grenoble Alpes, CNRS, Univ. Savoie Mont Blanc, Laboratoire d'Ecologie Alpine (LECA), Grenoble, France.

Ecography

42: 696–705, 2019

doi: 10.1111/ecog.04027

Subject Editor: Carsten Dormann

Editor-in-Chief: Miguel Araújo

Accepted 5 September 2018



A long-standing hypothesis in biogeography is that a species' abundance is highest at the centre of its geographical or environmental space and decreases toward the edges. Several studies tested this hypothesis and provided mixed results and overall weak support to the theory. Most studies, however, are affected by several limitations related to the sample size, the comparability among abundance measures, the definition of species geographic range and corresponding environmental space, and the proxy variables used to represent centrality/marginality gradients.

Here we test the abundant-centre hypothesis on 108 bird and mammal species and embrace the plural nature of the hypothesis by considering 9 geographic and ecological centrality/marginality measures. We analyse the species-specific effect sizes using a meta-analytical approach, and test whether the support for the hypothesis is mediated by species dispersal abilities, and the geographic and environmental coverage of the data.

The summary effect sizes estimated for the 9 measures are largely inconsistent with the theoretical expectations and show a significant amount of residual heterogeneity. Variables such as dispersal distance, geographic and environmental coverage of the data, appear important in explaining the variation observed between different species, but the results are contrary to those originally hypothesized, and inconsistent across centrality/marginality measures and the datasets used.

We show that addressing common pitfalls in previous studies does not provide more support to the abundant-centre hypothesis, with support being very dependent on the centrality/marginality measure tested, the geographic extent considered for the test, and geographic and environmental coverage of the data. The abundant-centre hypothesis so far remains an appealing speculation with little and variable empirical support.

Keywords: abundant-centre hypothesis, density, meta-analysis

Introduction

A long-standing hypothesis in biogeography and macroecology is that a species' abundance is highest at the centre of its geographical range and decreases toward the edges, i.e. abundant-centre hypothesis (Sagarin and Gaines 2002, Sexton et al. 2009, Pironon et al. 2017). This hypothesis goes back to Grinnell (1922) who proposed the idea of a general tendency of animals to disperse from the centre of the range to the edges in analogy to gas molecules, but several authors envisioned other explanations (see Pironon et al. 2017 for an historical perspective). Nowadays, the most widely accepted explanation is that the species geographic range would be a spatial representation of its ecological niche, so that abundance would decrease from optimal conditions at the centre of its range towards marginal conditions at the periphery (Hengeveld and Haeck 1982, Brown 1984, Colwell and Rangel 2009).

However, it is notorious that geographical and environmental gradients do not strictly overlap and, for this reason, abundance should not systematically follow a geographical or ecological centre-to-edge gradient (Soule 1973, Chardon et al. 2015, Pironon et al. 2015, but see Lee-Yaw et al. 2016). Indeed, research on several animal and plant taxa provided support for an environmentally-based abundant-centre hypothesis rather than the originally envisioned geographic one (VanDerWal et al. 2009, Martínez-Meyer et al. 2013, Van Couwenberghe et al. 2013), whereas others did not find any support for the two visions of the hypothesis (Pironon et al. 2015, Dallas et al. 2017). Many of these papers concluded that the expected geographic abundant-centre pattern is actually a reflection of the environmental gradients and corresponding variation in niche suitability. Yet, this hypothesis remains widely debated and more systematic studies are needed (Dallas et al. 2017, Pironon et al. 2017, Soberon et al. 2018).

Although its empirical bases are still weak, the abundant-centre hypothesis has been used to support many other ideas in conservation, disease ecology, evolution and genetics (Sagarin and Gaines 2002, Sagarin et al. 2006). Only a limited number of studies have directly tested this hypothesis and provided a mixed support to it (Sagarin and Gaines 2002, Sexton et al. 2009, Pironon et al. 2017). Most studies focused on one or few species, with only a few analyses testing the hypothesis on a large number of species: passerines (Blackburn et al. 1999a), forest birds (Emlen et al. 1986), and trees, fishes, birds and mammals (Dallas et al. 2017). However, while they pave the way for more research on the topic, they also suffer from important limitations and biases.

First, in most cases the abundance estimates being considered are incomparable in space. For example, some studies used bird abundance data derived from transect counts collected by volunteers (Martínez-Meyer et al. 2013, Dallas et al. 2017). Bird transect counts were either averaged over route totals (Martínez-Meyer et al. 2013) or standardized by travel time (Dallas et al. 2017), irrespective of distance covered and

individual observation biases. However, the variance component associated with the experience of the observer has been shown to be equivalent to the variance linked to sampling time (Johnston et al. 2017). Similarly, mammal trapping data have been standardized by trap nights, irrespective of the number of traps, sampling design and area surveyed. While trap night is a fair standardization measure for studies using the same sampling design, it cannot be used to standardize estimates from different studies. For example, the number of animals trapped depends on the distance between traps (Hayne 1950, Renzulli et al. 1980), and whether individuals have been marked to avoid pseudo-replicates or not. Finally, while different methodologies are known to provide different – often incomparable – estimates (Seddon et al. 2003, Gottschalk and Huetmann 2011), density estimates from multiple studies employing diverse methodologies have often been mixed to test the abundant-centre hypothesis (Martínez-Meyer et al. 2013). Using such data likely introduces substantial noises, preventing or confusing the detection of any potential geographic and/or environmental gradients, at least in intra-specific patterns. Second, many of these studies limited their analyses to a given geographic area, irrespective of the range-wide distribution of the species considered (Blackburn et al. 1999b, Martínez-Meyer et al. 2013, Dallas et al. 2017), which can result in important biases in the estimate of distance from centre to edge, and in large underestimation of species niches (Guisan and Thuiller 2005, Pironon et al. 2017). Third, several approaches have been used to define species geographic distribution. Dallas et al. (2017), for example estimated species geographic ranges by fitting a minimum convex polygon (MCP) around their abundance sampling sites, and repeated the analysis using an MCP fitted around occurrence records collected from the GBIF database (<www.gbif.org/>), which is notoriously affected by errors and geographic biases in species occurrences (Beck et al. 2013, Meyer et al. 2016). MCPs are highly sensitive to sample size and spatial biases/outliers in the data, potentially showing large commission/omission errors. Therefore, MCPs may exclude large parts of the species ranges that were not sampled, and include large areas in which the species are absent; by using an MCP these areas may be considered as part of the range (Burgman and Fox 2003, Gaston and Fuller 2009). This geographic error translates into an environmental error, because some environmental values will be included or excluded from the species' environmental niche. Fourth, studies have focused on different measures of centrality. Some have exclusively focused on centrality (Martínez-Meyer et al. 2013, Dallas et al. 2017) while marginality has often been considered a major driver of the pattern (Blackburn et al. 1999b, Sexton et al. 2009). However, centrality and marginality are not two sides of the same coin. Depending on the shape of the geographic range, these measures can largely differ as occurrences far from the geographic centroid can also be far from the edge, and vice-versa. Fifth, species niches can be estimated in several ways. Examples include centrality in the environmental niche (Dallas et al. 2017), probabilities

of presence as estimated from species distribution models (VanDerWal et al. 2009, Martínez-Meyer et al. 2013) or niche marginality measures (Soares et al. 2015). The consistency among these measurements can vary largely, so different conclusions might be reached depending on how centrality/marginality are estimated.

To address these limitations, here we test the abundant-centre hypothesis on 9 bird and 99 mammal species and embrace the plural nature of the hypothesis by considering several geographic and ecological centrality/marginality measures. We analyse the species-specific effect sizes using a meta-analytical approach, and repeat the analyses with more stringent criteria of data quality. We test whether the support for the hypothesis is mediated by species dispersal abilities and the geographic and environmental coverage of the data. We show that addressing the aforementioned limitations does not provide more support to the abundant-centre hypothesis and conclude that this hypothesis might simply not hold true.

Methods

Data collection

We obtained population density estimates for bird and mammal species from the TetraDENSITY database (Santini et al. 2018a). This database encompasses global georeferenced population density estimates for terrestrial vertebrates. We excluded all migratory birds from the selection, for which the geographic range is divided in breeding and non-breeding areas and the definition of a unique geographic or environmental space is challenging. To match the environmental predictors we only retained population density estimated after 1980. We defined independent locations as locations in different 1-degree cells (~ 110 km at the equator), and excluded all density records with a spatial precision lower than 1-degree. We selected species for which density estimates were available in at least 5 independent locations within the species range and retained multiple population density estimates at the same locations when available. This criterion ensured to have density estimates that were not clumped in both geographic and environmental space and covered several hundred kilometres. The final dataset included 9 bird and 99 mammal species (Supplementary material Appendix 1 Table A1).

We used species range polygons as mapped by the IUCN Red List assessment and BirdLife International (BirdLife International and Handbook of the Birds of the World 2017, IUCN 2017) resampled at 1-degree resolution to meet the spatial accuracy of the data. To model species niches we considered the bioclimatic variables available from CHELSA (Karger et al. 2016) for the period 1979–2013, the land cover variables available from HYDE 3.1 (Klein Goldewijk et al. 2011) for the period 1980–2010, and the AVHRR NDVI 3g v.1 time series from <<https://ecocast.arc.nasa.gov/>> for the period 1981–2015 as a measure of primary productivity.

We extracted 11 binary habitat maps from yearly land cover maps and averaged them across all years (i.e. polar desert, boreal forest, temperate forest, tropical forest, grassland, savannah, shrubland, desert, cropland, pasture). We estimated the mean annual NDVI and the NDVI seasonality (standard deviation) across all years. Finally, we resampled the 31 layers at 1-degree resolution and computed a principal component analysis to identify the major axes of variation. We extracted the first eight components explaining > 80% of the variance and we used them as predictors for estimating species niches in the following analyses (Supplementary material Appendix 1 Fig. A1).

Geographic and environmental abundant-centre estimates

We used three geographic and six ecological niche measures to describe the centrality and marginality gradients (Fig. 1). Geographic measures included: 1) distance from the range centroid (CD); 2) minimum distance from the edge (ED); 3) and the ratio between the two (CD/ED). Geographic distances were calculated as ‘great-circle distance’ using the un-projected (lat-long) vectorial range maps. Ecological niche measures included: 1) environmental suitability calculated as probability of occurrence from species distribution models (SUIT); 2) environmental marginality (MAR); 3) Mahalanobis distance (MAH); 4) Euclidean distance from the centroid of the environmental space (envCD); 5) minimum distance from the edge of the environmental space (envED); 6) and the ratio between the two latter measures (envCD/envED) (Fig. 1). We modelled habitat suitability (SUIT) using an ensemble modelling approach that included six models (Thuiller et al. 2009): generalized linear models, generalized additive models, random forests, multiple adaptive regression splines, artificial neural network, boosted regression trees. We considered all 1-degree cells of the geographic range as presence points and all the cells of the same continent(s) as background points. We estimated marginality (MAR) using the ecological niche factor analysis (ENFA) with the same procedure used for the ensemble modelling approach. ENFA quantifies environmental marginality of a species by comparing the species mean with the global mean of the environmental space considered (Hirzel et al. 2002). We estimated the Mahalanobis distance from the mean of the environmental space (Farber and Kadmon 2003), this has been argued to be a good proxy for marginality as it accounts for the covariance of the niche axes (Soberon et al. 2018). To estimate the environmental space, we ran a principle component analysis on the cell values of the eight principal components within species range polygons. Then, we generated a convex hull around the cell values distributed along the first two components. Finally, we estimated the Euclidean distance of all cells from the centroid (EnvCD) and the edge (EnvED) of the convex hull. To assess the agreement between these eight measures in representing centrality and marginality, we assessed their correlations across all density record locations.

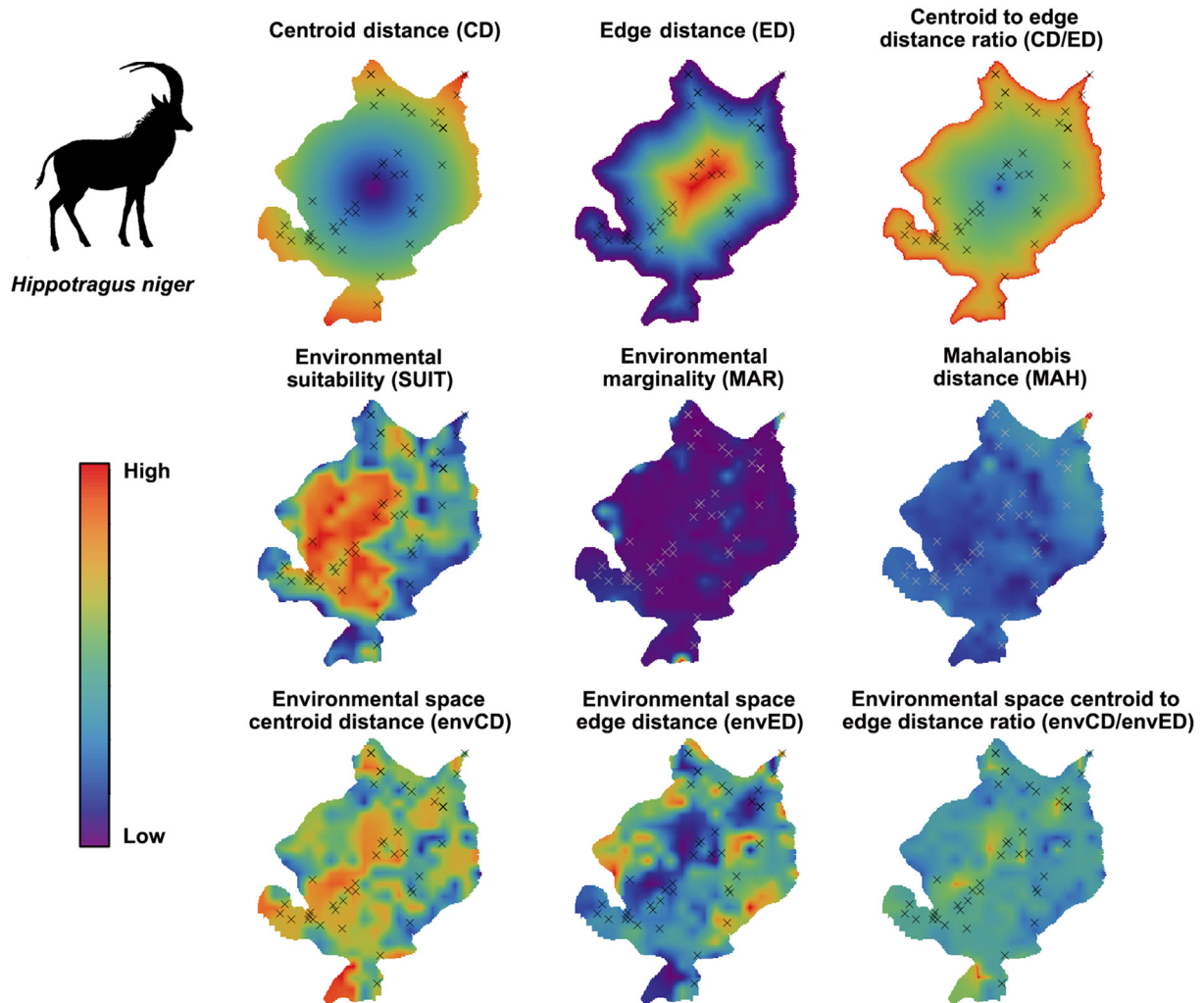


Figure 1. Example of the 9 measures of centrality and marginality used on the geographic range of the sable antelope *Hippotragus niger*. Crosses indicate the location of the density estimates. CD/ED and envCD/envED were log-transformed for visualization purposes, they were not log-transformed for the analysis as Spearman's rank correlation is insensitive to transformation.

We tested the relationship between population density and geographic and ecological measures for all species using Spearman rank correlation. To account for the multiple measurements within the same cells, we estimated a weighted correlation by weighting each observation as one divided by the number of observations per cell. This prevented to overweight oversampled locations in the correlation coefficients. In order to estimate the overall trend and agreement across species, we performed a meta-analysis on the species effect sizes (Borenstein et al. 2009). We transformed correlation coefficients to Fisher's z scores using the correlation sample size to obtain the effect size for each species. We conducted a mixed-effect meta-analysis on the transformed effect size values and the associated variance of each of the 9 measures to calculate a summary effect size. We used a nested random effect 'Class/Order/Family' to account for the different taxonomic groups in our sample. Finally, we transformed the summary effect and the

confidence intervals back to correlation coefficients for interpretability. We tested the residual heterogeneity for all models using the Q-statistic. Significant Q test indicates that a significant amount of variability exists between the effect sizes.

Sensitivity to more stringent criteria of data selection

Because density estimates are influenced by the sampling method and this may reduce their comparability, we repeated the analysis for a subset of density records that met more stringent criteria. For each species we selected only estimates derived from the same methodological approach (e.g. distance sampling, mark-recapture, censuses, ...) (Santini et al. 2018a). This restricted our dataset to 2 species of birds and 56 species of mammals, and included multiple correlation coefficients for some species (one per sampling method) (Supplementary material Appendix 1 Table A2). We ran a

second mixed-effect meta-analysis on this restricted dataset using 'Class/Order/Family/Species' as nested random effects, in order to control for the pseudo-replication at the species level.

Assessing the effect of species dispersal and the geographic and environmental coverage of the data

Species traits can determine whether species follow the abundant-centre pattern (Dallas et al. 2017). Similarly, the distribution of the density records can determine our ability to detect a signal in the data. Here we tested the effect of three possible variables: dispersal distance, geographic coverage and niche environmental tolerance. Dispersal distance reduces population isolation and tends to synchronize the demography of populations within species geographic distribution (Paradis et al. 1999). Therefore, we might expect a smaller effect in species that disperse over large areas. We used body mass to predict median dispersal distance in birds and mammals using the allometric models in Sutherland et al. (2000) and Santini et al. (2013), respectively. We obtained species body mass from the EltonTrait database (Wilman et al. 2014). The geographic distribution of data points over species range can influence our ability to detect a signal in the data. We measured geographic coverage as the area percentage of the minimum convex polygon around density records, over the minimum convex polygon around the species geographic range. Similarly, the extent to which environmental conditions vary across sampled locations within the species ranges can determine whether we should expect a corresponding gradient in population density. Following Dolédec et al. (2000), we calculated niche environmental tolerance (niche breadth) for each species using the values of mean annual temperature, temperature seasonality, annual precipitation and precipitation seasonality observed in the location of the density estimates. We expect the dispersal distance and geographic coverage to be particularly relevant for geographic measures (CD, ED, CD/ED), and the niche tolerance to be particularly relevant for environmental-niche measures (SUIT, MAR, MAH, envCD, envED, envCD/envED). To estimate whether these three variables could alter the meta-analysis results, we modelled them as covariates (moderators) of the mixed-effect meta-analytical model (metaregression). We ran a model selection using AIC to select moderators, and only retained them if $\Delta AIC < 2$ from the only-intercept model. Models were fitted using maximum likelihood (ML) for model selection, and using restricted maximum likelihood (REML) for coefficient estimation (Pinheiro and Bates 2000). We tested the overall effect of moderators using the omnibus test.

All analyses were entirely conducted in R 3.0.3 (R Core Team), using the package 'rgeos' (Bivand and Rundel 2013), 'geosphere' (Hijmans et al. 2015), 'maptools' (Lewin-Koh and Bivand 2011) and 'raster' (Hijmans and van Etten 2014) for GIS operations, 'biomod2' (Thuiller et al. 2009) and 'adehabitatHS' (Calenge 2006) for the species distribution and ENFA models, 'wCorr' package for estimating

weighted correlation coefficients (Emad and Bailey 2017), and 'metafor' (Viechtbauer 2010) for the meta-analysis.

Data deposition

Data available from Figshare Digital Repository: <<http://dx.doi.org/10.6084/m9.figshare.7117880.v1>> (Santini et al. 2018c).

Results

Some of the 9 measures of centrality and marginality gradients considered showed little agreement, while others were highly correlated. The correlation was low between SUIT and MAR and all other measures (Spearman's $\rho < 0.5$), whereas it was high between CD and MAH ($\rho = 0.59$), ED and CD/ED ($\rho = 0.89$), envCD and envED ($\rho = -0.67$), envCD and envCD/envED ($\rho = 0.89$) and envED and envCD/envED ($\rho = -0.92$) (Supplementary material Appendix 1 Fig. A2).

The Spearman's ρ coefficients of the relationship between density and the nine abundant-centre gradients tested were approximately equally distributed around zeros showing no clear support of any of the geographical predictions (Fig. 2, Supplementary material Appendix 1 Table A1). Several correlations were significant and were in the expected direction, but these cases were not more frequent than cases in which the correlations were significant and were in the opposite direction (Fig. 2; Supplementary material Appendix 1 Table A1).

The sign of the intercepts (summary effect size) of the mixed-effect meta-analyses seemed to support the expectations for CD, SUIT, MAR, envCD and envCD/envED but not in ED, CD/ED, MAH, and envED, further, none of the models was significant (Table 1). In all models, the Q test of residual heterogeneity was highly significant ($p < 0.001$) indicating a considerable amount of variation around the summary effect size (Table 1).

Not all moderators were selected in the meta-regressions. Dispersal distance was never selected, geographic coverage was selected and significant in the metaregressions for envCD and envCD/envED, and environmental niche tolerance was selected and significant in the metaregressions for CD and CD/ED (Table 2). In these four metaregression models the omnibus test was significant (Table 2), indicating that the moderators contributed to explain a significant amount of variation around the summary effect sizes. However, these results did not meet our original expectation that geographic coverage would be selected for geographic measures and environmental tolerance for niche-based measures. Further, all the effects detected were opposite to our original expectations, indicating that envCD and envCD/envED were less supported when geographic coverage was high, and that CD and CD/ED were less supported with environmental niche tolerance was high (Table 2, Fig. 3).

We repeated the correlations and the meta-analyses and metaregressions using the restricted dataset, which only

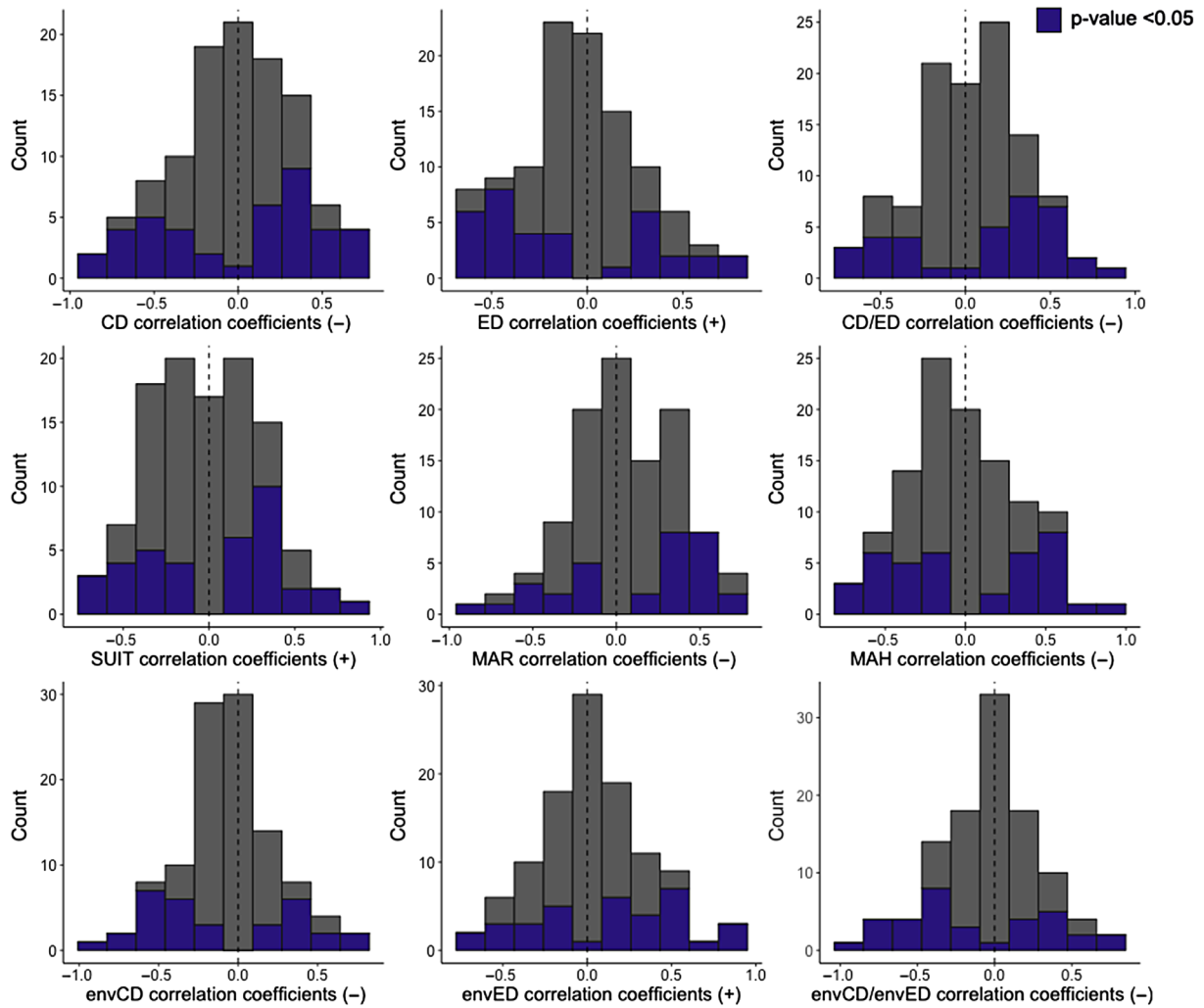


Figure 2. Distribution of the Spearman's correlation coefficients between species population density and the 9 centrality/marginality measures. Blue bars represent the significant coefficients. The sign in parentheses indicates the expected sign of the coefficient under the abundant-centre hypothesis. CD=centroid distance; ED=edge distance, SUIT=suitability; MAR=marginality; MAH=Mahalanobis distance; envCD=distance from the centroid of the environmental space; envED=distance from the edge of the environmental space.

Table 1. Z coefficients (SE) of the mixed effect meta-analyses. QE=QE statistic for the test of residual heterogeneity; p-values: *= $p < 0.05$; **= $p < 0.01$; ***= $p < 0.001$. CD=centroid distance; ED=edge distance, SUIT=suitability; MAR=marginality; MAH=Mahalanobis distance; envCD=distance from the centroid of the environmental space; envED=distance from the edge of the environmental space.

Measure	Expected relationship	Intercept (SE)	QE
CD	Negative	-0.083 (0.059)	420.256 ***
ED	Positive	-0.075 (0.05)	259.898 ***
CD/ED	Negative	0.055 (0.047)	282.283 ***
SUIT	Positive	0.063 (0.121)	334.826 ***
MAR	Negative	-0.018 (0.106)	366.27 ***
MAH	Negative	0.017 (0.042)	324.857 ***
envCD	Negative	-0.071 (0.066)	245.849 ***
envED	Positive	0.047 (0.065)	277.746 ***
envCD/envED	Negative	-0.050 (0.072)	285.063 ***

included density records estimated using comparable methodologies. As in the full dataset, the correlation coefficients were equally distributed around zero showing no clear support for any of the predictions (Supplementary material Appendix 1 Fig. A3, Table A2). Consistently, the meta-analyses results did not provide any support to the abundant-centre hypothesis (Supplementary material Appendix 1 Table A3). Only in CD, SUIT, envED and envCD/envED the sign of the summary effect size supported the expectations, but none of the test summary effect size was significant (Supplementary material Appendix 1 Table A3). We then tested the effect of the three moderators, but only environmental niche tolerance was selected through AIC in the metaregression for ED and CD/ED. However, consistently with the results from the full dataset, their effects were opposite to our expectations (Supplementary material Appendix 1 Fig. A4, Table A4).

Table 2. Z coefficients (SE) of the mixed effect meta-regressions. QE=QE statistic for the test of residual heterogeneity; QM= test statistic for the omnibus test of coefficients; p-values: *= $p < 0.05$; **= $p < 0.01$; ***= $p < 0.001$. CD=centroid distance; ED=edge distance, SUIT=suitability; MAR=marginality; MAH=Mahalanobis distance; envCD=distance from the centroid of the environmental space; envED=distance from the edge of the environmental space.

Measure	Expected relationship	Intercept (SE)	Tolerance (SE)	Geographic coverage (SE)	QE	QMp
CD	Negative	0.112 (0.072)	0.297 (0.063) ***		400.439 ***	22.109 ***
ED	Positive	-0.075 (0.05)			259.898 ***	2.239
CD/ED	Negative	0.149 (0.068) *	0.141 (0.061) *		279.595 ***	5.313 *
SUIT	Positive	0.063 (0.121)			334.826 ***	0.268
MAR	Negative	-0.018 (0.106)			366.27 ***	0.030
MAH	Negative	0.017 (0.042)			324.857 ***	0.160
envCD	Negative	-0.123 (0.071)		0.002 (0.001) *	239.597 ***	5.348 *
envED	Positive	0.047 (0.065)			277.746 ***	0.528
envCD/envED	Negative	-0.107 (0.069)		0.002 (0.001) *	281.776 ***	5.295 *

Discussion

Here we tested the abundant-centre hypothesis using 9 different geographical or environmental centre-periphery gradients. We found summary effect sizes across species to be largely inconsistent with the theoretical expectations. We also found some effect sizes to be dependent on the observed environmental variation captured and the data geographic coverage, yet the effects were contrary to those originally hypothesized. As a consequence, our results do not seem to provide support for any of the interpretation of the abundant-centre hypothesis.

Previous studies provided weak or no support for the geographical interpretation while they provided support for the niche suitability vision (VanDerWal et al. 2009, Martínez-Meyer et al. 2013, Van Couwenberghe et al. 2013). However, several methodological and conceptual differences between this work and previous studies are to be noted. First, previous studies tested the hypothesis on a smaller set of species (but see Dallas et al. 2017). Because apparently the relationship can be positive or negative (Fig. 2, Supplementary material Appendix 1 Fig. A3), the smaller the set of species the higher the probability the hypothesis is supported by chance. Second, previous studies focused on specific geographic area, which implies that the niche and distribution of many of the species considered were only partially covered (Blackburn et al. 1999b, Martínez-Meyer et al. 2013, Dallas et al. 2017). Here we estimated geographic and niche measures on the entire

geographic range of the species, therefore presumably better capturing the centrality/marginality patterns. Third, considering solely one metric to assess geographical or ecological centre-periphery gradients may be misleading (Pironon et al. 2015). Here we showed that 9 different geographical and ecological gradients are only partly correlated, which explains the lack of correspondence between geographical and ecological gradients as found previously (Martínez-Meyer et al. 2013, Chardon et al. 2015, Pironon et al. 2015, but see Lee-Yaw et al. 2016, Pironon et al. 2017) and suggests that the plurality of the notions of centre and periphery can drive the results in any particular direction. Further, our results do not confirm previous finding that used the same niche measures and provided support to the theory (VanDerWal et al. 2009, Martínez-Meyer et al. 2013). Yet, for example, VanDerWal et al. found a positive relationship only with the upper limit of local abundance (VanDerWal et al. 2009). Interestingly, although some studies supported a relationship between abundance and suitability under the abundant-centre hypothesis, a recent meta-analysis encompassing a wide range of taxa showed that the relationship between environmental suitability and abundance is often weak or absent, or even reverse in some cases (Weber et al. 2017). Similarly, studies focusing on environmental suitability and demographic rates in plants found contrasting results (Thuiller et al. 2014, Pironon et al. 2015, Csergő et al. 2017). These studies, together with our results, suggest that only some species conform to the abundant-centre pattern, and even among these

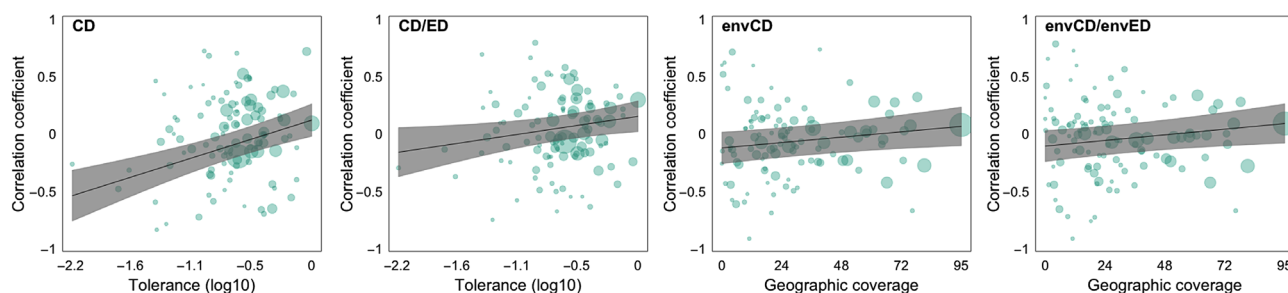


Figure 3. Partial effects of environmental niche tolerance and geographic coverage on the effect sizes of the centrality and marginality measures. Size of data points is proportional to the sample size. CD=centroid distance; ED=edge distance, SUIT=suitability; envCD=distance from the centroid of the environmental space; envED=distance from the edge of the environmental space.

the probability of finding spurious relationships is high. This, combined with a likely publication bias toward supportive and significant results in the literature (Leimu and Koricheva 2004, Parker et al. 2016) may be responsible of the contrasting accumulating evidence.

There might be several explanations for this lack of support, both biological and methodological. First, population density can be much more dependent on factors (e.g. inter-specific interactions) than environmental suitability (Boakes et al. 2017). Second, some animal populations can conform to an 'ideal despotic distribution' rather than an 'ideal free distribution'. The ideal despotic distribution predicts that highly suitable areas are monopolized by a relatively low number of superior competitors, so that density results high in low suitability areas and low in high suitability areas (Fretwell and Lucas 1969, Fretwell 1972). Categorizing species into these two extremes is challenging as territorial behaviour is uncertain for most species, or it can manifest only under certain circumstances, at the individual or group level, or only within one of the two sexes. Third, the decline in population abundance might not be linear from the centre to the edge (or from high to low suitability), but could rather show no relationship and present a sharp decline at the very edge (Pironon et al. 2017). If this is the case, the test of the theory would be extremely challenging because the same definition of geographic range boundaries is blurry and depends on the methodological approach used and the data available (Fortin et al. 2005). Fourth, populations fluctuate in time because of demographic and environmental stochasticity. This means that a series of repeated measures of density data would be needed in order to characterize the average abundance in a specific location, however abundance time-series are scarce and do not allow to test this hypothesis. Fifth, different species vital rates may respond to different ecological gradients and therefore exhibit different geographical distributions across the species range, ultimately affecting the distribution of its population growth rate or abundance in any or no direction (Pironon et al. 2018). Sixth, an implicit assumption of the abundant-centre theory is that species distributions are in equilibrium with the environment, and thus species geographic ranges reflect their niche tolerance (Colwell and Rangel 2009). However this is a simplification, as current geographic ranges can depend on many historical factors, geographical barriers or even recent human influence (Colwell and Rangel 2009, Di Marco and Santini 2015, Ficetola et al. 2017). For example the geographic range shape can be bounded by a change in habitat type (e.g. forest margin), or biogeographic barriers such as mountains or coastlines (but see Lee-Yaw et al. 2016, Khaliq et al. 2017). An influence of any of these factors would not only violate the premises of the geographical abundant-centre interpretation, but it would also bias the niche suitability measurements. This does not imply a lack of influence of climate and energy availability on large scale patterns of population density (Currie and Fritz 1993, Pettorelli et al. 2009, Santini et al. 2018b), but it rather suggests that the current geographic

range of species is a poor reflection of their Grinnellian niche (Soberon 2007), and suitable conditions for high population abundance can be at the margin of the current geographic range or the estimated environmental space from observed occurrences. Finally, as in the present study, the test of this hypothesis is normally performed on opportunistically collected data. This implies that density records do not cover the whole range of distances from edge-to-centre, or the whole range of environmental suitability values. This, combined with limited sample size per species, can hamper the detection of any large-scale pattern.

While the disagreement between multiple environmental centrality/marginality measures emphasizes the need of testing for multiple measures, it also implies that testing many measures can increase the chance of finding spurious results. Measures should thus be chosen carefully depending on their ecological meaning. Environmental suitability is modelled as the probability of presence, which does not necessarily need to be related to abundance. For example, in source-sink dynamic source habitats are expected to have higher probability of occurrence than sink habitats, but the opposite may be true for abundance (Pulliam 1988). Different environmental centrality/marginality measures can lead to substantially different estimates. The Euclidean and Mahalanobis distances are calculated on the observed distribution only, while the ENFA marginality contrasts used with available habitat. So depending on the delimitation of available habitat, the two groups of variables can strongly differ. Mahalanobis distance also accounts for the covariance of the multiple environmental axes (Farber and Kadmon 2003). As a consequence, depending on the environmental variables considered the two measures can again lead to different estimates. All in all, the Mahalanobis distance seems to be a more appropriate measure of niche marginality (Soberon et al. 2018), and less sensitive to the considered geographic area. However, more important perhaps is the data used to estimate these measures. Using opportunistically collected presence data should be avoided (Martínez-Meyer et al. 2013, Dallas et al. 2017) as it risks to greatly underestimate niche breadth and bias the estimation of niche centre. Further, the environmental variables chosen are crucial, as they are all weighted equally in distance measures, but likely they are not equally important in shaping species distributions. Yet, our knowledge is, in most cases, too limited to identify important variables a priori.

Our data come at a low spatial resolution (1-degree), however this is unlikely to influence our conclusions. In fact, finer resolution data would have allowed controlling for habitat and other local factors and explaining why density estimates were higher or lower, but no centrality pattern would have been observed at a range-wide level either. Indeed, the theory implicitly assumes that the pattern arises at a geographical scale, irrespective of local variations. A serious limitation of ours and previous studies is certainly the uneven and biased data distribution within the species geographic range. We attempted to control for this factor by focusing on the

geographic and environmental coverage, but results did not meet our expectations and appear to be spurious.

The main challenges to test the abundant-centre hypothesis remain the limited number of species on which the theory can be tested, and the uneven and opportunistic distribution of abundance data across species ranges. Our research cannot provide a definitive conclusion about the validity of the abundant-centre hypothesis, as – although to a lesser extent than previous papers – it is affected by some limitations. Yet it does not provide any support to it either, rather it suggests that the support for the hypothesis depends on the variable tested, the geographic area considered, and whether the distribution of the species actually reflects an environmental gradient. Given the variability of our results, we believe that any broad conclusion derived on a small sample of species can be largely misleading. We definitely need more research to better understand the drivers of the intra-specific variation of population density within species range, and the relative importance of factors acting at different spatial scales (Santini et al. 2018b). Finally, we should acknowledge that any measure of geographic or environmental space is only the realization of the fundamental niche of a species (Colwell and Rangel 2009), therefore we might never be able to clearly test the abundance-center hypothesis. So far, the abundant-centre hypothesis remains an appealing speculation with scarce empirical support.

Acknowledgements – We thank Ana Benítez-López for useful discussions on the analyses.

References

- Beck, J. et al. 2013. Online solutions and the “Wallacean shortfall”: what does GBIF contribute to our knowledge of species’ ranges? – *Divers. Distrib.* 19: 1043–1050.
- BirdLife International and Handbook of the Birds of the World 2017. Bird species distribution maps of the world. Version 6.0. – <<http://datazone.birdlife.org/species/requestdis>>.
- Bivand, R. and Rundel, C. 2013. rgeos: interface to geometry engine – open source (GEOS). – R package ver. 0.3-2, R package ver. 0.1-8.
- Blackburn, T. M. et al. 1999a. Geographic gradients in body size: a clarification of Bergmann’s rule. – *Divers. Distrib.* 5: 165–174.
- Blackburn, T. M. et al. 1999b. Do local abundances of British birds change with proximity to range edge? – *J. Biogeogr.* 36: 493–505.
- Boakes, E. H. et al. 2017. Examining the relationship between local extinction risk and position in range. – *Conserv. Biol.* 32: 229–239.
- Borenstein, M. et al. 2009. Introduction to meta-analysis. – John Wiley and Sons.
- Brown, J. H. 1984. On the relationship between abundance and distribution of species. – *Am. Nat.* 124: 255.
- Burgman, M. A. and Fox, J. C. 2003. Bias in species range estimates from minimum convex polygons: implications for conservation and options for improved planning. – *Anim. Conserv.* 6: 19–28.
- Calenge, C. 2006. The package “adehabitat” for the R software: a tool for the analysis of space and habitat use by animals. – *Ecol. Model.* 197: 516–519.
- Chardon, N. I. et al. 2015. Topographic, latitudinal and climatic distribution of *Pinus coulteri*: geographic range limits are not at the edge of the climate envelope. – *Ecography* 38: 590–601.
- Colwell, R. K. and Rangel, T. F. 2009. Hutchinson’s duality: the once and future niche. – *Proc. Natl Acad. Sci. USA* 106: 19651–19658.
- Csergő, A. M. et al. 2017. Less favourable climates constrain demographic strategies in plants. – *Ecol. Lett.* 20: 969–980.
- Currie, D. J. and Fritz, J. T. 1993. Global patterns of animal abundance and species energy use. – *Oikos* 67: 56–68.
- Dallas, T. et al. 2017. Species are not most abundant in the centre of their geographic range or climatic niche. – *Ecol. Lett.* 20: 1526–1533.
- Di Marco, M. and Santini, L. 2015. Human pressures predict species’ geographic range size better than biological traits. – *Global Change Biol.* 21: 2169–2178.
- Dolédec, S. et al. 2000. Niche separation in community analysis: a new method. – *Ecology* 81: 2914–2927.
- Emad, A. and Bailey, P. 2017. wCorr: weighted correlations. – R package ver. 1.9.1, <<https://cran.r-project.org/package=wCorr>>.
- Emlen, J. T. et al. 1986. Density trends and range boundary constraints of forest birds along a latitudinal gradient. – *Auk* 103: 791–803.
- Farber, O. and Kadmon, R. 2003. Assessment of alternative approaches for bioclimatic modeling with special emphasis on the Mahalanobis distance. – *Ecol. Model.* 160: 115–130.
- Ficetola, G. F. et al. 2017. Global determinants of zoogeographical boundaries. – *Nat. Ecol. Evol.* 1: 0089.
- Fortin, M. J. et al. 2005. Species’ geographic ranges and distributional limits: pattern analysis and statistical issues. – *Oikos* 108: 7–17.
- Fretwell, S. D. 1972. Populations in a seasonal environment. – Princeton Univ. Press.
- Fretwell, S. D. and Lucas, H. L. J. 1969. On territorial behaviour and other factors influencing habitat distribution in birds. – *Acta Biotheor.* 19: 16–36.
- Gaston, K. J. and Fuller, R. A. 2009. The sizes of species’ geographic ranges. – *J. Appl. Ecol.* 46: 1–9.
- Gottschalk, T. K. and Huettmann, F. 2011. Comparison of distance sampling and territory mapping methods for birds in four different habitats. – *J. Ornithol.* 152: 421–429.
- Grinnell 1922. The role of the “accidental”. – *Auk* 34: 373–381.
- Guisan, A. and Thuiller, W. 2005. Predicting species distribution: offering more than simple habitat models. – *Ecol. Lett.* 8: 993–1009.
- Hayne, D. 1950. Apparent home range of *Microtus* in relation to distance between traps. – *J. Mammal.* 31: 26–39.
- Hengeveld, R. and Haeck, J. 1982. The distribution of abundance. I. Measurements. – *J. Biogeogr.* 9: 303–316.
- Hijmans, R. J. and van Erten, J. 2014. raster: geographic data analysis and modeling. – R package ver. 2.3-12, <<http://CRAN.R-project.org/package=raster>>.
- Hijmans, R. J. et al. 2015. Geosphere: spherical trigonometry. – R package ver. 1.3-11.
- Hirzel, A. H. et al. 2002. Ecological-niche factor analysis: how to compute habitat-suitability maps without absence data? – *Ecology* 83: 2027–2036.
- IUCN 2017. The IUCN Red List of threatened species. Version 2017-1. – <www.iucnredlist.org> accessed 1 May 2017.

- Johnston, A. et al. 2017. Estimates of observer expertise improve species distributions from citizen science data. – *Methods Ecol. Evol.* 9: 88–97.
- Karger, D. N. et al. 2016. Climatologies at high resolution for the Earth land surface areas. – arXiv:1607.00217.
- Khaliq, I. et al. 2017. The influence of thermal tolerances on geographical ranges of endotherms. – *Global Ecol. Biogeogr.* 26: 650–668.
- Klein Goldewijk, K. et al. 2011. The HYDE 3.1 spatially explicit database of human-induced global land-use change over the past 12,000 years. – *Global Ecol. Biogeogr.* 20: 73–86.
- Lee-Yaw, J. A. et al. 2016. A synthesis of transplant experiments and ecological niche models suggests that range limits are often niche limits. – *Ecol. Lett.* 19: 710–722.
- Leimu, R. and Koricheva, J. 2004. Cumulative meta-analysis: a new tool for detection of temporal trends and publication bias in ecology. – *Proc. R. Soc. B* 271: 1961–6.
- Lewin-Koh, N. and Bivand, R. 2011. mapproj: tools for reading and handling spatial objects. – R packag. ver. 0.8-10.
- Martínez-Meyer, E. et al. 2013. Ecological niche structure and rangewide abundance patterns of species. – *Biol. Lett.* 9: 20120637.
- Meyer, C. et al. 2016. Multidimensional biases, gaps and uncertainties in global plant occurrence information. – *Ecol. Lett.* 19: 992–1006.
- Paradis, E. et al. 1999. Dispersal and spatial scale affect synchrony in spatial population dynamics. – *Ecol. Lett.* 2: 114–120.
- Parker, T. H. et al. 2016. Transparency in ecology and evolution: real problems, real solutions. – *Trends Ecol. Evol.* 31: 711–719.
- Pettorelli, N. et al. 2009. Energy availability and density estimates in African ungulates. – *Am. Nat.* 173: 698–704.
- Pinheiro, J. C. and Bates, D. M. 2000. Mixed-effects models in S and S-Plus. – Springer Science and Business Media.
- Pironon, S. et al. 2015. Do geographic, climatic or historical ranges differentiate the performance of central versus peripheral populations? – *Global Ecol. Biogeogr.* 24: 611–620.
- Pironon, S. et al. 2017. Geographic variation in genetic and demographic performance: new insights from an old biogeographical paradigm. – *Biol. Rev. Camb. Phil. Soc.* 92: 1877–1909.
- Pironon, S. et al. 2018. The ‘Hutchinsonian niche’ as an assemblage of demographic niches: implications for species geographic ranges. – *Ecography* 41: 1103–1113.
- Pulliam, H. R. 1988. Sources, sinks, and population regulation. – *Am. Nat.* 132: 652–661.
- Renzulli, C. et al. 1980. The effects of trapping design on demographic estimates in the meadow vole, *Microtus pennsylvanicus*. – *Am. Midl. Nat.* 104: 397–401.
- Sagarin, R. D. and Gaines, S. D. 2002. The “abundant centre” distribution: to what extent is it a biogeographical rule? – *Ecol. Lett.* 5: 137–147.
- Sagarin, R. D. et al. 2006. Moving beyond assumptions to understand abundance distributions across the ranges of species. – *Trends Ecol. Evol.* 21: 524–530.
- Santini, L. et al. 2013. Ecological correlates of dispersal distance in terrestrial mammals. – *Hystrix* 24: 181–186.
- Santini, L. et al. 2018a. TetraDENSITY: a database of population density estimates in terrestrial vertebrates. – *Global Ecol. Biogeogr.* 27: 787–791.
- Santini, L. et al. 2018b. Global drivers of population density in terrestrial vertebrates. – *Global Ecol. Biogeogr.* 27: 968–979.
- Santini, L. et al. 2018c. Data from: Addressing common pitfalls does not provide more support to geographical and ecological abundant-centre hypotheses. – Figshare Digital Repository, <<https://dx.doi.org/10.6084/m9.figshare.7117880.v1>>.
- Seddon, P. J. et al. 2003. A comparison of derived population estimate, mark–resighting and distance sampling methods to determine the population size of a desert ungulate, the Arabian oryx. – *Oryx* 37: 286–294.
- Sexton, J. P. et al. 2009. Evolution and ecology of species range limits. – *Annu. Rev. Ecol. Evol. Syst.* 40: 415–436.
- Soares, T. N. et al. 2015. Patterns of genetic variability in central and peripheral populations of *Dipteryx alata* (Fabaceae) in the Brazilian Cerrado. – *Plant Syst. Evol.* 301: 1315–1324.
- Soberon, J. 2007. Grinnellian and Eltonian niches and geographic distribution of species. – *Ecol. Lett.* 10: 1115–1123.
- Soberon, J. et al. 2018. A comment on “Species are not most abundant in the center of their geographic range or climatic niche”. – bioRxiv doi:<http://dx.doi.org/10.1101/266510>
- Soule, M. 1973. The epistasis cycle: a theory of marginal populations. – *Annu. Rev. Ecol. Syst.* 4: 165–187.
- Sutherland, G. D. et al. 2000. Scaling of natal dispersal distances in terrestrial birds and mammals. – *Conserv. Ecol.* 4: 16.
- Thuiller, W. et al. 2009. BIOMOD – a platform for ensemble forecasting of species distributions. – *Ecography* 32: 369–373.
- Thuiller, W. et al. 2014. Does probability of occurrence relate to population dynamics? – *Ecography* 37: 1155–1166.
- Van Couwenberghe, R. et al. 2013. Can species distribution models be used to describe plant abundance patterns? – *Ecography* 36: 665–674.
- VanDerWal, J. et al. 2009. Abundance and the environmental niche: environmental suitability estimated from niche models predicts the upper limit of local abundance. – *Am. Nat.* 174: 282–291.
- Viechtbauer, W. 2010. Metafor: meta-analysis package for R. – R package ver. 1.4-0.
- Weber, M. M. et al. 2017. Is there a correlation between abundance and environmental suitability derived from ecological niche modelling? A meta-analysis. – *Ecography* 40: 817–828.
- Wilman, H. et al. 2014. EltonTraits 1.0: species-level foraging attributes of the world’s birds and mammals. – *Ecology* 95: 2027–2027.

Supplementary material (Appendix ECOG-04027 at <www.ecography.org/appendix/ecog-04027>). Appendix 1.