Check for updates

Received: 10 March 2020 | Revised: 19 July 2021 | Accepted: 22 July 2021

DOI: 10.1111/geb.13376

RESEARCH ARTICLE

Global Ecology and Biogeography A Journal of Macroecology WILEY

Using a robust multi-settings inference framework on published datasets still reveals limited support for the abundant centre hypothesis: More testing needed on other datasets

Mathieu Chevalier[1,2] | Olivier Broennimann[1,3] | Antoine Guisan[1,3]

[1]Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland

[2]IFREMER, Centre de Bretagne, DYNECO, Laboratoire d'Ecologie Benthique Cotière (LEBCO), Plouzané, France

[3]Institute of Earth Surface Dynamics, University of Lausanne, Lausanne, Switzerland

Correspondence
Mathieu Chevalier, FREMER-DYNECO-LEBCO, Centre de Bretagne, CS 10070, 29280 Plouzané, France.
Email: mathieu.chevalier38@gmail.com

Funding information
Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung, Grant/Award Number: CR23I2_162754

Editor: Franziska Schrodt

Abstract

Aim: The abundant centre hypothesis (ACH) predicts a negative relationship between species abundance and the distance to the geographical range centre. Since its formulation, empirical tests of the ACH have involved different settings (e.g. the distance to the ecological niche or to the geographical range centre), but studies found contrasting support for this hypothesis. Here, we evaluate whether these discrepancies might stem from differences regarding the context in which the ACH is tested (geographical or environmental), how distances are measured, how species envelopes are delineated, how the relationship is evaluated and which data are used.

Location: The Americas.

Time period: 1800–2017.

Major taxa studied: Mammals, birds, fish, and tree seedlings.

Methods: Using published abundance data for 801 species, together with species range maps, we tested the ACH using three distance metrics in both environmental and geographical spaces with range and niche envelopes delineated using two different algorithms, totalling 12 different settings. We then evaluated the distance–abundance relationship using correlation coefficients (traditional approach) and mixed-effect models to reduce the effect of sampling noise on parameter estimates.

Results: Similar to previous studies, correlation coefficients indicated an absence of effect of distance on abundance for all taxonomic groups and settings. In contrast, mixed-effect models highlighted relationships of various strengths and shapes, with a tendency for more theoretically supported settings to provide stronger support for the ACH. The relationships were however not consistent across taxonomic groups and settings, and were sometimes even opposite to ACH expectations.

Main conclusions: We found mixed and inconclusive results regarding the ACH. These results corroborate recent findings, and suggest either that our ability to predict abundances from the location of populations within geographical or environmental spaces is low, or that the data used here have a poor signal-to-noise-ratio.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2021 The Authors. *Global Ecology and Biogeography* published by John Wiley & Sons Ltd.

The latter calls for further testing on other datasets using the same range of settings and methodological framework.

KEYWORDS

abundance, Bayesian inference, centroid, convex hull, distance, ellipsoids, kernel density, margins, mixed-effect models, nonlinearity

1 | INTRODUCTION

The assumption that species are most abundant in the centre of their range and decline in abundance toward the range edges – the so-called abundant centre hypothesis (ACH) – has a long history in the ecological literature (Sagarin & Gaines, 2002; Sagarin et al., 2006). In recent years some studies found support for the ACH (Martínez-Meyer et al., 2013; Osorio-Olvera et al., 2020; Yañez-Arenas et al., 2012), but others (Dallas et al., 2017; Santini et al., 2019) found very low or no support, casting doubts about the generality of this hypothesis. The controversy was recently revived following the study published by Dallas et al. (2017) highlighting consistent near-zero correlation between population abundance and the distance to the climatic niche or the geographical range centre for more than 1,400 species.

The ACH is implicitly grounded in the niche theory but was originally developed to explain distribution–abundance patterns within species geographical ranges (Brown, 1984). It relies on the main assumption that the range of a species is a geographical representation of its ecological niche (sensu Hutchinson, 1957), so that environmental conditions are optimal near the centre of the range and harsher at the periphery (Brown, 1984). Since then, several studies have shown that the centre of the geographical space does not necessarily correspond to the centre of the environmental space and that geographical and environmental gradients are not necessarily concordant (Hutchinson's niche-biotope duality; Colwell & Rangel, 2009; Soberón & Nakamura, 2009). This suggests that a different assessment of the ACH can be obtained depending on the space considered.

In recent years, several studies found stronger support for an environmental declination of the ACH (Martínez-Meyer et al., 2013; Osorio-Olvera et al., 2020; Yañez-Arenas et al., 2012; but see Dallas, Pironon, et al., 2020). However, this assessment is often based on an unclear characterization of species envelopes (and thus niches). For instance, envelopes are often characterized by projecting the geographical coordinates of species' observations (e.g. occurrence or abundance) in a two-dimensional environmental space characterizing variation in the global climate (Dallas et al., 2017; Santini et al., 2019). If these observational data were collected over the full species range, then envelopes could be considered as an estimate of the realized (also called ecological) niche, that is, the set of suitable abiotic conditions that are accessible to the species (Barve et al., 2011) constrained by biotic interactions (Jackson & Overpeck, 2000; Soberón & Nakamura, 2009). However, in many cases, observations only cover a subset of species ranges, implying that these realized niches are truncated (Chevalier et al., 2021; Dallas & Hastings, 2018; Osorio-Olvera et al., 2020; Yañez-Arenas et al., 2020). While this issue can be minimized by considering the full species range (Santini et al., 2019; Soberón et al., 2018), the ACH could more ideally be tested relative to the centre of the fundamental niche (Yañez-Arenas et al., 2020), which is defined in terms of population fitness as a function of abiotic conditions in the absence of biotic interactions and dispersal constraints (Austin, 1999; Hutchinson, 1957). Accordingly, and provided that abundance is an expression of fitness (which is not necessarily the case; see e.g. McGill, 2012; Samis & Eckert, 2009), then, a negative distance–abundance relationship should be observed (Osorio-Olvera et al., 2019). However, although some evidence suggests that species' fundamental environmental niches are convex in shape (Peterson et al., 2011; Soberón & Peterson, 2019), obtaining robust estimates using correlative methods has proved challenging (Peterson et al., 2011).

Geographical range and climatic niche envelopes are often delineated using the minimum convex polygon around sample points i.e. the convex hull (CH; Broennimann et al., 2007; Dallas et al., 2017). However, geographical ranges are usually irregular and fragmented, and thus potentially concave (Brown et al., 1996; Soberón & Nakamura, 2009), a type of shape that cannot be accommodated by the CH (Blonder et al., 2014). Delineating these complex geometries using more appropriate algorithms such as kernel density estimators (KDEs; Blonder et al., 2014; Broennimann et al., 2012) could yield new insights about the ACH. Furthermore, while convexity is theoretically expected for climatic niche envelopes (Soberon & Arroyo-Peña, 2017), CHs are sensitive to outliers (Blonder et al., 2014; Soberón et al., 2018), implying that the volume and the shape of CH-based climatic envelopes can be biased. Minimum volume ellipsoids (MVEs Van-Aelst & Rousseeuw, 2009) have been recently proposed as putatively closer estimates of fundamental niches than CHs (Escobar et al., 2018; Osorio-Olvera et al., 2020; Qiao et al., 2016; Soberón, 2019; Soberón & Peterson, 2019; Yañez-Arenas et al., 2018). Yet whether this assertion holds true remains difficult to test given the absence of data allowing the quantification of fundamental niches (Soberon & Arroyo-Peña, 2017).

Most ACH studies mainly focused on the distance to centres with little attention to margins (Santini et al., 2019). While these two distance metrics convey the same information for circular envelopes, this is not the case for non-circular envelopes where populations close to the centre are not necessarily far from margins (Blonder et al., 2014; Broennimann et al., 2021; Santini et al., 2019). In such cases, the distance to margins might be a better descriptor of the suitability gradient (Broennimann et al., 2021). The Mahalanobis

distance, which unlike the Euclidean distance considers the variance and correlation between environmental variables, has also been proposed to more accurately describe the distance to the centre of the environmental space (Soberón et al., 2018). These two distance metrics have however rarely been considered to test the ACH (Osorio-Olvera et al., 2020; Santini et al., 2019).

Abundances have traditionally been assumed to linearly decline as the distance from the centre increases (Brown, 1984), but recent evidence suggests that the relationship, although globally negative, may actually be nonlinear (Pironon et al., 2017; Santini et al., 2019). Such nonlinear effects have often been explained by a violation of another assumption of the ACH stating that environmental suitability is a continuously decreasing function of the distance to the centre (Brown, 1984). More recently, other factors have been advanced to explain these nonlinear relationships including biotic interactions (e.g. ideal despotic distribution), threshold effects, Allee effects or a nonlinear spatial structure of the environment (Osorio-Olvera et al., 2019; Pironon et al., 2017; Santini et al., 2019). Depending on the factor, different shapes, both concave and convex, can be expected. To date, few studies (Dallas et al., 2017; Osorio-Olvera et al., 2020) have considered the potential for a nonlinear distance–abundance relationship and most of them have used rank correlations (Spearman; but see Yancovitch Shalom et al., 2020), a measure that provides no information regarding the shape of the relationship.

The ACH has mostly been tested individually for multiple species, with the overall trend and agreement across species assessed by taking the mean of species effect sizes, regardless of the number of populations per species (Dallas et al., 2017; Yancovitch Shalom et al., 2020; but see Santini et al., 2019). However, not all species are equally informative about the distance–abundance relationship. For instance, data collected for rare or difficult to sample species can present a considerable level of noise, which can blur the overall relationship (Brown, 1984). Hierarchical (mixed-effect) models are emerging as powerful tools to reduce the effect of sampling noise on parameter estimates, and have been shown to improve both inferences (Schaub & Kéry, 2012) and forecasts (Chevalier & Knape, 2019). Given the heterogeneity of the data usually considered in ACH assessments (Dallas et al., 2017; Osorio-Olvera et al., 2020), this framework is expected to provide more robust estimates of distance–abundance relationships.

Another argument that is often advanced to explain the mixed support for the ACH relates to the quality of the data (Dallas, Pironon, et al., 2020; Knouft, 2018; Soberón et al., 2018). Two aspects are notably discussed: (a) that the centre and the margins of geographical ranges and climatic niches are biased because they are often estimated using populations covering only a subset of geographical ranges and (b) that abundance data are inappropriate owing to the use of non-standardized protocols or due to the special status of some species (e.g. invasive species that are not at equilibrium with their environment, migratory species). This aspect is particularly important as the use of different datasets could yield varying support for the ACH.

Here, we aim to re-analyse the dataset published by Dallas et al. (2017) for birds, mammals, fishes and trees, while accounting for the reported biases, to assess whether different conclusions could be obtained when using different methodological settings (see below). For birds, we also consider another dataset collected under a robust sampling protocol [North American Breeding Bird Survey (BBS) dataset; Osorio-Olvera et al., 2020] to evaluate the effect of the quality of abundance data on the distance–abundance relationship. We evaluate the support for the ACH both globally and for each species at the same time, using two different statistical frameworks: the traditional one based on correlation coefficients and another one based on hierarchical Bayesian nonlinear mixed-effect models. Using abundance data for 801 species totalling 357,317 records (Dallas et al., 2017; Osorio-Olvera et al., 2020) together with species range maps and related environmental information (Rocchini et al., 2011), we tested the ACH in both the geographical and the environmental space. In each space, we expressed the suitability gradient using three distance metrics (Euclidean and Mahalanobis distance to the centre, distance to margins), with two methods to delineate species envelopes in both geographical (CH and KDE) and environmental (CH and MVE) spaces. Overall, this led to the testing of the ACH under 12 different settings.

2 | MATERIAL AND METHODS

2.1 | Datasets

2.1.1 | Climatic data

Climatic data were obtained in the form of 19 bioclimatic variables from WorldClim (http://www.worldclim.org/), at a spatial resolution of 10 minutes (*c.* 340 km^2). From these variables, we performed a principal component analysis (PCA) from which we extracted the two first axes, which explained 55 and 19% of the total variance, respectively. PCA axes were then rasterized into climatic pixels with a spatial resolution of 10 minutes.

2.1.2 | Abundance data

We used the data described in Dallas et al. (2017), in which estimates of population density for mammals, birds, fish, and tree seedlings were extracted from different databases based on published work (Knouft & Anthony, 2016; Sullivan et al., 2009; Thibault et al., 2011; Woudenberg et al., 2010). For all groups, species abundance was estimated as the number of individuals within a sampling area, standardized by either sampling area or sampling intensity, thus representing a measure of population density. Owing to concerns expressed regarding the bird dataset (Soberón et al., 2018), which is based on a citizen science programme (eBird), we also tested the ACH using the BBS dataset, which follows a standardized survey protocol. This evaluation allowed us to assess the effect of the quality

of the sampling protocol on the estimated distance–abundance relationship. The BBS dataset was extracted from Osorio-Olvera et al. (2020). For further details, see Dallas et al. (2017) and Osorio-Olvera et al. (2020). Similarly, because for trees, the measure of density only included one age class that may not be representative of the whole population density, we also report global results with tree seedlings excluded.

For all species, when several populations fell within the same climatic pixel (*c*. 340 km^2), we averaged abundances across populations so that only one population was considered per pixel (i.e. we thinned the data; Knouft, 2018; Soberón et al., 2018). After this thinning procedure, we only retained species presenting at least 10 populations. Our analyses are thus based on 801 species, including 626 species of birds (599 species for the eBird dataset, 389 species for the BBS dataset – 362 species in common between the two datasets), 28 species of fish, 27 species of mammals and 120 species of trees. In our final dataset, the number of populations varied from 10 to 4,444 depending on species.

2.1.3 | Geographical range maps

The above-mentioned abundance data (including the BBS dataset) do not encompass the full geographical range for most species, potentially leading to niche truncation issues and biased location of both the centres and the margins of geographical and environmental spaces (Dallas, Pironon, et al., 2020; Knouft, 2018; Soberón et al., 2018). Thus, contrary to previous studies that built envelopes using the location of observed populations, we here delineated species envelopes using the entire distribution of species. For birds, mammals and fish, we used International Union for Conservation of Nature (IUCN) range maps (https://www.iucnredlist.org/), only considering the resident range for mammals and fish, but also including the breeding range for birds owing to the presence of migratory species in both the eBird (n = 246 species; 41% of the original number) and the BBS (n = 196 species; 50% of the original number) datasets. However, since the migratory status has been shown to have an influence on ACH assessment (Osorio-Olvera et al., 2020; but see Dallas et al., 2017), we also report results without migratory species. For trees, we used Little's distribution maps (Prasad & Iverson, 2003). In all cases, range maps were rasterized at the same spatial resolution as climatic data (i.e. 10 minutes).

2.2 | Defining geographical range and climatic niche envelopes

To obtain envelopes in the geographical space, we extracted the coordinates of pixels falling within distribution ranges. Following the traditional framework, we delineated species envelopes using the CH around the range coordinates (Figure 1). To account for the irregular and fragmented shapes of distribution ranges, we further delineated envelopes using a multidimensional KDE procedure (Blonder et al., 2014; Figure 1). The bandwidth of the KDE was estimated from the data using a multivariate generalization of the univariate plug-in bandwidth selector (Wand & Jones, 1994) and species envelopes were defined as the minimum threshold of probability density that included 99% of populations. To obtain envelopes in the environmental space, we extracted the two-dimensional coordinates (niche values) corresponding to the location of population estimates. Species envelopes were delineated using the CH and the MVE that included 99% of the coordinates (Figure 1). The MVE is expected to provide an estimate potentially closer to the fundamental niche than the CH.

For each envelope, we computed its centroid as the mean of the vertices and considered it as the centre of the geographical range or of the climatic niche. Note that because disjunctions in the hypervolume are possible with the KDE, envelopes can be characterized by multiple centroids.

2.3 | Distance metrics

To calculate distances, we projected observed populations within the estimated species envelopes. For CH envelopes, the distance to the geographical centre was calculated using the Haversine distance (i.e. Euclidean distance accounting for the curvature of the Earth), while the distance to the climatic niche centre was calculated using the Euclidean distance. We also used Euclidean distances to compute the distance to the centre of the MVE. For KDE envelopes, we computed distances similarly but in cases where species envelopes were composed of several polygons, using the centroid of the polygon in which the focal population was located. For all envelopes, we also computed the distance to the geographical and the climatic centre using the Mahalanobis distance, while the distance to the margins was calculated for each population as the orthogonal distance to the closest margin.

In all cases, distances were standardized by the maximum observed distance to ensure comparability across species and settings. Hence, values close to one indicate populations far from the centre (or from the margins depending on the type of distance considered) and values close to zero populations close to the centre (or to the margins).

2.4 | Distance–abundance relationship

We evaluated the distance–abundance relationship both globally and individually for each species using two different statistical frameworks.

For the traditional framework, we computed Spearman (to account for nonlinear effects) correlation coefficients for each species and then averaged the coefficients across species to obtain a global estimate of the relationship at the scale of the taxonomic group for each setting (Dallas et al., 2017; Osorio-Olvera

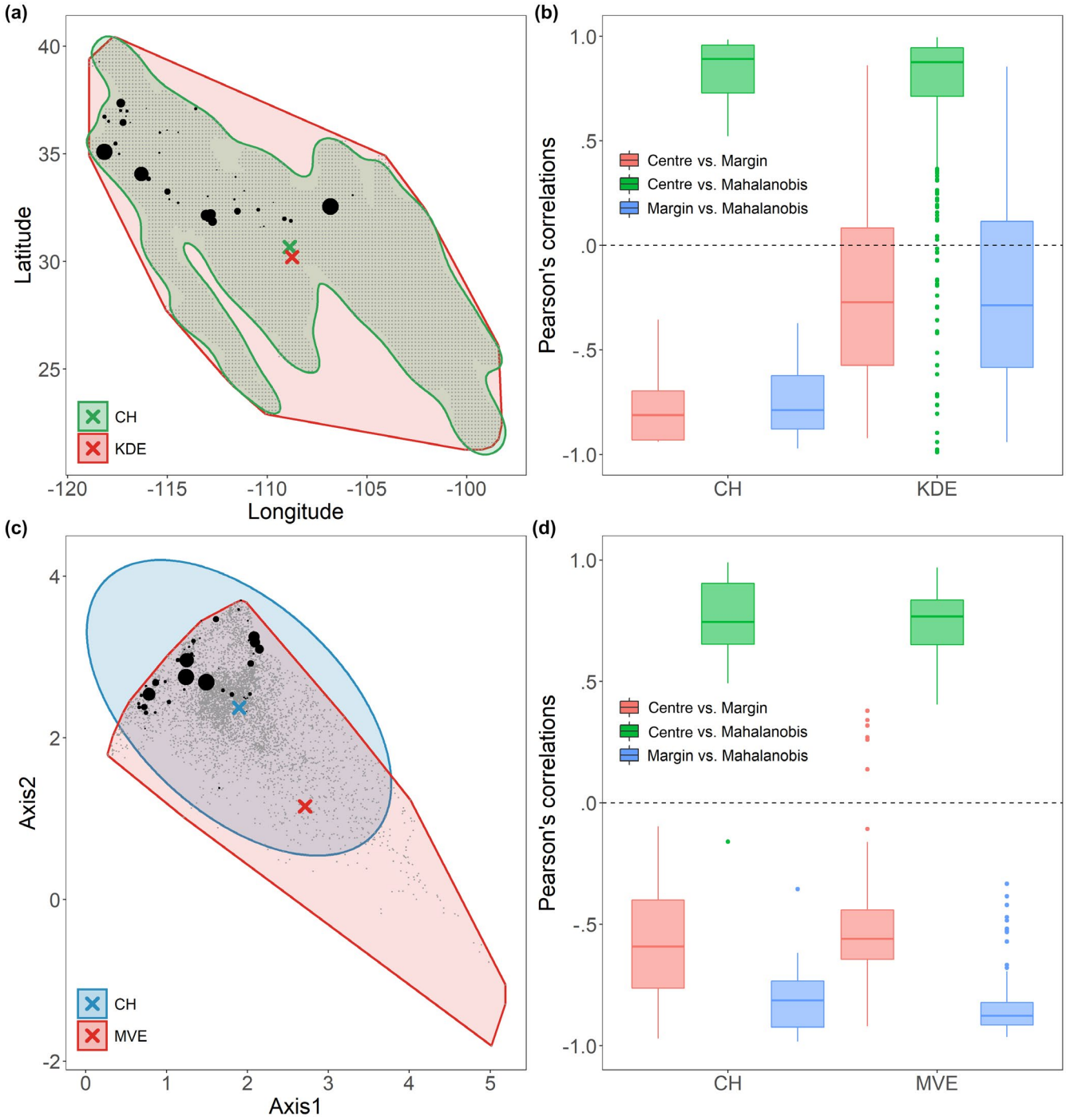

FIGURE 1 Illustration of the differences between the two envelopes in the geographical (a) and the environmental (c) space for the species *Dipodomys merriami* along with a measure of the correlation between the three distance metrics (across all species) in both spaces (b, d). In (a) and (c), grey dots represent the coordinates of pixels falling within species ranges and that were used to delineate species envelopes, whereas black dots represent the location of observed populations within each envelope with size proportional to the estimated population density. For all envelopes, centres are denoted with a cross. CH, convex hull; KDE, kernel density estimator; MVE, minimum volume ellipsoid

et al., 2020). We then tested whether averaged values differed from zero using Wilcoxon one-sample signed-rank tests. While this approach is expected to produce better species-specific estimates than a hierarchical approach when using sound datasets (e.g. standardized sampling protocol; Chevalier & Knape, 2019), this is not the case when using heterogeneous and noisy datasets that are prone to small sample sizes and measurement errors. By considering that species are not independent and potentially share some commonalities, for example due to large-scale ecological determinants such as climate, the hierarchical approach is less sensitive to noise and is expected to provide more robust estimates (Harrison et al., 2018).

The second framework is based on hierarchical mixed-effect models (Burnham & Anderson, 2002; Gelman et al., 2013). Mixed-effect models are composed of both fixed and random effects and are commonly used when measurements are made on clusters (e.g. taxonomic group) of related statistical units (e.g. species). This approach is particularly useful in settings with noisy data because it makes it possible to reduce the effect of sampling noise and to refine individual estimates by borrowing information across statistical units (Gelman et al., 2013; Harrison et al., 2018; Kruschke, 2014). This borrowing of information is made possible by using higher-level distributions with means representing commonalities shared by all species (fixed effects) and standard deviations representing departures from the means (random effects) that allow the model to partly adapt to features specific to the species. Thus, while each species is characterized by its own set of parameters (providing species-specific relationships), the relationships are assumed to come from the same higher-level distribution and therefore mutually inform each other. This hierarchical structure implies that species with a large number of observations have a stronger influence on the estimation of fixed effects and that species estimates tend to be pulled toward the fixed effect, a process known as shrinkage (Gelman et al., 2013; Kruschke, 2014). In general, the shrinkage is strongest for species with small sample sizes, as the paucity of within-species information to estimate fixed-effects is counteracted by the model using data from other species to improve the precision of the estimate. For each group, we investigated the distance–abundance relationship under the 12 different settings using Bayesian nonlinear mixed-effect models including species-specific effects (denoted *s*) on intercepts (α_s), slopes (β_s) and quadratic terms (γ_s). These three sets of species-specific parameters were assumed to follow normal distributions with hyper-parameters μ_α, μ_β and μ_γ representing the average value of the three parameters at the group scale (fixed effects) and hyper-parameters σ_α, σ_β and σ_γ representing among-species variation in parameter values (random effects). Our model thus includes fixed and random effects on intercepts, slopes and quadratic terms. A formal description of the model is provided in Supporting Information Text S1.

To evaluate the strength and the shape of the distance–abundance relationship at the scale of each taxonomic group along with its uncertainty, we extracted posterior samples of model parameters and computed the medians and the 95% highest posterior density (HPD_{95}) intervals of fixed parameters (μ_β for the slope and μ_γ for the quadratic term). For each setting and taxonomic group, the confidence about the estimated relationship (and thus the support for the ACH) was assessed by computing the posterior probability for a negative relationship (or positive if the distance to margins is considered). Since the model is nonlinear, the direction of the relationship must be evaluated by considering the sign of both the linear and the quadratic term. Given that the range of distance values is comprised between 0 and 1, then, when $\mu_\beta < 0$, the relationship is negative if $\mu_\gamma < |\mu_\beta|$ but positive if $\mu_\gamma > (-\mu_\beta)$. In contrast, if $\mu_\beta > 0$ the relationship is negative if $\mu_\gamma < (-\mu_\beta)$ but positive if $\mu_\gamma > |\mu_\beta|$. In practice, we computed the posterior probability for expected relationships (i.e. negative if the distance to the centre is considered, positive otherwise) depending on the sign of the posterior median of the linear term (μ_β). For instance, when considering the distance to the centre (Mahalanobis or Euclidean), we computed the posterior probability for the expected relationship (i.e. negative) as $P\,[\mu_\gamma < (-\mu_\beta)] = \frac{N^{samp}_{\mu_\gamma < (-\mu_\beta)}}{N^{samp}}$ (where $N^{samp}_{\mu_\gamma < (-\mu_\beta)}$ represents the number of Markov Chain Monte Carlo (MCMC) samples with $\mu_\gamma < (-\mu_\beta)$ and where N^{samp} represents the total number of MCMC samples) if the posterior median of μ_β was positive and as $P[\mu_\gamma < |\mu_\beta|] = \frac{N^{samp}_{\mu_\gamma < |\mu_\beta|}}{N^{samp}}$ (where $N^{samp}_{\mu_\gamma < |\mu_\beta|}$ represents the number of MCMC samples with $\mu_\gamma < |\mu_\beta|$) if the posterior median of μ_β was negative. If the distance to the margins is considered, the posterior probability for the expected relationship (i.e. positive) is thus equal to $P[\mu_\gamma > |\mu_\beta|] = 1 - P[\mu_\gamma < |\mu_\beta|]$ if $\mu_\beta < 0$ and $P[\mu_\gamma > (-\mu_\beta)] = 1 - P[\mu_\gamma < (-\mu_\beta)]$ if $\mu_\beta > 0$. Regardless of the type of distance considered, if the posterior probability for the expected relationship had a value above .9 (meaning a 90% probability), we considered this has strong support for the ACH. In contrast, if this probability had a value below .1 (meaning a 10% probability) this would indicate a strong support for an unexpected relationship (since $P_{positive} = 1 - P_{negative}$; then $P_{unexpected} = 1 - P_{expected} = 90\%$ probability for the opposite trend) and therefore a low support for the ACH. We conducted a similar analysis using the posterior distribution of species-specific coefficients (β_s and γ_s) to estimate the proportion of species presenting the expected relationship under the different settings.

For each model, we used posterior predictive checks (Gelman et al., 1996) to assess their goodness of fit. Specifically, we calculated the sum of squared standardized Pearson residuals for both the observed data and a replicated dataset derived from model estimates. From this χ^2 discrepancy metric, we computed the Byesian *p*-value, which quantifies the proportion of posterior samples in which the distance of observed data to the model is greater than the distance of replicated data to the model. Values close to .5 suggest a good model fit, whereas values close to 0 or 1 indicate a lack of fit. These Bayesian *p*-values provide information about the plausibility of a model given the data. To obtain information regarding the proportion of variance explained we computed the marginal (variance explained by fixed effects only) and the conditional (variance explained by fixed and random effects) R^2 of each model (Nakagawa & Schielzeth, 2013).

To investigate the potential for higher-order terms (i.e. higher degree of nonlinearity) in the distance–abundance relationship we also evaluated the distance–abundance relationship using generalized additive mixed-effect models (Hastie & Tibshirani, 1986), allowing up to three degrees of freedom for smoothing terms.

3 | RESULTS

3.1 | Influence of methodological settings on distance measures

The three distance metrics were only moderately correlated under the different settings, confirming that they do not convey the same

information in either the geographical or the environmental space (Figure 1). Our results further confirm that the method used to delineate species envelopes (CH or KDE in the geographical space; CH or MVE in the environmental space) yields distinct estimates of the relative position of populations to both the centre and the margins (Figure 1).

3.2 | Influence of the statistical framework on ACH evaluation

Using the traditional framework (i.e. correlation coefficients averaged across species), we found little support for the ACH, regardless of the methodological setting or the taxonomic group ($\overline{\rho}_{\text{mammals}} = .015$, $\overline{\rho}_{\text{eBird}} = -.01$, $\overline{\rho}_{\text{fish}} = .03$ and $\overline{\rho}_{\text{trees}} = .005$; Figure 2 – note that the signs of coefficients for margins were reversed before averaging in order to be comparable to the two other measures of distance). While the correlations were larger for the BBS dataset and were all statistically different from zero, they were still weak ($\overline{\rho}_{\text{BBSbirds}} = -.09$; Figure 2). Regarding the alternative framework (i.e. mixed effect models), all models fitted the data well (Bayesian *p*-values close to .5), thus suggesting that the data are equally plausible under the different settings (Supporting Information Table S1). Importantly, we found strong evidence for non-null relationships under most settings with this framework. For all groups, the proportion of variance explained was rather high (conditional R^2_{mammals}=.23, $R^2_{\text{eBird}} = .41$, $R^2_{\text{fish}} = .61$, $R^2_{\text{trees}} = .54$, $R^2_{\text{BBSbirds}} = .86$) and was comparable across models (except for the tree and the eBird dataset), indicating that all models performed well in explaining abundance variations across species (Supporting Information Table S1). However, the models greatly differed regarding their support for the ACH. Similar results were obtained with generalized additive mixed-effect models (Supporting Information Figure S1).

3.3 | Influence of methodological settings on ACH evaluation

Over the 60 relationships tested (5 datasets × 12 settings), 23 (38%) showed support for the ACH (expected relationships, Figure 3a), 18 of which were related to the two bird datasets. In contrast, seven relationships (11.6%) presented trends opposite to ACH expectations (i.e. unexpected relationships, Figure 3a). When excluding tree seedlings, the percentage of expected relationships increased to 45.8% whereas the percentage of unexpected relationships dropped to 6.2% (Supporting Information Figure S2a). A similar pattern was observed when removing migratory birds (Supporting Information Figure S3a). Overall, the proportion of expected relationships was larger in the geographical space (46%) than in the environmental space (30%) while the proportion of unexpected relationships was lower in the former (10%) than in the latter (13%) (Figure 3b). These trends were reinforced when excluding migratory birds (Supporting Information Figure S3b) or tree seedlings (Supporting Information Figure S2b). Regarding envelopes, the proportion of expected relationships was higher with the MVE (33%) than with the CH (26%) in the environmental space whereas the same proportions were obtained with the CH and the KDE in the geographical space (46%; Figure 3c). However, a much higher proportion of unexpected relationships was obtained with the CH (13% in the geographical space, 26% in the environmental space) than with the KDE (6.6%) or the MVE (0%). In both spaces, we found a larger proportion of expected relationships with the Mahalanobis distance (45%) and the distance to the margins (55%) than with the Euclidean distance (15%; Figure 3d). The proportion of unexpected relationships was also larger with the Euclidean distance (15%) than with the two other measures of distance (10% each). Similar patterns were obtained when excluding tree seedlings or migratory birds for both envelopes (Supporting Information Figures S2c and S3c) and distance measures (Supporting Information Figures S2d and S3d).

3.4 | Influence of the dataset (taxonomic group) on ACH evaluation

All taxonomic groups showed support for the ACH under some settings. However, the relationships were not always consistent and were sometimes unexpected (Figure 3e). For mammals, we found support for the ACH for three out of the twelve settings but also found evidence for one unexpected relationship (Figures 3e, 4, Supporting Information Table S1). The number of expected (unexpected) relationships were one (two) for the fish dataset, one (four) for the tree dataset, ten (zero) for the eBird dataset and eight (zero) for the BBS dataset (Figures 3e, 4, Supporting Information Table S1). For the two bird datasets, removing migratory species yielded stronger support for the ACH under most settings (Supporting Information Figures S3e and S4, Table S2).

3.5 | Among-species variation in distance–abundance relationships

Despite evidence for global patterns, large variations were visible across species (Figure 5a, Supporting Information Figures S5–S9). Over the 60 relationships tested, 33% of species displayed expected relationships whereas 8.4% displayed unexpected relationships and 58.6% presented undetermined relationships. These proportions greatly varied across settings and datasets but confirmed the observed trends. Only marginal changes were observed when excluding tree seedlings (Supporting Information Figure S10a). Larger changes were detected after removing migratory birds with, for example, the proportion of expected and unexpected relationships increasing to 40.5 and 9.7%, respectively (Supporting Information Figure S11a). On average (considering all settings), 52% of the birds displayed the expected relationship with the BBS dataset (8% of unexpected relationships), whereas this proportion only reached 26% for the eBird dataset (3.3% of unexpected relationships; Figure 5b). When excluding migratory species, the proportion of expected

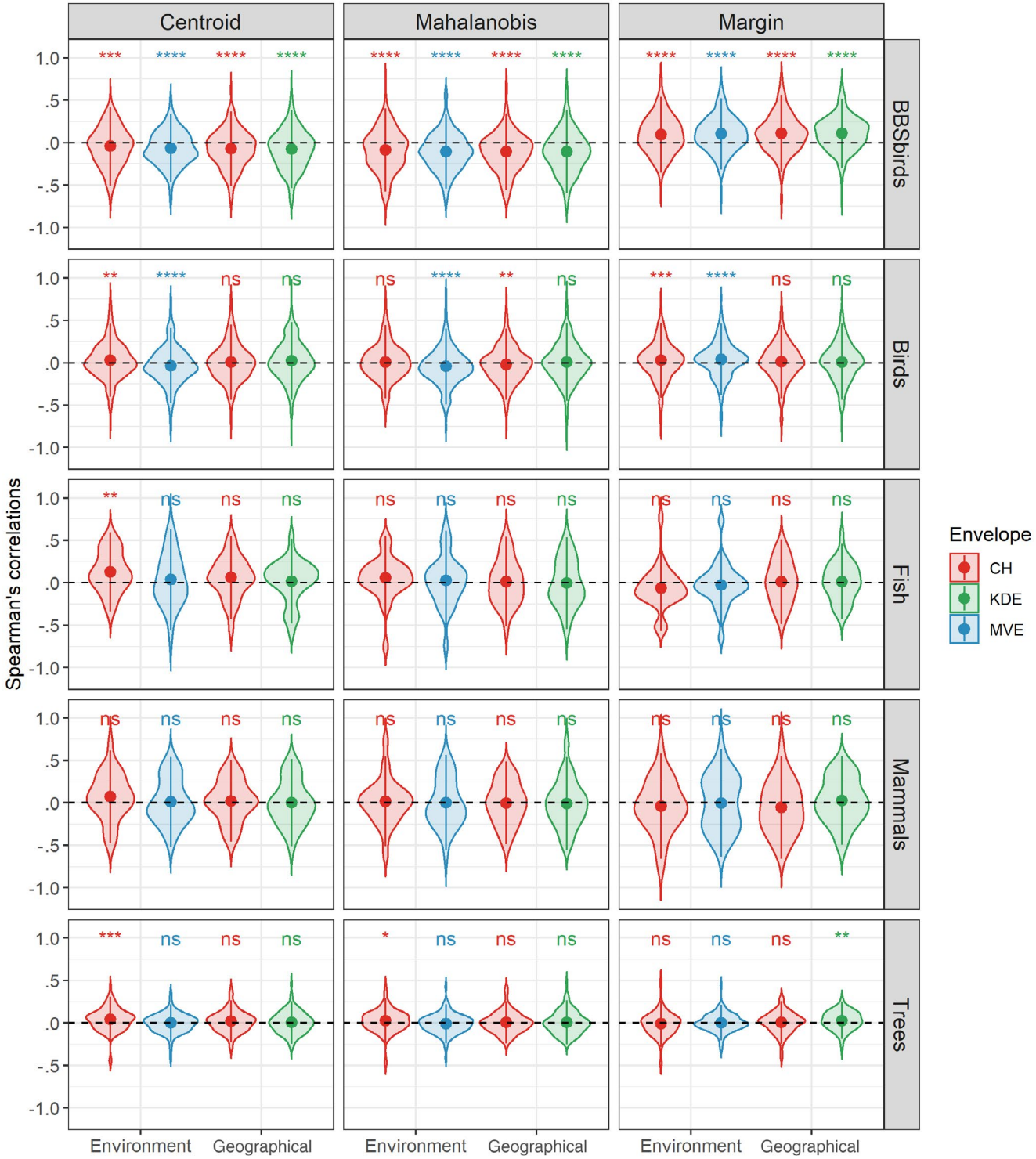

FIGURE 2 Violin plots showing the probability density of Spearman's correlation coefficients computed between distance and population density for mammals, fish, birds [both eBird and North American Breeding Bird Survey (BBS) data] and trees under the 12 different settings. The horizontal dashed lines point to the zero value. Note that the coefficients are expected to be positive when considering the distance to margins but negative for the two other distance metrics (Mahalanobis and Euclidean). For each setting and taxonomic group, we tested whether the average value differed from zero using Wilcoxon tests (ns: $p > .05$. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$. $^{****}p < .0001$). CH, convex hull; KDE, kernel density estimator; MVE, minimum volume ellipsoid

relationships strongly increased for both the BBS (59%) and the eBird (43.8%) datasets whereas the proportion of unexpected relationships remained similar (8% for the BSS dataset and 2.3% for the eBird dataset) (Supporting Information Figure S11b). For mammals, fish and trees, these proportions were 17 (5.8), 12 (12.8) and 12% (35%), respectively. A higher proportion of species presented expected

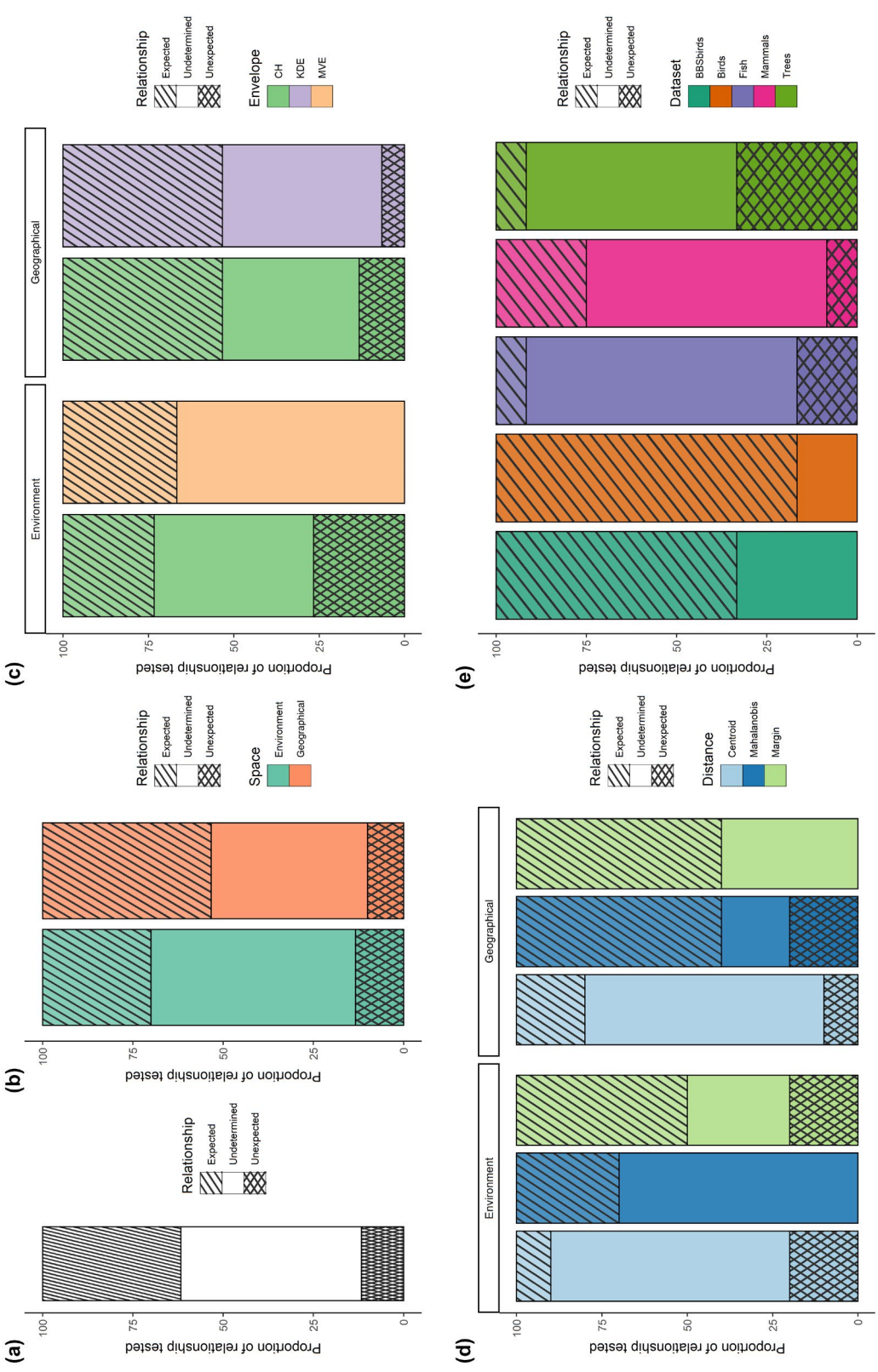

FIGURE 3 Barplots showing the proportion of expected (hatched areas), unexpected (crosshatched areas) and undetermined (blanked areas) relationships globally (a), in each space (b), for both envelopes in each space (c), for the three measures of distance in each space (d) and for each dataset (e). BBS, North American Breeding Bird Survey; CH, convex hull; KDE, kernel density estimator; MVE, minimum volume ellipsoid

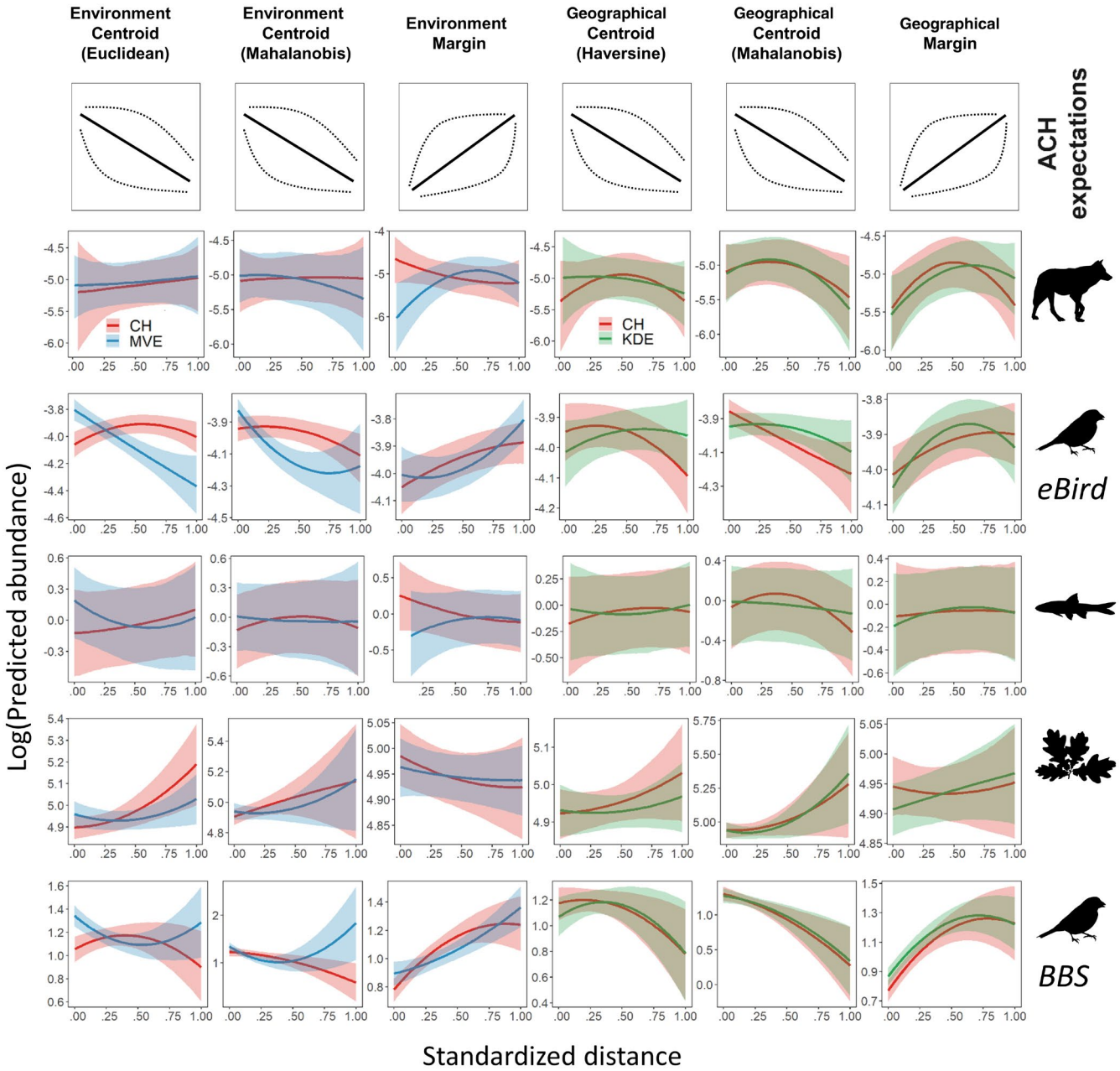

FIGURE 4 Predicted distance–abundance relationships estimated from Bayesian nonlinear mixed-effect models for mammals, fish, birds [both eBird and North American Breeding Bird Survey (BBS) data] and trees under the 12 different settings. Lines represent the median of the posterior predictive distribution of log-scaled abundances as a function of distance (standardized between 0 and 1) while the shaded area shows the corresponding 95% highest posterior density interval. Following the abundant centre hypothesis (ACH), the relationship is expected to be negative when distances are computed relative to the centre but positive when distances are computed relative to the margins. The top row gives an overview of the expected relationships depending on the distance metric considered with dotted lines representing nonlinear relationships (we here provide two examples of such relationships but different shapes can occur in real settings). For predicted distance–abundance relationships estimated with generalized additive models (GAMs), see Supporting Information Figure S1. CH, convex hull; KDE, kernel density estimator; MVE, minimum volume ellipsoid

relationships (and a lower proportion of unexpected relationships) under some settings than others in both the geographical (e.g. distance to the margins of the KDE) and the environmental (Euclidean distance to the centre of the MVE) space (Figure 5c), particularly after removing tree seedlings (Supporting Information Figure S10c) or migratory birds (Supporting Information Figure S11c).

4 | DISCUSSION

We investigated how sensitive the testing of the abundant centre hypothesis (Sagarin & Gaines, 2002; Sagarin et al., 2006), a central theory in biogeography (Brown, 1984), was to different contextual and methodological choices. For this, we investigated the strength

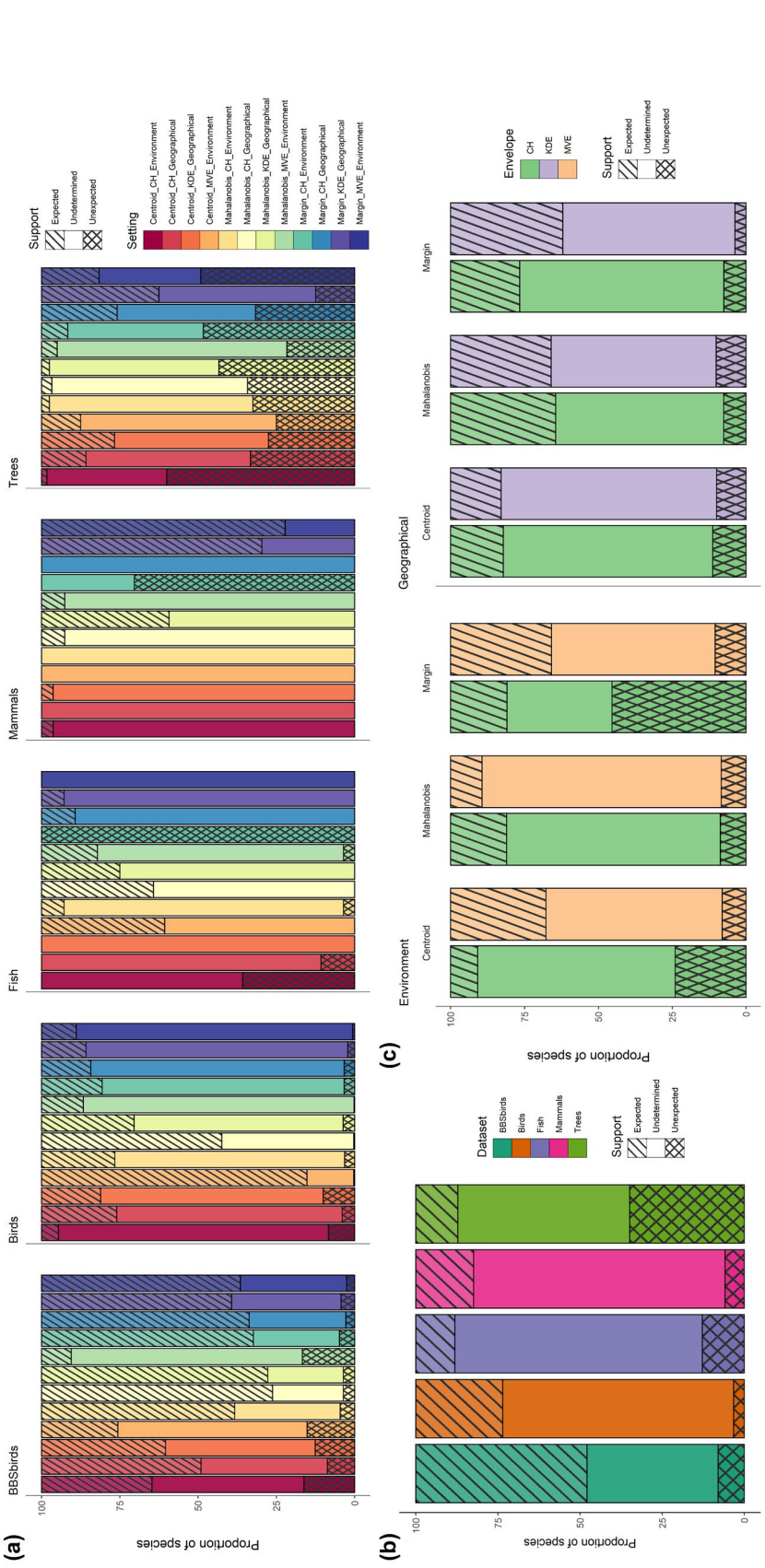

FIGURE 5 Barplots showing the proportion of species displaying expected (hatched areas), unexpected (crosshatched areas) and undetermined (blanked areas) relationships across the 12 settings for the five datasets (a), for each dataset (b) and for the three measures of distance computed for the two types of envelopes in each space (c). Species-specific distance–abundance relationships estimated under the 12 settings are presented in Supporting Information Figures S2–S6. BBS, North American Breeding Bird Survey; CH, convex hull; KDE, kernel density estimator; MVE, minimum volume ellipsoid

and support of distance–abundance relationships calculated under 12 different settings. Using published datasets (Dallas et al., 2017; Osorio-Olvera et al., 2020), we specifically explored whether the support for the ACH varied with respect to the context in which the distance–abundance relationship is evaluated (environmental or geographical), how species envelopes are delineated (CH or KDE in the geographical space, CH or MVE in the environmental space), how distances are measured (relative to the centre using Euclidean and Mahalanobis distances or to the margins), and the soundness of the data used (e.g. use of a standardized sampling protocol; measure of density/abundance at the population level instead of at a particular age class), using two different statistical frameworks (correlation coefficients and mixed-effect models).

4.1 | Mixed-effect models to unravel distance–abundance relationships

As an alternative to correlation coefficients used in previous studies, we introduced here the use of mixed-effect models to assess the distance–abundance relationship both globally and for each species at the same time. A key difference is that while simple correlations assume all species to be both equal and independent (statistically speaking), mixed-effect models use the commonalities (e.g. common response to climatic factors or anthropogenic pressures) shared between species to reduce the effect of sampling noise on parameter estimates (Gelman et al., 2013; Kruschke, 2014). Using correlation coefficients, we found very low support for the ACH, regardless of the settings and the taxonomic group considered, a result in line with recent studies (Dallas et al., 2017; Santini et al., 2019) but contradicting others (Martínez-Meyer et al., 2013; Osorio-Olvera et al., 2020; Yañez-Arenas et al., 2012). Only when considering the BBS dataset did we find a tendency toward support for the ACH, but the correlations were still weak. This low support can either be explained by all species showing no signal (scenario 1) or by large variations across species where positive trends are cancelled out by negative trends, resulting in a low signal after averaging coefficients (scenario 2). In the first scenario, we would not expect results to show any difference between correlation coefficients and mixed-effect models because all species are showing no signal, regardless of how much they are affected by sampling noise. Differences between methods are however expected in the second scenario owing to the shrinkage procedure where the estimate of species affected by sampling noise is pulled toward the main effects (Gelman et al., 2013). The fact that we found support for distance–abundance relationships of various shapes and strengths when using Bayesian mixed-effect models supports the second scenario, that is, that the low signal found when averaging correlation coefficients is partly due to the effect of sampling noise. Thus, our results suggest that mixed-effect models are a promising alternative to unravel common biogeographical patterns (either expected or not) and should be more routinely used in ACH studies (see also Chevalier & Knape, 2019; Schaub & Kéry, 2012). However, although using this statistical framework made it possible to unravel patterns that were undetected with correlations, we still found large variations in distance–abundance relationships across settings, taxonomic groups and species, highlighting the difficulty in finding a universal trend for the ACH.

4.2 | The ACH is sensitive to methodological settings: robust settings are needed

Overall, we found a tendency for some settings to provide stronger support for the ACH than others. Yet, no settings provided unambiguous support for the ACH, confirming the ongoing controversy (Dallas et al., 2017; Dallas, Pironon, et al., 2020; Knouft, 2018; Santini et al., 2019; Soberón et al., 2018). The ACH has traditionally been tested by measuring the distance of populations to the centre of the CH in the geographical space but empirical tests of this idea yielded mixed results (Sagarin & Gaines, 2002; Sagarin et al., 2006). More recently, the ACH shifted toward an environmentally based paradigm following the publication of results showing stronger support for the theory when considering the distance to the environmental niche centre instead of the geographical centre (Martínez-Meyer et al., 2013; Osorio-Olvera et al., 2020; Yañez-Arenas et al., 2012). However, other studies found no support for either of the two versions of the ACH – geographical or environmental – using the traditional CH/Euclidean setting (Dallas et al., 2017; Pironon et al., 2015; Santini et al., 2019). Following the study of Dallas et al. (2017), several authors called for alternative tests of the ACH using more theoretically/statistically supported settings (Knouft, 2018; Soberón et al., 2018). Osorio-Olvera et al. (2020) recently tested the ACH in the environmental space using the Mahalanobis distance to the centre of the MVE and found strong support for the theory. This study was however criticized on statistical grounds undermining the credibility of their results (Dallas, Pironon, et al., 2020). Overall, it appears that the distance–abundance relationship is highly sensitive to methodological and conceptual choices, which may explain why it is so difficult to find general evidence for the ACH in the literature. Our results clearly support this view and further suggest that the classical setting where distances are measured relative to the centre of the CH in either the geographical or the environmental space is not the most appropriate to describe the suitability gradient and that stronger support for the theory can be obtained by considering other settings such as the distance to the margins of the KDE and MVE.

4.2.1 | Stronger support for the ACH in the geographical space: the difficulty to characterize environmental niches

Overall, using these data, we surprisingly found stronger support for the ACH in the geographical space than in the environmental space. This result, which contrasts with previous studies supporting an environmentally based ACH (Martínez-Meyer et al., 2013;

Osorio-Olvera et al., 2020; Yañez-Arenas et al., 2012), could be due to several factors. In particular, most previous studies were delimiting the ecological (i.e. realized environmental) niche (Guisan et al., 2017; Peterson et al., 2011), which is constrained by the effect of biotic interactions, geographical barriers and the limited availability of environmental conditions. In contrast, we here attempted to focus on an estimate of the putative fundamental niche, whose properties (e.g. shape and volume) can be different from the ones of the ecological niche (Soberon & Arroyo-Peña, 2017). We made this choice because there is less theoretical underpinning for the ACH to hold in the ecological niche where the centre and the margins can be biased (i.e. they do not necessarily correspond to an optimum or a limit; Yañez-Arenas et al., 2020) owing to the constraints listed above (e.g. the margin can reflect a geographical barrier). Indeed, while ecological niches can match ACH expectations in some cases (e.g. species not affected by dispersal limitations or by the availability of the environment), it remains unclear whether abundances should decrease from the centre of the ecological niche toward the margins, notably because the link between abundance, fitness and habitat suitability is not clearly established and even controversial (Dallas & Hastings, 2018; McGill, 2012; Samis & Eckert, 2009) but also because marginal populations can be strongly affected by various factors including source–sink dynamics, climate change and evolutionary processes (Angert, 2009; Pironon et al., 2017). Further tests would nevertheless be welcome to determine under which circumstances the ACH could hold for ecological niches. Ideally, this would require disentangling the effects of biotic and abiotic factors to provide a comprehensive understanding of the factors involved in abundance-distribution patterns within ecological niches (see e.g. Pironon et al., 2018 for a similar approach). Another factor that could explain the low support for the ACH in the environmental space relates to the use of IUCN range maps to delimit species envelopes. Indeed, while considering the whole distribution of the species is important (Dallas, Pironon, et al., 2020; Knouft, 2018; Soberón et al., 2018), range maps are known to suffer from rough approximations (Rocchini et al., 2011) and to be derived based on distributions that may be strongly affected by human influences (Faurby & Svenning, 2015; Pacifici et al., 2020) with potential consequences for the location of populations relative to both the centre and the margins. Furthermore, although the use of a coarse resolution to extract climatic information from range maps (*c.* 340 km^2) is warranted to prevent misleading inferences (Hurlbert & Jetz, 2007), the drawback is that ecological niches are likely to be poorly resolved as well, regardless of the method used to delineate them. Finally, we note that – unlike for the geographical space – the climatic niche can contain much more dimensions (Hutchinson, 1957). Similar to most studies (e.g. Dallas et al., 2017; Santini et al., 2019), we here only considered two dimensions of the niche. Further studies are needed to test whether considering more dimensions (e.g. three; Osorio-Olvera et al., 2020) could help better characterize the distance of populations to the centre or the margins of the climatic niche.

4.2.2 | Reflecting the complex shape of species envelopes and using appropriate distance measures

Regardless of the space (geographical or environmental) in which the ACH is tested, most previous studies delineated species envelopes using the CH. This choice has strong implications on subsequent analyses owing to its influence regarding the location of populations relative to the centre and the margins of the envelope and therefore on the distance measured (Soberón et al., 2018). In general, we found lower support for the ACH when envelopes were delineated using the CH than with alternative methods (KDE and MVE). Ultimately, the choice of the method should be motivated by the objective of the study and particularly the space in which the ACH is tested. For the geographical space, we recommend using the KDE (Blonder et al., 2014) to account for the complex geometry of geographical ranges (Brown, 1984). For the environmental space, the choice of the envelope will depend on the type of niche one wants to consider. For instance, if the focus is on trying to estimate the fundamental niche, we recommend using the MVE because this algorithm is theoretically better suited to reproduce the shape of the fundamental niche than the CH (Soberón & Peterson, 2019). We note however that while the MVE is expected to provide a closer estimate than the CH (Qiao et al., 2016; Soberón, 2019; Soberón & Peterson, 2019; Yañez-Arenas et al., 2020), a formal test of the efficiency of MVE to estimate fundamental niches is still lacking. Alternatively, if the focus is on the ecological niche, we recommend using the KDE to account for the irregular and potentially fragmented shape of this type of niche (Soberón & Nakamura, 2009). Indeed, while ecological niches may not necessarily conform to the ACH (Yañez-Arenas et al., 2020), studying abundance-distribution patterns within this niche can shed lights on interesting biogeographical and ecological processes (Pironon et al., 2017).

The choice of the envelope ultimately has implications regarding how distances are measured. For instance, the complex geometry of the KDE implies that the Euclidean distance to the centre is not a good descriptor of the suitability gradient because populations can be close to the centre without necessarily being far from margins (Santini et al., 2019). Accordingly, we found stronger support for the theory when using alternative measures of distance (Mahalanobis distance to the centre, distance to margins).

4.2.3 | Accounting for nonlinear effects in distance–abundance relationships

Importantly, our results suggest that the distance–abundance relationship is nonlinear. Several factors have been advanced to explain this pattern including Allee effects, a nonlinear structure of the environment, threshold effects, demographic stochasticity, metapopulation dynamics, environmental heterogeneity, and biotic interactions (Dallas et al., 2017; Dallas & Santini, 2020; Holt, 2020; Osorio-Olvera et al., 2019; Pironon et al., 2017; Santini et al., 2019). Depending on the factor responsible, different shapes both concave

and convex can be expected. For instance, previous studies have shown that birds and mammals seldom adhere to an ideal-free distribution because social interactions such as dominance and territoriality play a major role in habitat use (Boyce et al., 2016). As a result, abundances in highly productive habitats (i.e. toward the centre) can be lower than in less productive habitats, which can materialize as a hump-shape relationship with a peak closer to the centre, and an abundance higher at the centre than at the margin. Such a pattern matches well with the expectation of an ideal despotic distribution of individuals (Pironon et al., 2017; Santini et al., 2019) and we found evidence for such a pattern for some taxonomic groups under some settings (see e.g. mammals and birds considering the distance to the margins of the KDE in the geographical space). Assessing the shape of the distance–abundance relationship can thus not only help us understand why a particular species or taxonomic group does not conform well to the ACH; it can also shed light on processes that could help explain abundance-distribution patterns across geographical ranges or climatic niches and ultimately propose hypotheses to explain departures from the theory.

4.3 | The ACH strongly varies across taxonomic groups and species: true pattern or inappropriate data?

Despite some commonalities, we found considerable variation among species and taxonomic groups. This result has classically been interpreted as evidence that the distance–abundance relationship does not hold in natural populations and thus that the ACH cannot be considered a biogeographical rule (Dallas et al., 2017; Dallas & Hastings, 2018; Santini et al., 2019). Several factors have been advanced to explain such an inconsistent pattern including biotic interactions, the presence of geographical barriers, metapopulation dynamics, a weak link between abundance and fitness, and many others (Dallas et al., 2017; Dallas & Santini, 2020; Dallas, Santini, et al., 2020; Holt, 2020; McGill, 2012; Osorio-Olvera et al., 2019; Sagarin et al., 2006).

However, an alternative explanation for this inconsistent pattern could be that the data used here and in previous studies (Dallas et al., 2017; Osorio-Olvera et al., 2020) are inappropriate to test the ACH (Dallas, Pironon, et al., 2020; Knouft, 2018; Santini et al., 2019; Soberón et al., 2018; Yañez-Arenas et al., 2020). Five features of the data seem worth further assessment. First, abundance data collected under a standardized sampling protocol should be favoured. We here tested the ACH using two different datasets for birds: eBird data, which does not follow any sampling protocol, and BBS data, which follows a standardized sampling protocol (Sauer et al., 2005). Using BBS data increased the overall support of the ACH under most settings owing to a larger proportion of variance explained and a larger number of species displaying the expected relationship. Second, considering larger sample sizes (i.e. species with a large number of populations) should help reduce the effect of sampling noise and yield more robust estimates of parameters. Indeed, although mixed-effect models made it possible to unravel undetected patterns for all taxonomic groups, relationships were still uncertain for fish and mammals, the groups with the lowest sample sizes, precluding us from drawing any firm conclusions for these groups. Third, the ACH should be tested using data that reflect abundance (and presumably fitness) at the population level and not at particular size or age classes. For instance, for trees, we found strong signals for unexpected relationships under a number of settings. While these results can be explained by the sampling protocol or by a weak relationship between fitness and abundance (McGill, 2012), they may also be explained by the fact that inferences were based on the abundance of seedlings, which can rather be regarded as a demographic parameter (e.g. recruitment success; Pironon et al., 2018). Indeed, species can present opposite geographical trends in individual vital rates, through different responses to environmental gradients; a process known as demographic compensation (Csergő et al., 2017; Kunstler et al., 2020; Pironon et al., 2018; Thuiller et al., 2014). Accordingly, recruitment (i.e. the number of seedlings) can be higher at the margins than at the centre of the geographical range where another demographic rate (e.g. adult survival) could be higher. For instance, because seedlings have low competitive abilities, it is likely that the most productive areas, where the abundance is expected to be higher, are dominated by mature individuals presenting high survival rates. If adult survival is decreasing toward range margins, the release of the intraspecific competition pressure may allow seedlings to thrive; a process that can ultimately promote species range expansion (Hampe & Petit, 2005). For plants, although considering abundance at the population level could be more appropriate, a more robust assessment of the ACH would likely be obtained using a measure of biomass density, which has proved to be a good proxy of individual fitness (e.g. the average diameter at breast height within a given area; Younginger et al., 2017). Fourth, abundance data should be collected over the whole species range (Dallas, Pironon, et al., 2020; Knouft, 2018; Soberón et al., 2018). This was not the case for any of the datasets considered in this study. We attempted to cope with this problem by estimating species envelopes based on IUCN range maps instead of the pool of populations for which abundance data were available but these range maps are not ideal (Herkt et al., 2017). Fifth, a robust estimation of species envelopes in both spaces is paramount since all inferences are based on this assessment. The fact that we found lower support for the ACH when including migratory birds supports this view. Indeed, because the geographical range of migratory species is divided into breeding and non-breeding areas, the definition of a unique geographical or environmental space is challenging (Santini et al., 2019). This problem is even more acute regarding the definition of the niche because a different treatment should be applied depending on the breeding (environmental conditions during the breeding period) and the non-breeding range (environmental conditions during the whole year). This is far from trivial given that species can have different breeding periods. Unless one can account for these specificities, these species should be considered separately (Dallas et al., 2017; Osorio-Olvera et al., 2020).

4.4 | A multi-setting inference framework to test the ACH?

We considered various settings to provide a comparative overview of their impact on the estimated distance–abundance relationship. For most taxonomic groups, mixed-effect models displayed a similar explanatory power, despite providing varying support for the ACH, thus suggesting that standard evaluation metrics (R^2 and Bayesian *p*-values in our case) can hardly discriminate between robust and bad settings. This similar performance arose because all models performed well in capturing among-species variations in the distance–abundance relationship (i.e. most of the variance was captured by the random effects), a pattern that was expected given the variations observed in previous studies (e.g. Dallas et al., 2017). This does not mean that differences between models cannot emerge with other datasets (e.g. the tree dataset) or with other measures of model performance [e.g. Akaike's information criterion (AIC), Bayesian information criterion (BIC)]. In theory, this would make it possible to run various models with distances computed under different settings and to select the model that best fits the data (Bolker et al., 2009). Using this approach, only settings supported by the theory and based on robust estimators (e.g. KDE instead of CH) should be considered. Otherwise, there is a risk of false inference because (a) multiple-testing could increase the probability of finding spurious relationships (type-I error rate) and (b) bad settings (or bad predictors; see e.g. Fourcade et al., 2018) can sometimes fit the data just as well or even better than more robust settings, for example because the data or the performance measures are inappropriate. In line with studies that advise choosing predictors based on the known ecology or physiology of species (instead of using a model selection procedure based on a large set of predictors; Burnham & Anderson, 2002), an alternative approach would be to choose the settings beforehand based on the theory, the availability of robust statistical tools and the purpose of the study (e.g. using the KDE and a measure of distance to margins if the focus is to test the ACH in the geographical space; Dallas, Santini, et al., 2020).

5 | CONCLUSIONS

Many ecological and evolutionary assumptions have been derived based on the abundant centre hypothesis (ACH) but several studies conducted across a range of taxa and locations failed to find general evidence for abundant centre patterns (Dallas et al., 2017; Sagarin & Gaines, 2002; Sagarin et al., 2006; Santini et al., 2019). Importantly, our results suggest that the noise inherent to ecological data can strongly hamper our ability to detect theory-based biogeographical patterns. Mixed-effect models can help in this regard and we advise future ACH studies to more regularly use this framework. Here, it yielded mixed and inconclusive results, even though we found a tendency for more theoretically supported settings (e.g. the distance to the margins of an MVE) to provide stronger support for the ACH. We nonetheless propose that a way to help resolve the debate would be to test the ACH using (a) robust abundance datasets to limit the effect of sampling noise on parameter estimates and increase the statistical power of models and (b) settings that have stronger theoretical support.

ACKNOWLEDGMENTS

We sincerely thank Dallas and colleagues for having made their data and R code publicly available, allowing us to repeat their analyses and to go a step further in testing the ACH. AG obtained support from the Swiss National Science Foundation, grant number CR23I2_162754.

AUTHOR CONTRIBUTIONS

MC conceived the study, performed the analyses with help from OB and wrote the first draft of the manuscript. OB and AG provided critical comments on different versions of the manuscript.

DATA AVAILABILITY STATEMENT

All the data used in this study have already been published [see Dallas et al. (2017) and Osorio-Olvera et al. (2020) for abundance data] or are freely available to download on the web (climatic data can be downloaded at https://www.worldclim.org/data/worldclim21.html; range maps for trees can be downloaded at https://www.fs.fed.us/nrs/atlas/littlefia/ and range maps for mammals, birds and fish can be downloaded at https://www.iucnredlist.org/resources/spatial-data-download). Nonetheless, to ensure reproducibility, all these data have been compiled into one unique folder available to download on Dryad (https://doi.org/10.5061/dryad.bzkh18993). The R code, the JAGS model and the functions used to draw inferences from these data are available to download on GitHub (https://github.com/Mathieu-Chevalier/ACH-GEB; DOI: https://doi.org/10.5281/zenodo.5113900).

ORCID

Mathieu Chevalier https://orcid.org/0000-0002-1170-5343
Antoine Guisan https://orcid.org/0000-0002-3998-4815

REFERENCES

Angert, A. L. (2009). The niche, limits to species' distributions, and spatiotemporal variation in demography across the elevation ranges of two monkeyflowers. *Proceedings of the National Academy of Sciences USA*, *106*, 19693–19698. https://doi.org/10.1073/pnas.0901652106

Austin, M. P. (1999). A silent clash of paradigms: Some inconsistencies in community ecology. *Oikos*, *86*, 170–178.

Barve, N., Barve, V., Jiménez-Valverde, A., Lira-Noriega, A., Maher, S. P., Peterson, A. T., Soberón, J., & Villalobos, F. (2011). The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling*, *222*, 1810–1819. https://doi.org/10.1016/j.ecolmodel.2011.02.011

Blonder, B., Lamanna, C., Violle, C., & Enquist, B. J. (2014). The n-dimensional hypervolume. *Global Ecology and Biogeography*, *23*, 595–609.

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., & White, J. S. S. (2009). Generalized linear mixed models: A practical guide for ecology and evolution. *Trends in Ecology and Evolution*, *24*, 127–135. https://doi.org/10.1016/j.tree.2008.10.008

Boyce, M. S., Johnson, C. J., Merrill, E. H., Nielsen, S. E., Solberg, E. J., & van Moorter, B. (2016). Can habitat selection predict abundance? *Journal of Animal Ecology*, *85*, 11–20.

Broennimann, O., Fitzpatrick, M. C., Pearman, P. B., Petitpierre, B., Pellissier, L., Yoccoz, N. G., Thuiller, W., Fortin, M. J., Randin, C., Zimmermann, N. E., Graham, C. H., & Guisan, A. (2012). Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography*, *21*, 481–497. https://doi.org/10.1111/j.1466-8238.2011.00698.x

Broennimann, O., Petitpierre, B., Chevalier, M., González-Suárez, M., Jeschke, J. M., Rolland, J., Gray, S. M., Bacher, S., & Guisan, A. (2021). Distance to native climatic niche margins explains establishment success of alien mammals. *Nature Communications*, *12*, 2353.

Broennimann, O., Treier, U. A., Müller-Schärer, H., Thuiller, W., Peterson, A. T., & Guisan, A. (2007). Evidence of climatic niche shift during biological invasion. *Ecology Letters*, *10*, 701–709. https://doi.org/10.1111/j.1461-0248.2007.01060.x

Brown, J. H. (1984). On the relationship between abundance and distribution of species. *The American Naturalist*, *124*, 255–279. https://doi.org/10.1086/284267

Brown, J. H., Stevens, G. C., & Kaufman, D. M. (1996). THE GEOGRAPHIC RANGE: Size, shape, boundaries, and internal structure. *Annual Review of Ecology and Systematics*, *27*, 597–623. https://doi.org/10.1146/annurev.ecolsys.27.1.597

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multi-model inference: A practical information-theoretic approach*. Springer.

Chevalier, M., Broennimann, O., Cornuault, J., & Guisan, A. (2021). Data integration methods to account for spatial niche truncation effects in regional projections of species distribution. *Ecological Applications*, In press. https://doi.org/10.1002/eap.2427

Chevalier, M., & Knape, J. (2019). The cost of complexity in forecasts of population abundances is reduced but not eliminated by borrowing information across space using a hierarchical approach. *Oikos*, *129*(2), 249–260. https://doi.org/10.1111/oik.06401

Colwell, R. K., & Rangel, T. F. (2009). Hutchinson's duality: The once and future niche. *Proceedings of the National Academy of Sciences USA*, *106*, 19651–19658. https://doi.org/10.1073/pnas.0901650106

Csergő, A. M., Salguero-Gómez, R., Broennimann, O., Coutts, S. R., Guisan, A., Angert, A. L., Welk, E., Stott, I., Enquist, B. J., McGill, B., Svenning, J. C., Violle, C., & Buckley, Y. M. (2017). Less favourable climates constrain demographic strategies in plants. *Ecology Letters*, *20*, 969–980. https://doi.org/10.1111/ele.12794

Dallas, T., Decker, R. R., & Hastings, A. (2017). Species are not most abundant in the centre of their geographic range or climatic niche. *Ecology Letters*, *20*, 1526–1533. https://doi.org/10.1111/ele.12860

Dallas, T. A., & Hastings, A. (2018). Habitat suitability estimated by niche models is largely unrelated to species abundance. *Global Ecology and Biogeography*, *27*, 1448–1456. https://doi.org/10.1111/geb.12820

Dallas, T., Pironon, S. & Santini, L. (2020). The abundant-centre is not all that abundant: A comment to Osorio-Olvera et al. 2020. *bioRxiv [Preprint]*. https://doi.org/10.1101/2020.02.27.968586

Dallas, T. A., & Santini, L. (2020). The influence of stochasticity, landscape structure and species traits on abundant–centre relationships. *Ecography*, *43*, 1341–1351. https://doi.org/10.1111/ecog.05164

Dallas, T. A., Santini, L., Decker, R., & Hastings, A. (2020). Weighing the evidence for the abundant-center hypothesis. *Biodiversity Informatics*, *15*, 81–91. https://doi.org/10.17161/bi.v15i3.11989

Escobar, L. E., Qiao, H., Cabello, J., & Peterson, A. T. (2018). Ecological niche modeling re-examined: A case study with the Darwin's fox. *Ecology and Evolution*, *8*, 4757–4770. https://doi.org/10.1002/ece3.4014

Faurby, S., & Svenning, J. C. (2015). Historic and prehistoric human-driven extinctions have reshaped global mammal diversity patterns. *Diversity and Distributions*, *21*, 1155–1166. https://doi.org/10.1111/ddi.12369

Fourcade, Y., Besnard, A. G., & Secondi, J. (2018). Paintings predict the distribution of species, or the challenge of selecting environmental predictors and evaluation statistics. *Global Ecology and Biogeography*, *27*, 245–256. https://doi.org/10.1111/geb.12684

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis*. Chapman and Hall/CRC.

Gelman, A., Meng, X.-L., & Stern, H. (1996). Posterior predictive assessment of model fitness via realized discrepancies. *Statistica Sinica*, *6*, 733–807.

Guisan, A., Thuiller, W., & Zimmermann, N. E. (2017). *Habitat suitability and distribution models: With applications in R* (1st ed.). Cambridge University Press.

Hampe, A., & Petit, R. J. (2005). Conserving biodiversity under climate change: The rear edge matters. *Ecology Letters*, *8*, 461–467. https://doi.org/10.1111/j.1461-0248.2005.00739.x

Harrison, X. A., Donaldson, L., Correa-Cano, M. E., Evans, J., Fisher, D. N., Goodwin, C. E. D., Robinson, B. S., Hodgson, D. J., & Inger, R. (2018). A brief introduction to mixed effects modelling and multi-model inference in ecology. *PeerJ*, *2018*, 1–32. https://doi.org/10.7717/peerj.4794

Hastie, T., & Tibshirani, R. (1986). Generalized additive models. *Statistical Science*, *10*, 354–363.

Herkt, K. M. B., Skidmore, A. K., & Fahr, J. (2017). Macroecological conclusions based on IUCN expert maps: A call for caution. *Global Ecology and Biogeography*, *26*, 930–941. https://doi.org/10.1111/geb.12601

Holt, R. D. (2020). Reflections on niches and numbers. *Ecography*, *43*, 387–390. https://doi.org/10.1111/ecog.04828

Hurlbert, A. H., & Jetz, W. (2007). Species richness, hotspots, and the scale dependence of range maps in ecology and conservation. *Proceedings of the National Academy of Sciences USA*, *104*, 13384–13389. https://doi.org/10.1073/pnas.0704469104

Hutchinson, G. (1957). Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, *22*, 415–427. https://doi.org/10.1101/SQB.1957.022.01.039

Jackson, S. T., & Overpeck, J. T. (2000). Responses of plant populations and communities to environmental changes of the late Quaternary. *Paleobiology*, *26*, 194–220. https://doi.org/10.1017/S0094837300026932

Knouft, J. H. (2018). Appropriate application of information from biodiversity databases is critical when investigating species distributions and diversity: A comment on Dallas et al.. *Ecology Letters*, *21*, 1119–1120.

Knouft, J. H., & Anthony, M. M. (2016). Climate and local abundance in freshwater fishes. *Royal Society Open Science*, *3*, 160093. https://doi.org/10.1098/rsos.160093

Kruschke, J. (2014). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan*. Academic Press.

Kunstler, G., Guyennon, A., Ratcliffe, S., Rüger, N., Ruiz-Benito, P., Childs, D. Z., Dahlgren, J., Lehtonen, A., Thuiller, W., Wirth, C., Zavala, M. A., & Salguero-Gomez, R. (2020). Demographic performance of European tree species at their hot and cold climatic edges. *Journal of Ecology*, 1–14.

Martínez-Meyer, E., Díaz-Porras, D., Peterson, A. T., & Yáñez-Arenas, C. (2013). Ecological niche structure and rangewide abundance patterns of species. *Biology Letters*, *9*, 20120637. https://doi.org/10.1098/rsbl.2012.0637

McGill, B. J. (2012). Trees are rarely most abundant where they grow best. *Journal of Plant Ecology*, *5*, 46–51. https://doi.org/10.1093/jpe/rtr036

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*, 133–142.

Osorio-Olvera, L., Soberón, J., & Falconi, M. (2019). On population abundance and niche structure. *Ecography*, *42*, 1415–1425. https://doi.org/10.1111/ecog.04442

Osorio-Olvera, L., Yáñez-Arenas, C., Martínez-Meyer, E., & Peterson, T. A. (2020). Relationships between population densities and

niche-centroid distances in North American birds. *Ecology Letters*, *23*(3), 555–564. https://doi.org/10.1111/ele.13453

Pacifici, M., Rondinini, C., Rhodes, J. R., Burbidge, A. A., Cristiano, A., Watson, J. E. M., Woinarski, J. C. Z., & Di Marco, M. (2020). Global correlates of range contractions and expansions in terrestrial mammals. *Nature Communications*, *11*, 1–9. https://doi.org/10.1038/s41467-020-16684-w

Peterson, A. T., Soberón, J., Pearson, R. G., Anderson, R. P., Martínez-Meyer, E., Nakamura, M., & Araújo, M. B. (2011). *Ecological niches and geographic distributions*. Princeton University Press.

Pironon, S., Papuga, G., Villellas, J., Angert, A. L., García, M. B., & Thompson, J. D. (2017). Geographic variation in genetic and demographic performance: New insights from an old biogeographical paradigm. *Biological Reviews*, *92*, 1877–1909. https://doi.org/10.1111/brv.12313

Pironon, S., Villellas, J., Morris, W. F., Doak, D. F., & García, M. B. (2015). Do geographic, climatic or historical ranges differentiate the performance of central versus peripheral populations? *Global Ecology and Biogeography*, *24*, 611–620. https://doi.org/10.1111/geb.12263

Pironon, S., Villellas, J., Thuiller, W., Eckhart, V. M., Geber, M. A., Moeller, D. A., & García, M. B. (2018). The 'Hutchinsonian niche' as an assemblage of demographic niches: Implications for species geographic ranges. *Ecography*, *41*, 1103–1113. https://doi.org/10.1111/ecog.03414

Prasad, A. M., & Iverson, L. R. (2003). *Little's range and FIA importance value database for 135 eastern US tree species*. Northeastern Research Station, USDA Forest Service.

Qiao, H., Peterson, A. T., Campbell, L. P., Soberón, J., Ji, L., & Escobar, L. E. (2016). NicheA: Creating virtual species and ecological niches in multivariate environmental scenarios. *Ecography*, *39*, 805–813. https://doi.org/10.1111/ecog.01961

Rocchini, D., Hortal, J., Lengyel, S., Lobo, J. M., Jiménez-Valverde, A., Ricotta, C., Bacaro, G., & Chiarucci, A. (2011). Accounting for uncertainty when mapping species distributions: The need for maps of ignorance. *Progress in Physical Geography*, *35*, 211–226. https://doi.org/10.1177/0309133311399491

Sagarin, R. D., & Gaines, S. D. (2002). The "abundant centre" distribution: To what extent is it a biogeographical rule? *Ecology Letters*, *5*, 137–147. https://doi.org/10.1046/j.1461-0248.2002.00297.x

Sagarin, R. D., Gaines, S. D., & Gaylord, B. (2006). Moving beyond assumptions to understand abundance distributions across the ranges of species. *Trends in Ecology and Evolution*, *21*, 524–530. https://doi.org/10.1016/j.tree.2006.06.008

Samis, K. E., & Eckert, C. G. (2009). Ecological correlates of fitness across the northern geographic range limit of a Pacific coast dune plant. *Ecology*, *90*, 3051–3061. https://doi.org/10.1890/08-1914.1

Santini, L., Pironon, S., Maiorano, L., & Thuiller, W. (2019). Addressing common pitfalls does not provide more support to geographical and ecological abundant-centre hypotheses. *Ecography*, *42*, 696–705. https://doi.org/10.1111/ecog.04027

Sauer, J. R., Link, W. A., & Royle, J. A. (2005). Hierarchical models and Bayesian analysis of bird survey information. In C. Ralph & T. Rich (Eds.), *Bird conservation implementation and integration in the Americas: Proceedings of the Third International Partners in Flight Conference* (pp. 762–770). United States Department of Agriculture.

Schaub, M., & Kéry, M. (2012). Combining information in hierarchical models improves inferences in population ecology and demographic population analyses. *Animal Conservation*, *15*, 125–126. https://doi.org/10.1111/j.1469-1795.2012.00531.x

Soberón, J. (2019). A Grinnellian niche perspective on species-area relationships. *The American Naturalist*, *194*, 760–775. https://doi.org/10.1086/705898

Soberon, J., & Arroyo-Peña, B. (2017). Are fundamental niches larger than the realized? Testing a 50-year-old prediction by Hutchinson. *PLoS ONE*, *12*, 1–14. https://doi.org/10.1371/journal.pone.0175138

Soberón, J., & Nakamura, M. (2009). Niches and distributional areas: Concepts, methods, and assumptions. *Proceedings of the National Academy of Sciences USA*, *106*, 19644–19650. https://doi.org/10.1073/pnas.0901637106

Soberón, J., & Peterson, A. T. (2019). What is the shape of the fundamental Grinnellian niche? *Theoretical Ecology*, *13*(1), 105–115. https://doi.org/10.1007/s12080-019-0432-5

Soberón, J. M., Peterson, T. A., Osorio-Olvera, L., Townsend Peterson, A., Osorio-Olvera, L., & Osorio-Olvera, L. (2018). A comment on "Species are not most abundant in the centre of their geographic range or climatic niche". *Rethinking Ecology*, *3*, 13–18. https://doi.org/10.3897/rethinkingecology.3.24827

Sullivan, B. L., Wood, C. L., Iliff, M. J., Bonney, R. E., Fink, D., & Kelling, S. (2009). eBird: A citizen-based bird observation network in the biological sciences. *Biological Conservation*, *142*, 2282–2292. https://doi.org/10.1016/j.biocon.2009.05.006

Thibault, K. M., Supp, S. R., Giffin, M., White, E. P., & Morgan Ernest, S. K. (2011). Species composition and abundance of mammalian communities. *Ecology*, *92*, 2316. https://doi.org/10.1890/11-0262.1

Thuiller, W., Münkemüller, T., Schiffers, K. H., Georges, D., Dullinger, S., Eckhart, V. M., Edwards, T. C., Gravel, D., Kunstler, G., Merow, C., Moore, K., Piedallu, C., Vissault, S., Zimmermann, N. E., Zurell, D., & Schurr, F. M. (2014). Does probability of occurrence relate to population dynamics? *Ecography*, *37*, 1155–1166. https://doi.org/10.1111/ecog.00836

Van-Aelst, S., & Rousseeuw, P. (2009). Minimum volume ellipsoid. *Wiley Interdisciplinary Reviews: Computational Statistics*, *1*, 71–82. https://doi.org/10.1002/wics.19

Wand, M. P., & Jones, M. C. (1994). *Kernel smoothing*. Chapman and Hall/CRC.

Woudenberg, S. W., Conkling, B. L., Connell, B. M. O., Lapoint, E. B., Turner, J. A., Waddell, K. L. & Service, F. (2010). *The Forest Inventory and Analysis Database: Database description and users manual version 4.0 for Phase 2*. United States Department of Agriculture.

Yancovitch Shalom, H., Granot, I., Blowes, S. A., Friedlander, A., Mellin, C., Ferreira, C. E. L., Arias-González, J. E., Kulbicki, M., Floeter, S. R., Chabanet, P., Parravicini, V., & Belmaker, J. (2020). A closer examination of the 'abundant centre' hypothesis for reef fishes. *Journal of Biogeography*, *47*, 2194–2209. https://doi.org/10.1111/jbi.13920

Yañez-Arenas, C., Martín, G., Osorio-Olvera, L., Escobar-Luján, J., Castaño-Quintero, S., Chiappa-Carrara, X., & Martínez-Meyer, E. (2020). The abundant niche-centroid hypothesis: Key points about unfilled niches and the potential use of supraspecfic modeling units. *Biodiversity Informatics*, *15*, 92–102. https://doi.org/10.17161/bi.v15i2.13218

Yañez-Arenas, C., Martínez-Meyer, E., Mandujano, S., & Rojas-Soto, O. (2012). Modelling geographic patterns of population density of the white-tailed deer in central Mexico by implementing ecological niche theory. *Oikos*, *121*, 2081–2089. https://doi.org/10.1111/j.1600-0706.2012.20350.x

Yañez-Arenas, C., Rioja-Nieto, R., Martín, G. A., Dzul-Manzanilla, F., Chiappa-Carrara, X., Buenfil-Ávila, A., Manrique-Saide, P., Correa-Morales, F., Díaz-Quiñónez, J. A., Pérez-Rentería, C., Ordoñez-Álvarez, J., Vazquez-Prokopec, G., & Huerta, H. (2018). Characterizing environmental suitability of Aedes albopictus (Diptera: Culicidae) in Mexico based on regional and global niche models. *Journal of Medical Entomology*, *55*, 69–77. https://doi.org/10.1093/jme/tjx185

Younginger, B. S., Sirová, D., Cruzan, M. B., & Ballhorn, D. J. (2017). Is biomass a reliable estimate of plant fitness? *Applications in Plant Sciences*, *5*, 1600094. https://doi.org/10.3732/apps.1600094

BIOSKETCHES

Mathieu Chevalier, Olivier Broennimann and Antoine Guisan are ecologists interested in spatio-temporal patterns and processes acting at large spatial scales.

Mathieu Chevalier's research is grounded in population and community dynamics. His current research focus on the processes acting on populations and communities within species geographical ranges and ecological niches.

Olivier Broennimann's research focus on niche theory and invasive species. **Antoine Guisan** has his main research interest in niche theory and spatial modelling of species distributions.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the Supporting Information section.

How to cite this article: Chevalier, M., Broennimann O., & Guisan A. (2021). Using a robust multi-settings inference framework on published datasets still reveals limited support for the abundant centre hypothesis: More testing needed on other datasets. *Global Ecology and Biogeography*, 00, 1–18. https://doi.org/10.1111/geb.13376