MONOGRAPH

Frontiers of Biogeography
the scientific journal of
the International Biogeography Society

A tale of two niches: methods, concepts, and evolution

Jason L. Brown[1] and Ana C. Carnaval[2]

1. Cooperative Wildlife Research Laboratory & The Center for Ecology, Southern Illinois University, Carbondale, IL 62901, USA. 2. City College of New York and The Graduate Center, City University of New York, New York, NY 10031, USA. *Corresponding Author: Jason L. Brown (jason.brown@siu.edu)

Abstract. Being snapshots in time, the ranges of species may fall short of representing all of the geographic or environmental-space that these taxa are able to occupy. This has important implications for niche studies, yet most comparative studies overlook the transient nature of species' distributions and assume that they are at equilibrium. We review the methods most widely used for niche comparisons today and suggest a modified framework to describe and compare niches based on snapshot range data of species. First, we introduce a new environmental-space-based Niche Equivalencestatistic to test niche similarity between two species, which explicitly incorporates the spatial distribution of environments and their availability into statistical tests. We also introduce a new Background statistic to measure the ability of this Niche Equivalence statistic to detect differences based on the available environmental-space. These metrics enable fair comparisons between different geographies when the ranges of species are out of equilibrium. Based on distinct parameterizations of the new Equivalence and Background statistics, we then propose a Niche Divergence test and a Niche Overlap test, which allow assessment of whether differences between species emerge from true niche divergences. These methods are implemented in a new R package, 'humboldt' and applied to simulated species with pre-defined niches. The new methods improve accuracy of niche similarity and associated tests – consistently outperforming other tests. We show that the quantification of niche similarity should be performed only in environmental-space, which is less sensitive than geographic space to the spatial abundance of key environmental variables. Further, our methods characterize the relationships between non-analogous and analogous climates in the species' distributions, something not available previously. These improvements allow assessment of whether the different environmental-spaces occupied by two taxa emerge from true niche evolution, as opposed to differences in life history and biological interactors, or differences in the variety and configuration of environments accessible to them.

Highlights:

- The distributions of most species are either shrinking or expanding in responses to changes in the environment. However, most methods that compare the climatic niches that they occupy assume that the species have reached an equilibrium – which can lead to spurious conclusions
- We present a new method that quantifies how much two species differ in their climatic niches, while not assuming that they are at equilibrium. For that, it incorporates information about the environments that each species can access
- The method allows scientists to more accurately evaluate if two species have actually evolved different niches or if they occupy separate climates because of differences in life history, biological interactions, or in the environments available
- We show that this novel method is more accurate than other available tests when applied to simulated data

Keywords: Ecospat, ENMtools, fundamental niche, niche evolution, NicheA, niche divergence, niche similarity, niche truncation, potential niche

Introduction

Understanding the drivers of species' ranges remains a fundamental aim across ecology and evolution (Lomolino et al. 2017). A key goal is to characterize and compare the ecological niches of species, with the ultimate aim of assessing how niches evolve. Most of these studies follow Soberón and Nakamura's (2009) definition of the Grinnellian niche (Grinnell 1914, 1917), which is a subset of environmental conditions in which populations of a species have positive growth

rates (James et al. 1984, Soberón 2007). Thanks to advances in the methods that quantify and compare species' distributions, studies of Grinnellian niches – including how niches differ between species and how niches evolve over time – have flourished in recent years (Peterson et al. 1999; Wiens and Graham 2005, Losos 2008, Pearman et al. 2008). They are included in disciplines as diverse as conservation biology, historical biogeography, evolution, and community ecology (Dennis and Stefan 2009, McCormack et al. 2010, Pellissier et al. 2013, Guisan et al. 2014), all of them rely on interpreting patterns and drivers of species' distributions across landscapes. Despite widespread interest and application, the field remains young and has yet to coalesce on lexicon, methods, and theory (Elton 1927, Jackson and Overpeck 2000, Soberón and Nakamura 2009, Guisan et al. 2014, Qiao et al. 2017).

One limitation faced by many comparative niche studies stems from the fact that niche similarity is often quantified in geographic space (Warren et al. 2008, 2010) as opposed to environmental-space (Broennimann et al. 2012, Di Cola et al. 2017). Studies that are focused on geographic space (G-space) compare niches by building correlative models of species'distributions from environmental descriptors and locality data, and by subsequently comparing their ranges when the inferred environmental envelope is projected in geographic space (Fig. 1; Soberón 2007, Colwell and Rangel 2009). The more similar the geographic distributions of the species being compared, the higher the inferred niche similarity. Though superficially this is true, this approach is handicapped by the fact that such measurements are only accurate when habitats that span the ecological tolerances of both species are equally represented in geographic space. This assumption makes measuring niche similarity in geographic space problematic – particularly, though certainly not exclusively, in the case of invasive species. This is because it requires two species to occupy the same geographical area before one can assess niche similarity - despite the fact that both analogous environments and the species

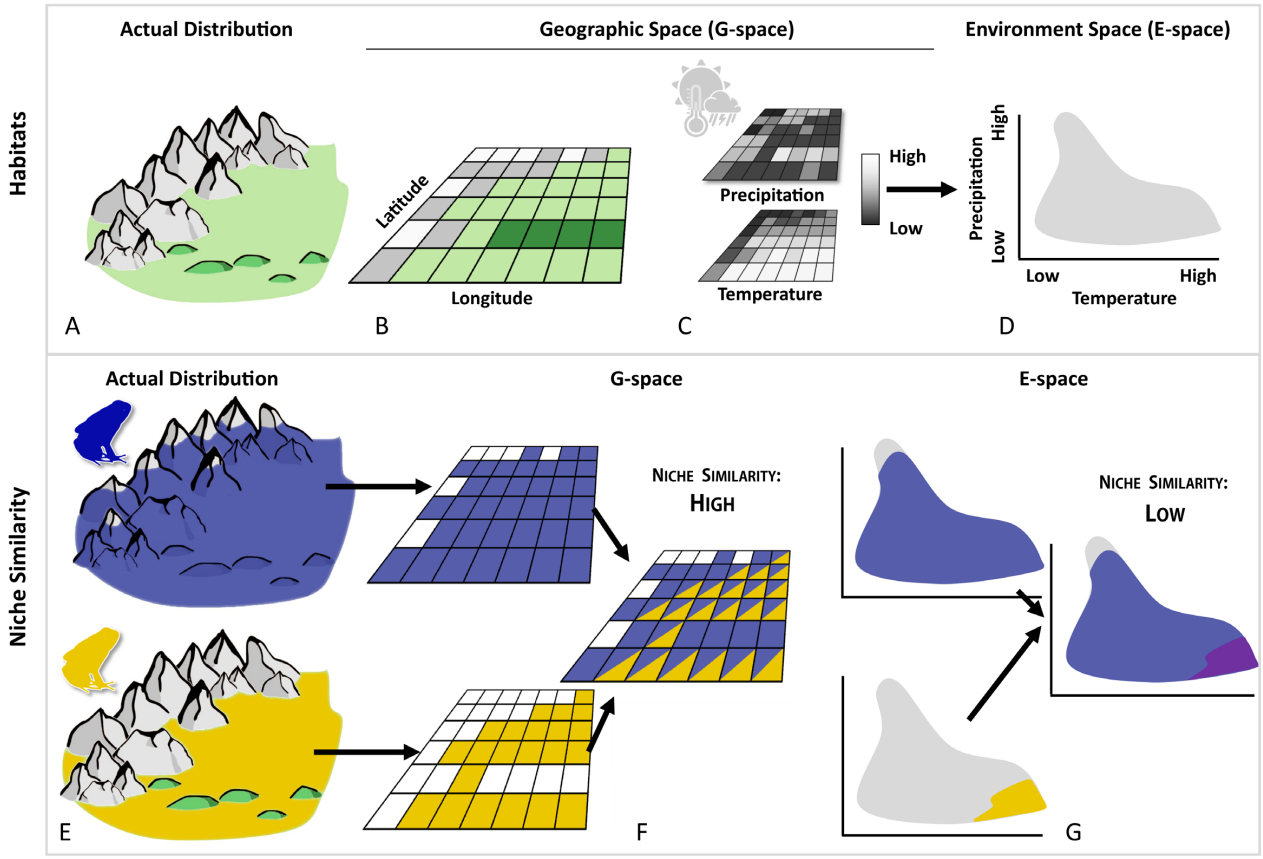


Fig. 1. Relationships between spatial, environmental and niche similarity measurements. Habitats (A) can be characterized several ways. One of the most common methods is to import measurements of raw environmental data into a GIS (B) and plot it in geographic space (C). Habitats can also be characterized by their environmental-space (E-space) (D) represented within the geographic region. In the example, we plotted annual precipitation against the annual temperature of the landscape. **Niche Similarity**. Two species' distributions (E) can be quantified in G-space (F) or in E-space (G) and then similarity can be measured. Here we present a simple habitat with mountains, hills, and lowland plains. In this example, lowland habitats are common and the observed niche similarity between the lowland-only frog and the lowland and highland frog is relatively high if based only on G-space. Measurements in E-space result in relatively low niche similarity.

might exist elsewhere. Lastly, because comparisons in G-space count the number of shared cells occupied (or in continuous analysis, the relative suitability), the resulting niche similarity measurement is biased towards more common habitat types.

To illustrate this point, picture a scenario where three landscape types reflect different environmental conditions (e.g., mean temperatures) and are present in a given area: [i] a montane environment in which topography changes sharply, such as a tall, steep mountain; [ii] a montane environment in which topography changes only slightly, for instance that of rolling hills, and [iii] a lowland region or basin (Fig. 1A-D). Let us assume that a pair of sister taxa occurs throughout the lowland landscape but only one of the species occupies the rolling hills and the tall mountain (Fig. 1E). If these three landscape configurations were equally represented in geographic space, then the use of an environmentally based correlative distribution model followed by geographic projection (akin to G-space metrics, Warren et al. 2008) would conclude that the niches of the two species are 33% similar. However, if the montane and hilly landscapes each represented only 3% of the area in geographic space, with 94% of the region being occupied by lowlands, then the same method would infer that these species' niches are 94% similar (Fig. 1F). Although this example oversimplifies the calculations, it demonstrates the potential pitfalls associated with counting pixels in geography when ranges should be inferred from environmental correlations. Because landscapes (thus habitats and environments) are rarely equally distributed in natural systems, quantifying niches in geography likely over- or under-estimates niche similarity merely based on the geographic coverage of key environmental parameters - unless the distributions of the two species being compared are identical.

Several studies have proposed fixes to G-space limitations by focusing on analyses in environmental-space (E-space; Broennimann et al. 2012, Qiao et al. 2016, Nunes and Pearson 2017). At time of writing, the most popular E-space methods are those of Broennimann et al. (2012). They are comprised of a pair of statistical analyses: [1] a Monte Carlo resampling statistic aimed to assess how similar two niches are, which is called an 'Equivalence' statistic, and [2] a spatial randomization statistic aimed to assess the power to detect a significant Equivalence statistic based on the range of environments included in the analyses, called a 'Background' statistic. More recently, Nunes and Pearson (2017) proposed a single test for inferring Phylogenetic Niche Divergence (or conservatism) based on a Random Translation and Rotation (RTR) statistic. We consider this a variant of the Background statistic because the strength of the test is dependent on the observed niche similarity values relative to the surrounding environment (vs. inter-taxon comparisons). Further, Qiao et al. (2016) also published a software packaged called NicheA, which provides a suite of tools to quantify and visualize E-space and G-space, and their explicit connections, but not to perform statistical analyses of niches.

Despite these many methodological improvements in niche quantification, a majority of researchers continue to overlook the transient nature of species distributions and assume that species have achieved equilibrium distributions and that their current geographic distributions reflect the nexus between suitable biotic space and suitable abiotic space (Fig. 2A). It is well known, however, that species' distributions are in a non-equilibrium state in most real-life situations be it due to seasonal differences (weather or biotically-related; Araújo and Pearson 2005, Peterson et al. 2011, Peterson and Soberón 2012) or to long-term dynamism in climate and barriers over evolutionary times (e.g., as glacial or monsoon cycles; Galbreath et al. 2009, Cheng et al. 2013, Calatayud et al. 2019). Species' ranges are snapshots in time, and likely fall short of representing all of the geographic or environmental-space that species are able to occupy in a given region, merely due to this non-equilibrium state. This has important implications for niche studies. For instance, if the range of a species is not at equilibrium (Fig. 2B), then current distributions may potentially fail to reflect the total range of physiological tolerances of the taxon in question. The practice of using range data to infer physiological limits (e.g., Wiens and Graham 2005, Kozak and Wiens 2006, Bonetti and Wiens 2014) is, therefore, risky. Similarly, describing and comparing niches based on snapshots of species' ranges poses a challenge that needs to be addressed (e.g., Hortal et al. 2008, Saupe et al. 2017).

To better characterize discussions of niches, particularly those in non-equilibrium states, we propose to expand Jackson and Overpeck's (2000) term 'potential niche' to characterize the portion of the existing fundamental niche that includes all favorable abiotic and biotic conditions present in a given region and time (Fig. 2A-B). Such biotic conditions could include the identity and abundance of mutualists, facilitators (e.g., pollinators, seed dispersers), predators, parasites, pathogens, and competitors that constrain or facilitate a species' distribution (Gaston 2003). Jackson and Overpeck's definition of a 'potential niche' was initially restricted to the favorable abiotic conditions available in a geographic area, which was recently renamed by Peterson et al. (2011) as the 'existing fundamental niche', providing a direct E-space analog to abiotically suitable area present in G-space. Peterson et al. (2011) also proposed the term 'biologically reduced niche' that is almost identical to our definition of potential niche. Peterson et al. (2011), critiqued Jackson and Overpeck's (2000) 'potential niche' stating: 'The term "potential niche" may be somewhat unfortunate, however, since it represents the currently existing manifestation of the fundamental niche (…) that is in reality available at the moment, rather than the species' potential'. We completely agree, and our proposed changes directly address this critique. We argue for this semantic change — using our updated definition of 'potential niche' instead of 'biologically reduced niche' — given that it more efficiently characterizes the core concepts and does not require a detailed understanding of a BAM diagram for casual comprehension. Lastly, our

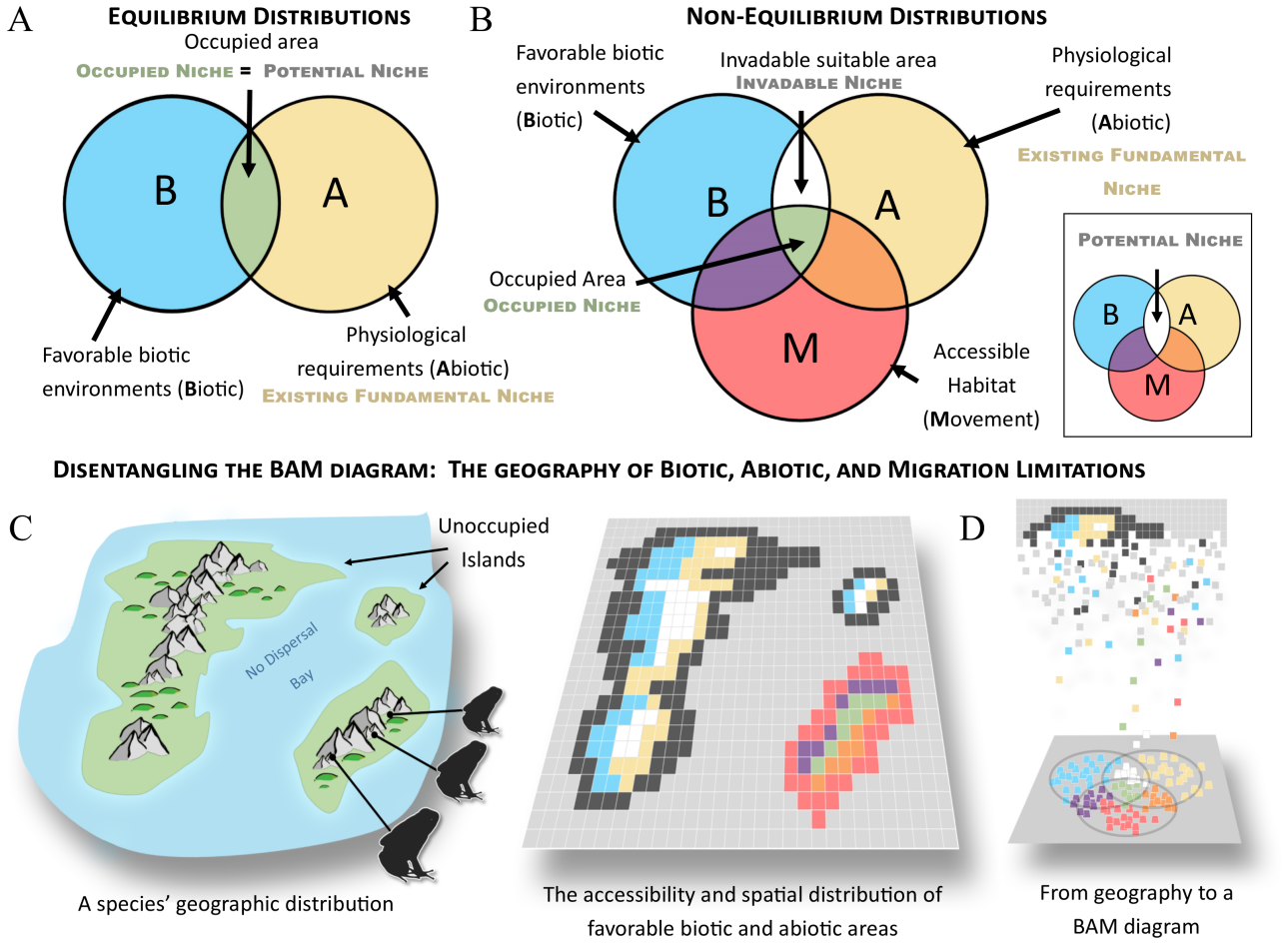


Fig. 2. Comparisons of equilibrium and non-equilibrium distributions and its role in occupancy of a species' existing fundamental niche. (A) If a species has an equilibrium distribution, it is occupying all of the potentially suitable habitats in the world (the potential niche is completely filled). (B) If a species is in a non-equilibrium distribution, its potential niche is not fully occupied due to seasonal and long-term dynamism of habitats. The BAM is cast in geographic space, corresponding E-space areas are labeled in colors and capital letters. BAM Fig. and terms adapted from Soberón and Nakamura (2009) and Peterson et al. (2011). (C) BAM units in connection to a species' geographic distribution. Dark grey pixels represent regions that are abiotically and/or biotically unsuitable areas, whereas light grey pixels are aquatic, unoccupiable habitats. (D) The spatial distribution of favorable factors is translated to a BAM diagram by pooling of pixels of corresponding regions into the diagram.

proposed change renders potential niche as a direct E-space analog to a species' 'potential distributional area' in G-space (Table 1) and, thus, discussions between E-space and G-space and their relationships are more intuitive.

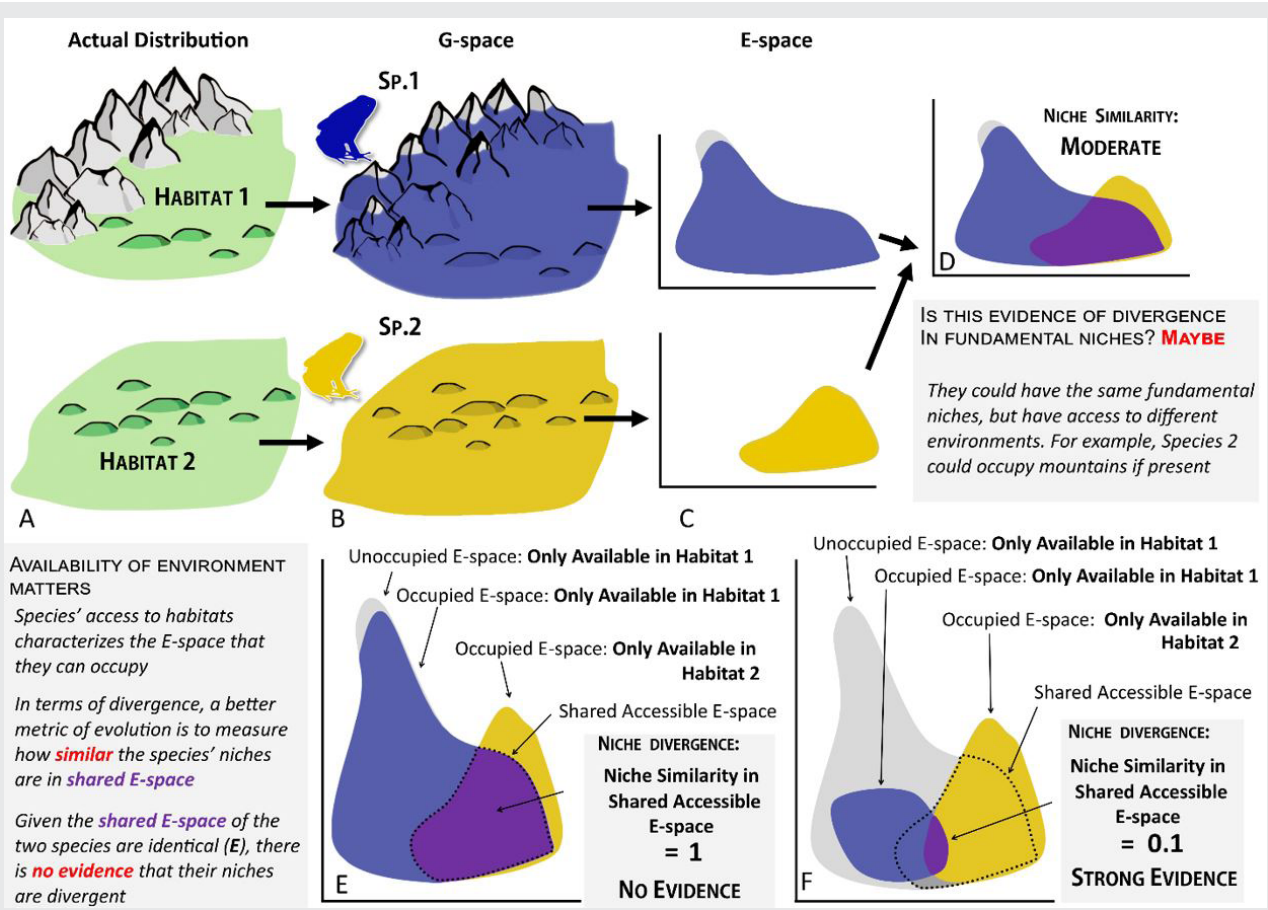
Recently, several researchers have begun to acknowledge issues associated with niche quantification in non-equilibrium distributions. For example, Petitpierre et al. (2012) demonstrated the importance of quantifying analog climates when comparing niche shifts among terrestrial plant invaders. Qiao et al. (2017) directly addressed non-equilibrium nature of species' distributions by restricting statistical analyses to accessible analogous climate space. The limited incorporation of non-equilibrium distributions into analyses, in part, is caused by the fact that no available software provides an intuitive or accessible way for researchers to quantify and incorporate non-equilibrium conditions into statistical tests of niche similarity.

In this paper, we aim to further progress by introducing an environmental-space (E-space) based Niche Equivalence Statistic that builds on the methods and statistics proposed by Broennimann et al. (2012), Petitpierre et al. (2012), and Qiao et al. (2017). Our method explicitly incorporates the spatial distribution of environments into statistical tests - particularly their availability, and whether environments are analogous to accessible climates of both species (Box 1, Fig. 3). Building on the methods of Warren et al. (2012), Beale et al. (2012), and Nunes and Pearson (2017), we introduce a new Background statistic to measure the ability of this Niche Equivalence statistic to detect differences based on the available E-space (Box 1, Fig. 3). These metrics enable fair comparisons between different

Table 1. Key relationships between geographic distributions and niches.

Geographic Space	*Ecological Space*
Abiotically suitable area	Existing fundamental niche
Potential suitable area	Potential niche
Invadable suitable area	Invadable niche
Occupied suitable area	Occupied niche*

*Hutchinson's definition of a 'realized niche' (Hutchinson 1957) is closely related to our definition of 'potential niche'. However, in past decades, the term 'realized niche' has been widely used to describe a species' occupied niche. To avoid confusion, we use 'occupied niche' here and not 'realized niche'.



Box. 1. Assessing niche divergence. If we simply look at the occupied E-space of each species (C), we would conclude they are quite different (D). However, we actually do not know if the yellow frog species is able to occupy mountains or not because no mountains exist in its current distribution (A & B). Given this spatial context of species' environments directly affect its distribution, any analysis of niche divergence must consider the spatial availability of habitats and make comparisons in only habitats that are available to both taxa (purple in D-F). If niches are very similar in shared accessible E-space (E), there is little evidence that niches have diverged. If they diverge in shared accessible E-space (F), then this is strong evidence that the species' niches have diverged.

geographies and when the ranges of species are out of equilibrium.

Based on distinct parameterization of the new Equivalence and Background statistics (Fig. 4), we then propose two corrected E-space-based statistical tests: a Niche Overlap Test (NOT) and a Niche Divergence Test (NDT) that, jointly, allow scientists to recognize differences between species that emerge from true niche divergence instead of other confounding causes such as differences in life history (e.g., mating systems or parental care types), differences in their biological interactors, or in the variety and configuration of accessible environments. Specifically, the NOT estimates the similarity between the occupied niches of the species; it considers the total accessible environmental-space represented within the geographic distribution of the species (for a general overview of entire study see Fig. 4). In turn, the NDT estimates the portion

of the accessible environment space that is shared by two species (herein called analogous accessible environments or analogous accessible E-space; Figs. 1, Supplementary Figs. S1-S3); it allows us to ask how equivalent (or not) the occupied niches of two species are given a common environmental background. When the NOT indicates significant differences in the total environmental-spaces occupied by the two species, there is support for the hypothesis that they currently occupy different niches - but we cannot state if the niches have diverged or if the similarity (or lack thereof) is due to other causes (see Table 2). If the NDT results in a significant value, it indicates that the niches of two species that share common accessible

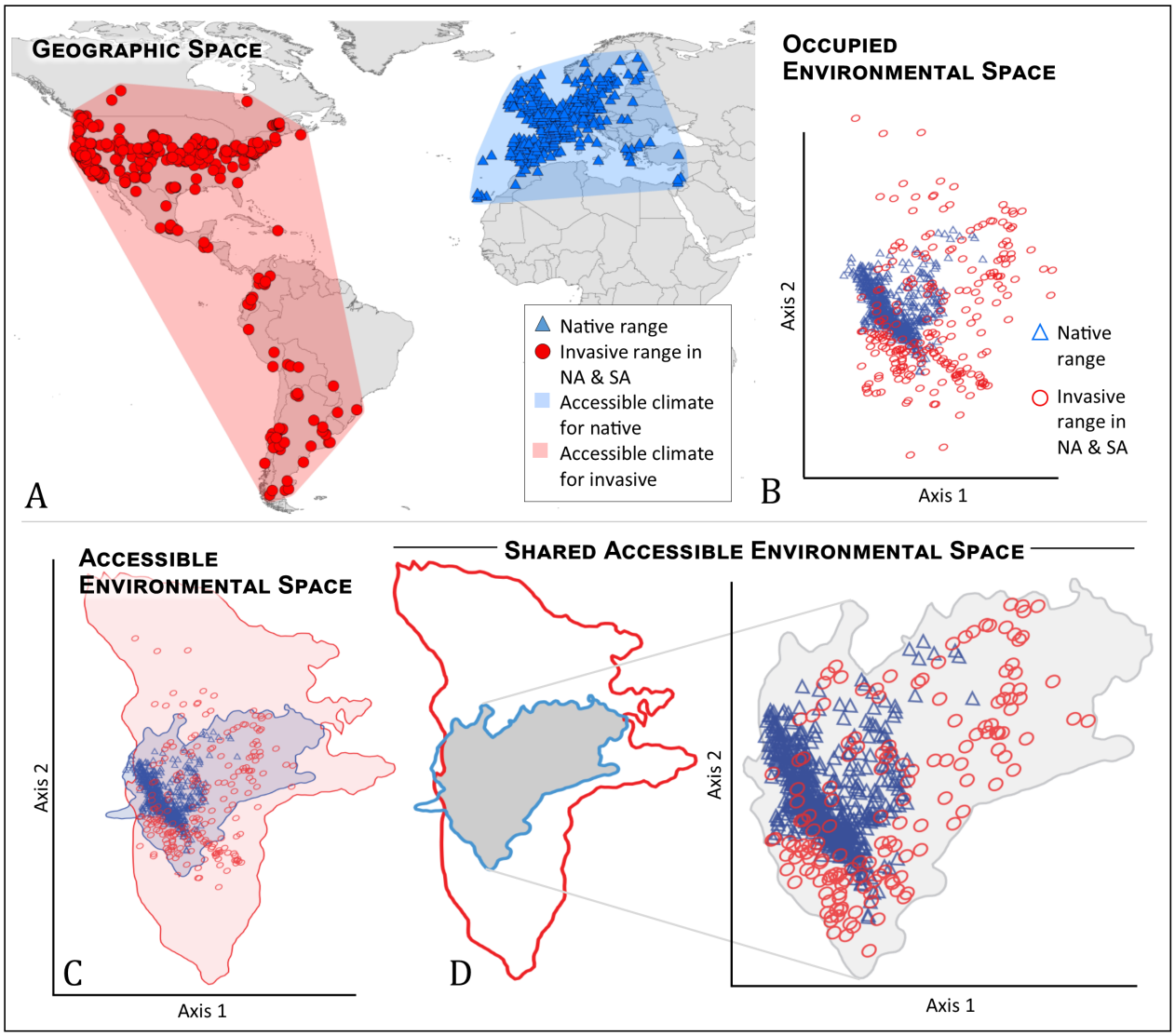


Fig. 3. Geographic space, accessible environments and environmental-space. A. Species interact with the environment via their geographic distributions. Within each species distribution is a subset of the total environmental-space which it has access to. This environmental-space is estimated by creating a buffered minimum convex polygon of each species' distribution. Pictured is the geographic distribution of *Conium maculatum* (poison hemlock) that is native to Europe and invasive in North and South America. **B.** Every geographic distribution (G-space) can be characterized by its environmental-space (E-space), which displays the distribution of occurrences in environmental data (as opposed to their physical geographic locations). Pictured is the occupied E-space of poison hemlock and accessible E-space. **C.** Every species has access to some E-space characterized by its distribution potential, biotic factors, and the composition of environments surrounding its realized distribution. **D.** When comparing two species (or populations), typically a portion of the available environmental-space is shared with the E-space of the other species (or population). The portion of shared E-space (shaded and blown-up on the right with corresponding localities of Poison Hemlock) we call accessible analogous E-space. On the other hand, the accessible E-space unique to each species' accessible environments, if present, we call non-analogous E-space (non-shaded areas).

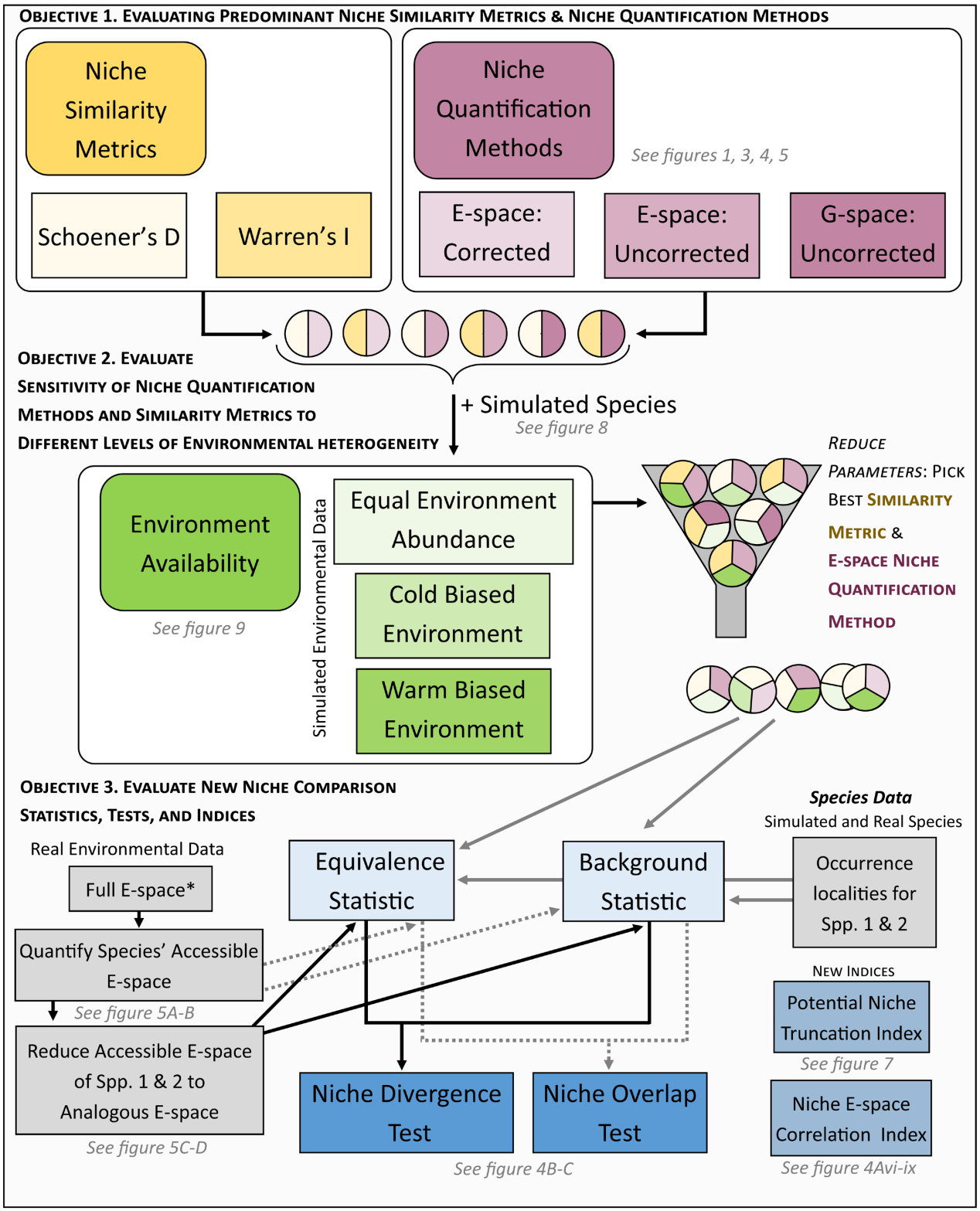


Fig. 4. Study overview. We explore how varying levels of habitat heterogeneity affect niche similarity indices and niche quantification methods (Objective 1), and which parameters best predict known relationships between our two simulated species (Objective 2). Then, we apply the best performing niche similarity metrics and E-space quantification methods to evaluate our new methods. We compare results from our new tests to two of the most commonly used niche divergence methods available, using both simulated species in real environments and a real species in real environments.

Table 2. Key to interpreting 'humboldt' results. **significant test, NS= non-significant test, *a significant divergence test could equally be reflective of differences in favorable biotic factors between test regions with equal fundamental niches. The biotic factors can include the identity and abundance of facilitators (e.g., pollinators, seed dispersers), predators, parasites, pathogens, and competitors that constrain or facilitate a species distribution (Gaston, 2003).

Niche Overlap Test		Niche Divergence Test		Interpretation
Equivalency Statistic	**Background Statistic**	**Equivalency Statistic**	**Background Statistic**	
**	** or NS	**	** or NS	Strong evidence that the species' niches have diverged*
**	** or NS	NOT POSSIBLE-NO OVERLAP	-	Strong evidence that the species' niches have diverged*
NS	**	**	**	The species' niches have diverged; however, there are no significant differences when looking at total distribution. Typically, this result suggests the relationship in the NDT is marginally significant ($0.05 < P < 0.1$). This can occur if data were not properly rarefied or environmental correlations need to be removed from the niches of species (see Corrected E-space).
**	**	NS	**	This supports the hypothesis that their current niches are not equivalent. However, these results do not support that the hypothesis is the result of divergent evolution; rather, differences are simply a result of different access to environments.
NS	NS	**	** or NS	Evidence that the species' niches have diverged*. In this case, the NOT is not significant in both the Equivalence statistic and Background statistic and is inconclusive. The NOT Background statistic suggests that the E-space available is similar to one or both distributions. Because of this, there is limited power to detect a difference even if one exists.
**	** or NS	NS	NS	This supports the hypothesis that their current niches are not equivalent. However, these results do not support that the hypothesis is the result of divergent evolution; rather, differences are simply a result of different access to environments. The NDT Background statistic suggests that the E-space of one species is similar to one or both species occupied distributions. Because of this, there is limited power to detect a difference even if one exists.
NS	NS	NS	NS	These results are inconclusive but suggests the niches are equivalent. However, this may be due to limited E-space available in both distributions.
NS	**	NS	**	Strong evidence that the niches are equivalent.

environmental-space are not equivalent, lending support for the hypothesis that their fundamental niches are the result of divergent evolution.

These novel tests are implemented in a new R package, 'humboldt', introduced here. Our methods differ from existing methods in several key aspects (Table 3). This comprehensive R package [1] facilitates quantification of a species' accessible E-space (not present in ENMtools, 'ecospat', and RTR), [2] provides a flexible framework to quantify analogous environments into statistical tests, which is important for assessing niche divergence in non-equilibrium distributions (not present in ENMtools, 'ecospat', and RTR), [3] provides statistical tests for comparing niches between species that occur in different geographic regions (analyses are restricted to the same region in ENMtools and RTR; no statistical niche tests are present in NicheA) and [4] between any taxa with suitable spatial data (in contrast to RTR, which is restricted to sister taxa). For a discussion of additional differences between 'humboldt' and 'ecospat', please see Discussion.

To evaluate the performance of the new tests, we use two simulated species with pre-defined niches: one able to tolerate both cool and warm habitats (akin to a species distributed both in lowlands and highlands), and one unable to tolerate cold conditions (lowland specialist). Using the simulated range of these species, we explore which parameters and settings provide the most accurate estimate of niche similarity among the two taxa, evaluating the impact of the choice of niche similarity index (Schoener's D and Warren's I), choice of Niche Equivalence statistics (G-based, uncorrected E-space, corrected E-space), and environmental availability (equal, warm-biased, cold-biased). Then, we apply the best performing metrics and parameters to compare the performance of NDT and NOT to two of the most commonly used niche divergence methods available – Warren et al.'s (2008, 2010) G-space analysis and Broennimann et al.'s (2012) E-space analysis, using both simulated species in real environments and a real species in real environments (see Fig. 4 for an overview of the study). We complement the statistical tests with an index that quantifies the potential for a species' occupied E-space to be truncated by the available E-space in its environment (thus providing context for cross-species comparisons) and a second index to reduce type 1 errors associated with different abundances of E-space across two species' distributions.

Materials and methods

Quantitative methods and statistical tests

Improving E-space-based metrics of niche equivalence, given background environments and the state of the art. To implement novel tests that

Table 3. Comparison of Niche Analysis Programs

Niche Analysis Programs	G-space	E-space	Based on ENMs	Tests between different regions	Not Restricted to sister taxa	Accounts for analogous E-space
EnmTools	X		X		X	
Ecospat		X		X	X	
RTR		X	X			
NicheA		X		X	X	
Humboldt		X		X	X	X

comparisons continued below

Niche Analysis Programs	Accounts for species' accessible E-space	Performs Equivalence test	Performs a Power Test or Background test	Assesses Potential Niche Truncation	N° potential niche dimensions	Citation
EnmTools		X	X		2	Warren et al. 2012
Ecospat		X	X		2	Di Cola et al. 2017
RTR			X		not restricted	Nunes & Pearson 2017
NicheA	X				3	Qiao et al. 2015
Humboldt	X	X	X	X	not restricted, but tests are performed in 2 dimensions	this manuscript

evaluate the similarity of niches between two species, we created the R package 'humboldt'[1], building upon the work of Jackson and Overpeck (2000), Warren et al. (2008), Beale et al. (2009), Broennimann et al. (2012), Petitpierre et al. (2012), Qiao et al. (2017), and Nunes and Pearson (2017). In 2008, Warren and colleagues proposed a pair of quantitative statistics and associated tests to assess niche similarity: [i] an Equivalence test, which assessed whether two niches are equivalent based on correlative distribution models and [ii] a Similarity test (to be applied when Identity tests were non-significant), which asked whether the two niches are simply more similar than expected by chance. The Similarity test is aimed to test the power of the Equivalence test, asking whether two distribution models are equivalent due to matching environments available in the habitat. If habitats contain identical environments, then species' niches could be statistically equivalent solely due to the lack of difference in the environments to which both species were exposed. These tests were implemented in the software ENMtools (Warren et al., 2010), where the Equivalence test was renamed as the Identity test, and the Similarity test was renamed as the Background test. For an introduction to the syntax and a visual explanation of the parameter options in 'humboldt', see Appendices 1 and 2. For a visual guide to interpreting the associated analyses and output figures see Appendix 3.

In 2012, Broennimann et al. introduced two complementary tests in E-space, seeking to address the caveats associated with comparisons in G-space (Warren et al. 2008, 2010). Though the two tests were conceptually similar to Warren et al.´s. (2010) Identity and Background tests, Broennimann et al. (2012) maintained the original test names of Equivalence and Similarity tests proposed by Warren et al. (2008), respectively. Yet while advancing on issues of environment availability and addressing the G-space testing limitation, Broennimann et al.'s test did not explicitly deal with non-equilibrium conditions in species' distributions. In 2017, Di Cola et al. incorporated Broenniman et al.'s (2012) methods into the R package 'ecospat'.

To provide an E-space-based framework to statistically compare niches that is neither impacted by the spatial distribution of environments (differently from Warren et al. 2008) nor relies on assumptions of that species' distributions are in equilibrium states (differently from Broennimann et al. 2012 and Di Cola et al. 2017), we modified the existing quantitative statistics, as described below. However, we recycled the nomenclature used by both Warren et al. (Warren et al. 2008, 2010) and Broennimann et al. (2012), choosing statistic names that best describe the actual underlying statistical procedures. Thus, we propose a modified E-space-based Equivalence Statistic (re-using the name and general resampling methods of both Warren et. al. 2008, Broennimann et al. 2012, and Di Cola et al. 2017) but use the term Background statistic (used by Warren et al. 2010) to evaluate the power to detect differences between the two groups, based on available environmental conditions.

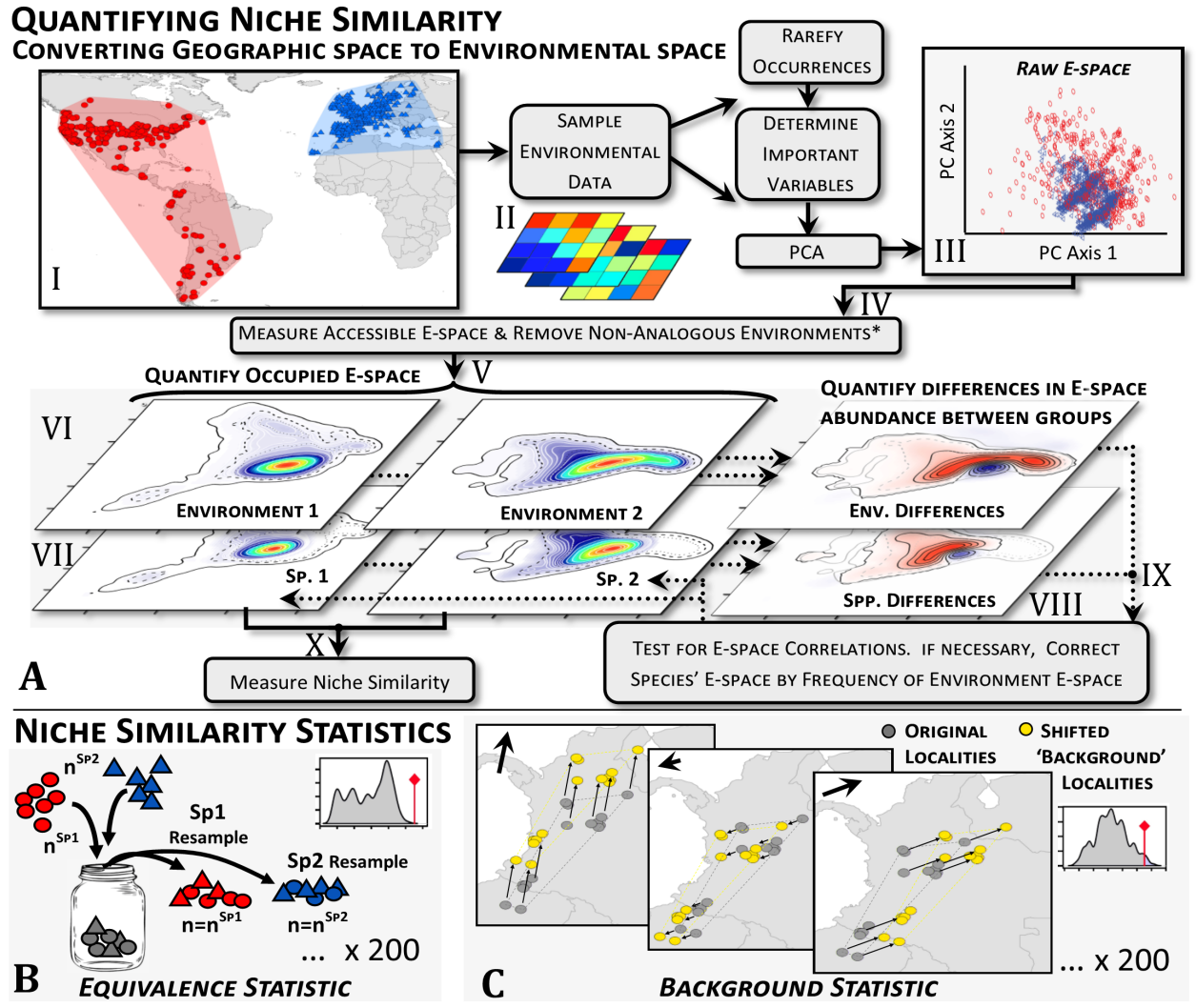
To avoid confusion hereafter, all discussions of Identity/Equivalence and Background/Similarity statistics and tests will use 'equivalence' and 'background' in reference to the corresponding statistics and tests. Also, in this manuscript we distinguish between 'statistic', referencing to the mathematical function (the statistical algorithm) and 'test', to characterize that a statistic is being used to test a hypothesis. We do so in an effort to reduce confusion between the two because here we implement distinct parameterizations of Equivalence and Background statistics as two separate statistical tests: a Niche Overlap Test (NOT) and a Niche Divergence Test (NDT, discussed below).

Quantifying E-space and Niches in E-space. We characterized E-space as two axes of a Principal Component (PC) analysis of input environmental variables across an entire study region of both species. As implemented in 'humboldt', this can include any combination of two output PCs. However, since our simulated species' fundamental niches are defined by two bioclimatic variables, we limited our analyses to the first two PCs. Following Broennimann et al. (2012), a kernel density function (Benhamou and Cornélis 2010) was used to create a continuous E-space surface in a grid of 100 x 100 cells, using the PC values from either the input occurrence localities or study region data to estimate the occupied E-space of the focal species or its environment, respectively.

Using Warren's I and Schoener's D to estimate the similarity between the niches of two species. To compare our results with those provided by G-space-based methods (Warren et al. 2008) and an existing E-space based method (Broenimann et al. 2012), we quantified the degree of similarity between the niches of two species in either G-space or E-space using two common metrics: Warren's I and Schoener's D. Both metrics output niche similarity values from 0-1. A value of 1 signifies niche equivalency, while a value 0 signifies perfect niche divergence. Warren's I (Warren et al. 2008) is a measurement derived from Hellinger (Hellinger 1909) and equals one minus Hellinger's distance (as measured between two niches). Schoener's *D* (Schoener 1968, Schoener and Gorman 1968) equals one minus the total variation distance between two niches.

A new Equivalence statistic. The R-package 'humboldt' runs a modified niche Equivalence statistic based on the niches of two species, quantified in E-space, and estimates the portion of the accessible environment space that is shared by both species (herein called shared analogous environments (SAE) or shared analogous E-space). The statistic calculates how similar the occupied niches of two species are to each other, by calculating Warren's I and Schoener's D, and compares these indices to those obtained when the occurrences of the two species are resampled (Box 2). During each resampling iteration, occurrences of species 1 and 2 are pooled and then randomly assigned

1 https://github.com/jasonleebrown/humboldt.git



Box. 2. Assessment of niche similarity. A. Quantifying niche similarity. I) Occurrence records and accessible environments are sampled. II) Corresponding environmental data are sampled at each species' occurrence sites. Occurrence records are rarefied to reduce spatial autocorrelation of localities. Relevant environmental variables are determined for each species and total environment data are reduced to these variables, which are in turn reduced to two dimensions in a standardized principal component analysis. III) The first two principal components (PCs), or any other pairs of PCs, of relevant environment data are plotted in two dimensions to characterize the raw environmental-space occupied. IV) Depending on test, Niche Overlap test or Niche Divergence test, the occupied environmental-space may be reduced. V) Raw environmental-space is converted to a kernel density representing the species' occupied E-space. The E-space of both VI) environments and VII) species are quantified. VIII) The difference between environmental E-space and species' E-space is quantified, and IX) correlations among these differences are assessed. The red and blue coloration in plots depict areas where that E-space is more abundant in environment 1 and environment 2, respectively. Niche and Environmental Correlation Index: If a high correlation exists in environmental difference between sites, then species' E-space should be corrected by the availability of E-space in their respective habitats. Niche similarity is quantified between both species' E-space. **B. Equivalence statistic**. To assess the significant equivalence of both species' distributions, occurrence localities are pooled and resampled in two groups equaling the number of localities in each. Niche similarity is then assessed and compared to the niche similarity of the observed data. This reshuffling is repeated several hundred times, each time comparing the resampled data to the observed. Significance is determined by the frequency that the observed overlap is greater than the reshuffled datasets. **C. Background statistic.** This statistic asks if the two distributed species are more different than expected given the underlying environmental differences between the regions in which they occur. The function compares the observed niche similarity between species 1 and 2 to the overlap between species 1 and the random shifting of the spatial distribution of species 2 in geographic space. It then measures how that shift in geography changes the occupied environmental-space. This statistic maintains most of the spatial structure of the input localities and thereby retains the nuances associated with each dataset's spatial autocorrelation.

to one of two groups. The number of occurrences in the two groups match the number of observations for species 1 or species 2, and, in each iteration, Warren's I and Schoener's D are measured between the two reshuffled groups. A null distribution is thus created from all values obtained from the reshuffled occurrences. The empirically derived measurements of similarity of the niches between species 1 and 2 (given by Warren's I and Schoener's D) are then assessed against the corresponding null distribution. When used as statistical test, a significantly small value of the empirical measurements, relative to the null distribution, rejects the null that species' niches are equivalent.

A new Background statistic. The Background statistic implemented in 'humboldt' measures the ability of the Equivalence statistic to detect differences based on the available E-space. It estimates the total environmental-space represented within the geographic distribution of the species and asks if the two species are more different than would be expected *given* the underlying environmental differences between the landscapes in which they occur. For that, the function compares the similarity of the niches of species 1 and 2, measured through Warren's I and Schoener's D, to the similarity between species 1 and the random shifting of the spatial distribution of species 2 in geographic space. Its goal is, thus, to evaluate how that shift in geography changes the occupied environmental-space (Box 2). The repeated random spatial shifting of localities, followed by the quantification of niche similarity between this shifted distribution and that of species 1, creates a null distribution of available E-space in the habitat of species 2. Note that this statistic maintains most of the spatial structure of the input localities and thereby retains the nuances associated with each dataset's spatial autocorrelation. If any points are initially shifted into areas without environment data, the points without environment data are shifted iteratively. Each round, if environment data are present in the new location, the environment is sampled, and that point is added back to the original dataset. This is repeated until all points have sampled areas with existing environment data. In practice, when clusters of points are shifted to areas of no environmental data, the entire cluster is subsequently shifted back into an area with data. Thus, in most cases the regional spatial autocorrelation is maintained.

A *non-significant* Equivalence statistic and a *significant* Background statistic support the hypothesis that the species niches are equivalent. A *significant* Equivalence statistic, regardless of the significance of the Background statistic, allows us to reject the null hypothesis of niche equivalence, and supports that species' niches are different. If both the Equivalence statistic and Background statistic are *non-significant*, it suggests that the perceived niche equivalency could be a result of the fact that the total environmental-space present in one or both landscapes is identical to one or both species' occupied niche(s) (Table 2). In these situations, there is limited power for the Equivalence statistic to actually detect significant differences among taxa, even if they exist. Importantly, however, the Background statistic does not provide any evidence that niches are not equivalent; it simply quantifies the power to detect significance based on the input environmental data.

A new metric to quantify the degree of potential niche truncation. Inferring the fundamental niche from a species' occupied niche remains a great challenge (Saupe et al.2017); most studies of niche divergence overlook how well (or how badly) the occupied niches characterized from contemporary distributions potentially characterize a species' fundamental niche. To provide a first step towards understanding this relationship, 'humboldt' provides a way to quantify the potential for a species' occupied E-space to be truncated by the available E-space in its environment (Fig. 5). The larger the proportion of the occupied niche that is truncated in E-space, the higher the risk that the occupied niche may poorly reflect the fundamental niche. Based on the relationship between the species' E-space and that available in adjacent habitats, we can assess the risk that the observed E-space is truncated and how likely we are underestimating the species' fundamental niches (Fig. 5). Here we introduce a new quantitative method to measure this: the Potential Niche Truncation Index (PNTI). It describes the amount of observed E-space of the species that is truncated by the available E-space. Specifically, it is a measurement of the overlap between the 5% kernel density isopleth of the species' E-space and the 10% kernel density isopleth of the accessible environment's E-space. The PNTI is the portion of the species' isopleth that falls outside of the environmental isopleth. This value physically measures how much of the perimeter of the species' E-space abuts, overlaps, or is outside the margins of the environment's E-space. If the value is large, there is moderate risk (PNTI= 0.15-0.3) or high risk (PNTI>0.3) that the measured occupied niche does not reflect the species' fundamental niche due to niche truncation driven by limited available E-space.

Addressing the non-equilibrium challenge - distinguishing between differences in niche similarity and significantly divergent niches. Most available studies of niche evolution or niche overlap assume that species have achieved equilibrium distributions and that their current geographic distributions reflect the nexus between suitable biotic space and suitable abiotic space (Fig. 2A; but see Petitpierre et al. 2012, Qiao et al. 2017). However, in most situations, species' distributions are likely in non-equilibrium (Fig. 2B). We address this through two new statistical tests based on distinct parameterizations of the Equivalence and Background functions implemented in 'humboldt'.

Niche Divergence Test (NDT). The first test, which we call the Niche Divergence test, estimates the portion of the accessible environmental-space that is shared by both species. NDT is, thus, the Equivalence and Background statistics performed in only analogous accessible environmental-space (Figs. 2, 5, 6); it allows us to ask if the species' niches are equivalent given a common environmental background (Figs. 3D-E, 5).

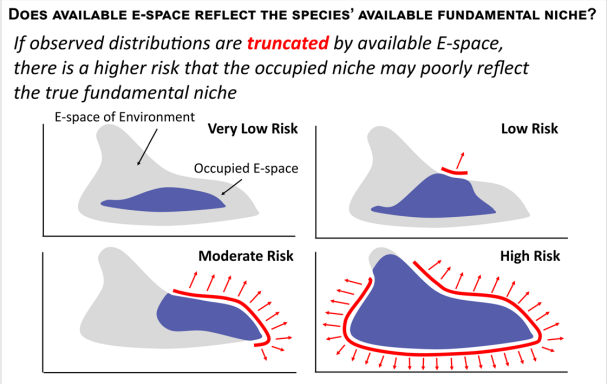


Fig 5. The quest for the fundamental niche. The paired visualization of the E-space of a region and its species provides insight into the risk that the focal species' observed niches (blue color) are truncated by the climate available in the species' distribution (grey). If a large portion of the species' E-space abuts the edge of the region's E-space, then this suggests that the species' E-space could be truncated and is larger. Pictured are four different scenarios and the risk that the observed distribution underestimates the species' potential and fundamental niches.

Niche Overlap Test (NOT). The second test, called the Niche Overlap test, estimates the total environmental-space represented within the geographic distribution of the species (Fig. Box 1D-E, Fig. 3). It corresponds to an Equivalence statistic performed in the total accessible environmental-space of both species' geographic distributions and allows us to ask how equivalent the two species' occupied niches are.

Note that these tests have different inference power. If the NDT results in a significant Equivalence statistic, it indicates that the niches shared accessible environmental-space is non-equivalent; thus, there is support for the hypothesis that their occupied niches are the result of divergent evolution. In turn, when the Equivalence statistic is performed in the scope of the NOT, it indicates significant differences in the total environmental-spaces occupied by the two species; there is support for the hypothesis that they occupy different niches, but one cannot affirm whether the niches differ due to divergent evolution or to asymmetries in habitat accessibility (see Table 2 and Supplementary Table S1, Box 1D-E), or other reasons. As typically implemented, the tests of ENMtools and 'ecospat' represent a form of a NOT.

Simulating niches to evaluate the performance of the new Niche Divergence and Niche Overlap tests

To test the performance of the tests implemented in 'humboldt', we compared the outcomes of the new NOT and NDT in E-space against NOT in G-space, as implemented in ENMtools v1.4.4 (Warren et al. 2010) and against NOT in E-space, as implemented in 'ecospat' (Di Cola et al. 2017). For that, we first used simulated species with pre-defined niches. With the R package 'virtualspecies' (Leroy et al. 2016), we created two simulated species whose tolerances were defined in two bioclimatic dimensions: maximum temperature of the warmest month (BIO5) and annual precipitation (BIO12, both variables from WorldClim 2.0; Hutchinson et al. 2009, Fick and Hijmans 2017). To evaluate how landscape complexity and the availability and abundance of environments differentially impact the new and existing niche metrics, we simulated a species that occupies both cool and warm conditions (simulated species 1, akin to a species that occupies both lowland and montane environments) and a species that does not occupy cool environments (simulated species 2, akin to a lowland species). The ecological tolerance of simulated species 1 was defined by a normal distribution of values corresponding to the maximum temperature of the warmest month, where environmental suitability is zero at 21^{0}C, increases to highest suitability at 26^{0}C and then decreases again until reaching zero at 31^{0}C (Fig. 6A top). For simulated species 2, the ecological tolerances were defined by cutting the normal distribution created for simulated species 1 in half at 26^{0}C, with values below 26^{0}C being unsuitable and values above 26^{0}C perfectly matching the suitability of simulated species 1. Although these simulated temperature affinities may not be realistic, we implemented them to ensure that the fundamental niche for species 1 is twice as large as simulated species 2 (Fig. 6A top). Both species share the same second niche dimension, represented by a logistic curve with a sigmoid midpoint at 1.6m of annual rainfall, between 0-0.6m from which suitability is zero. This created two rainforest species whose suitability is zero below 1 meter of annual rainfall, which goes to 0.5 habitat suitability at 1.6m of annual rainfall and achieves highest suitability at 2m annual rainfall and above (Fig. 6A bottom).

Exploring the impact of index choice and environmental availability on inferences of niche similarity

To test for Niche Divergence or Niche Overlap, we must first choose one index to measure niche similarity (Schoener's D or Warren's I). Maximizing our ability to accurately quantify niche similarity is of fundamental importance and is the foundation of the new Equivalence and Background tests. Thus, it is important that we carefully tune our niche quantification methods and use only the best performing indices (Fig. 4, objectives 1 and 2).

Uncorrected and Corrected E-space. To guide this choice, and its implementation, we ran exploratory analyses to assess the performance of these two Niche Similarity indexes through their direct quantification on a range of environmental datasets that the two simulated species. We also evaluated the impact of correcting niches in E-space, based on the availability of environments within a species' range, on the performance of the two indexes. Unlike niche quantifications in G-space, niche quantification in E-space can be rescaled based on the abundance of environments throughout the landscape. Following the E-space adjustments of

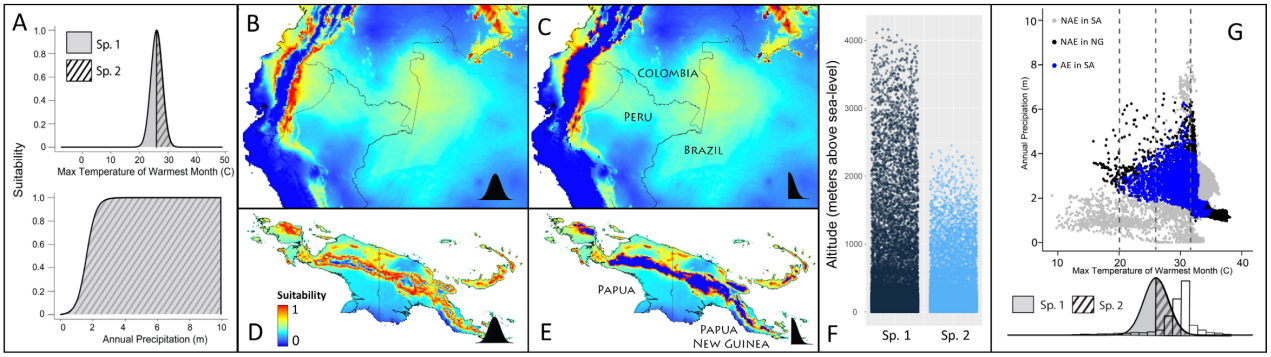


Fig. 6. Simulated species distributions in North-Western South America (SA) and the Island of New Guinea (NG). **A**. Response curves comprising the fundamental niches of simulated species 1 and 2 (Sp.1 and Sp.2, respectively). Top: The niche of Sp.1 is twice as large as Sp.2 with regards to the maximum temperature of the warmest month. Bottom: The niches of simulated species 1 and 2 are identical in annual precipitation requirements. **B.** Simulated istribution of Sp. 1 in SA. **C.** Simulated distribution of Sp.2 in SA. **D.** Simulated distribution of Sp.1 in ING. **E.** Simulated distribution of Sp.2 in ING. **F.** Altitudinal distribution of both species in SA and NG. **G.** Top: Environmental-space occupied by Sp.1 and Sp.2. X-axis is the maximum temperature of the warmest month and Y-axis is the annual precipitation. Grey and black dots depict non-analogous climates in SA and ING, respectively. Blue dots are analogous climate in both SA and ING. Bottom: species's fundamental niches and histogram of the frequency of values of the maximum temperature of the warmest month, which is the axis in which Sp.1 and Sp.2 differ. Lowland habitats are more abundant than cooler, higher elevation habitats in both landscapes (the histogram represents the frequency in both habitats).

Broennimann et al. (2012), we quantified species' niches in environmental-space in two different ways: (1) an Uncorrected E-space method that calculates a raw kernel density of environmental-space occupied and (2) a Corrected E-space quantification that standardizes the species' kernel density by the abundance of that E-space in the species corresponding E-space. The latter adjusts species' niches by the frequency the E-space is observed in the input landscape. Thus, it upweights observations in rarer E-space and downweighs observations in abundant E-space.

Quantifying Statistical Bias in Uncorrected E-space. A second issue associated with how habitat abundance affects niche quantification in E-space relates to environment-driven biases in statistical testing. To quantify species/environment correlations (Box 2A vi-ix) and determine if the occupied environmental-space of both species should be standardized by the abundance of habitats throughout their accessible environments, we created a new index, which we name the Niche E-space Correlation Index (NECI), and implemented it in 'humboldt'. The NECI first quantifies the abundance of E-space of accessible habitats and how the abundance of environments differs between the two study areas (Equation 1, ΔE). It then quantifies a standardized kernel density of E-space for both species and quantifies how the two species' E-space densities differ (Equation 2, ΔS). Whenever correlations between ΔE and ΔS are sufficiently high (e.g., >0.5; NECI), it is possible that the outcome of the Niche Similarity quantification (for example, low inferred similarity) is in fact driven by differences in the available environmental-space rather than true differences in the species' niches. Under this scenario, users should correct species' niches by the frequency of E-space in accessible environments to reduce type I errors (see discussion). Conversely, when the correlation is low between Δ_{Env} and Δ_{SPP}, differences in the availability of environmental-space are not correlated with differences between the two species' niches, and it may not be necessary to correct species' niches by the frequency of E-space in accessible environments (Box 2A vi-ix).

$$\text{Equation 1: } \Delta E = \left[\text{Espace}^{\text{Env1*}}\right] - \left[\text{Espace}^{\text{Env2*}}\right]$$

$$\text{Equation 2: } \Delta S = \left[\text{Espace}^{\text{Sp1*}}\right] - \left[\text{Espace}^{\text{Sp2*}}\right]$$

* E-space can be uncorrected or corrected

$$\text{Niche Espace Correlation Index} = \frac{n(\sum \Delta E \Delta S) - (\sum \Delta E)(\sum \Delta S)}{\sqrt{\left[n\sum \Delta E^2 - (\sum E)^2\right]\left[n\sum \Delta S^2 - (\sum S)^2\right]}}$$

The G-space and E-space niche similarity measurements reported here were calculated in ENMtools and 'humboldt', respectively. We chose to report a single value to keep the focus on the two similarity indices (as the equations for the niche similarity metrics are the same among 'ecospat' and 'humboldt'). These two methods often result in different test statistic values based on how the niche quantification is parameterized. However, in these particular comparisons, 'ecospat' resulted in similar values that were not significantly different among the four treatments (ANOVA, df=3, F= 1.189, p=0.441) when compared to Niche Similarity values generated in 'humboldt' (those reported here).

Lastly, we explored how different scenarios of environmental availability impacted the performance

of the two Niche Similarity indexes when applied under a G-space, uncorrected E-space, and corrected E-space (Fig. 4, Warren et al. 2010; Broennimann et al. 2012). To do this we created three different landscapes reflecting the maximum temperature of warmest month (BIO5): [1] a cold-biased landscape, [2] a warm-biased landscape, and [3] a landscape with equal environment abundance (Fig. 7). The three landscapes differ in the abundance of temperature values corresponding to the maximum temperature of the warmest month (the factor for which simulated species 1 and simulated species 2 differ). All three landscapes possess values ranging from 21.0^{0}C—31.0^{0}C in 0.1^{0}C increments, but differ in the frequency of the values. In the 'equal abundance' landscape, all temperature values were equally represented (Fig. 7A). In the 'cold-biased' landscape, 21.0^{0}C was a majority of values in the landscape, and the frequency of warmer values in landscape gradually decreases so that 31.0^{0}C, the warmest value, is the least frequent in the environment (Fig. 7B). In the 'warm-biased' landscape, 31^{0}C was a majority of values, with the frequency of cooler values gradually decreasing to 21.0^{0}C, the coldest value, which has the lowest frequency (Fig. 7C). Note that these two latter scenarios, though mirroring each other, are quite different; in the 'cold-biased' scenario, habitats suitable for simulated species 2 are rare, whereas in the 'warm-biased' scenario, habitats suitable for simulated species 2 are abundant. Because our tests required two dimensions of climate data, in all three scenarios the second environmental dimension represented annual rainfall with values 2.0 –10.0m (all values reflect maximum suitability for both simulated species), with each whole number being equally represented in each temperature values in the 'equal abundance' landscape. For each rainfall value, a single decimal place was randomized. This prevented this axis from binning in environmental-space (causing rows of densities for rainfall dimension) at environmental-space resolution of 100 x 100 grid.

Comparing the performance of the new niche overlap and niche divergence tests relative to former tests with simulated species projected into real landscapes

Niches of a simulated species and real environments. We also performed tests in real environments and projected both simulated species into two existing geographic regions of the world that would have been highly suitable for both simulated species: north-western South America and the Island of New Guinea. For that, we translated the niche of each simulated species into sampling localities for use in Niche Overlap and Niche Divergence tests by converting all grid cells with suitability values above 0.1 to 'presence' and then converting the raster pixels to individual points. Within each geographical region, 600 points were randomly selected from the range of each simulated species and used to test Niche Divergence and Niche Overlap. Analyses between north-western South America and the Island of New Guinea were not possible in G-space; therefore, comparisons in G-space were performed within each region, but not between.

We performed a pair of analyses among and within these two geographic regions. First, we compared the NOT and NDT performed in 'humboldt' to the analogous G-space tests performed in ENMtools and the analogous E-space tests in 'ecospat'. We performed two sets of comparisons, a (ideal) comparison where only 'true' environmental parameters were used to test Niche Overlap (i.e.,only maximum temperature of warmest month (BIO5) and annual precipitation (BIO12) were used for characterizing the niche). Both of these were

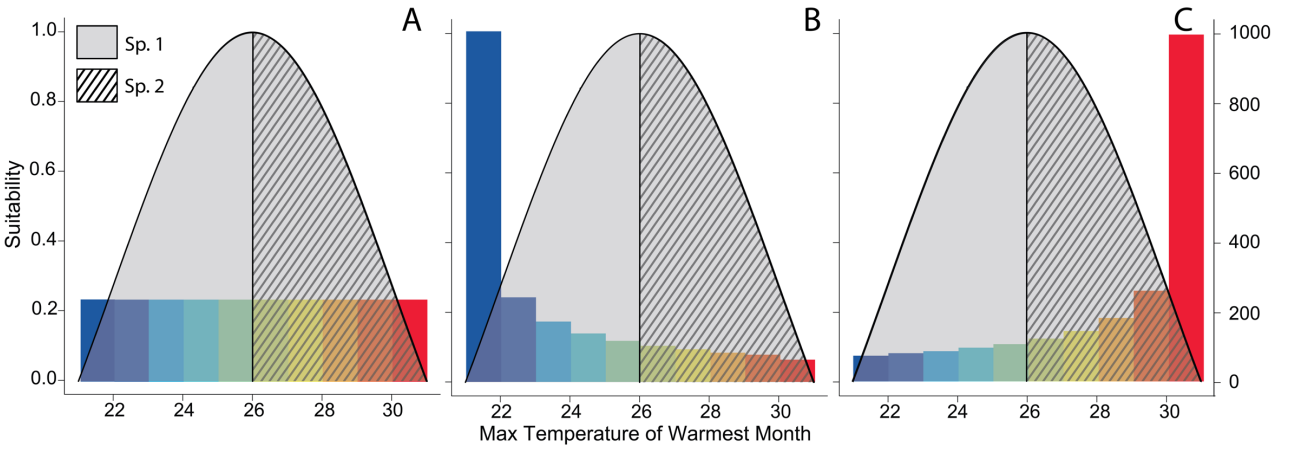


Fig. 7. Tests of environment heterogeneity on niche similarity measurements. We created three different landscapes reflecting the maximum temperature of warmest month (BIO5): **A.** equal environment abundance landscape, **B.** cold-biased landscape, and **C.** warm-biased landscape. The three landscapes differ in the abundance of temperature values corresponding the maximum temperature of the warmest month (the factor for which simulated species 1 and simulated species 2 differ). All three landscapes possess values ranging from 21.0^{0}C-31.0^{0}C in 0.1^{0}C increments but differ in the frequency of the values.

also used to define the species' niche; these results are presented in Supplementary Tables S2-S4). The second set compared how the above tests performed in a situation in which all 19 bioclimatic variables widely used in the biogeographical community (Austin and Van Niel 2011) were used for niche inference (Fick and Hijmans 2017). We did this because the latter scenario is more frequently followed by scientists as the true physiological limits are often unknown and researchers aim to estimate them based on patterns in the observed localities. In both cases, we employed environmental data at a 2.5 arc-minute spatial resolution.

The new NOT and NDT were run in E-space characterized only by those environmental variables that contributed 10%, or more, to a generalized boosted regression model of either of the two species (humboldt::humboldt.top.env). This prevented the naïve incorporation of all possible environmental variables into each species' environmental-space and defined a species' niche based only on environmental variables ranked as important for the aims of characterizing the species' range. This is recommended for all types of niche quantifications and comparisons in 'humboldt'. A similar process occurs when using species distribution models in ENMtools. In contrast, because variable selection is not common practice for 'ecospat' analyses, it was not used here to maintain general consistency with recommended practices (Di Cola et al. 2017).

To demonstrate how the new NOT and NDT perform on empirical data, and to evaluate if and how they can promote insight in real situations, we applied them to *Conium maculatum* (poison hemlock), a plant native to Europe and invasive in North and South American, and Asia (Vetter 2004). For that, we downloaded occurrence records from GBIF (GBIF, 2017) and vetted for accuracy, resulting in 4,977 occurrences records from the species' native range and 484 localities in North and South America. We rarefied the points at 40km^2 and sampled climate at a resolution of 5 min, using SDMtoolbox v2.3c (Brown et al. 2017). With this dataset in hand, we ran Niche Divergence tests, Niche Overlap tests, and the similar quantitative niche tests in E-space (using the methods of Di Cola et at. 2017), applying the same 'humboldt' input parameters to characterize and quantify niches used for the tests of the simulated species (Appendix 1).

Results

Assessing the accuracy of niche similarity indices with simulated species

Schoener's D was the index that most accurately inferred niche similarity of the simulated species, showing superior performance relative to Warren's I in all tests but one (the Warm-Biased landscape, see Table 3). These results were consistent in both simulated and real environments (*simulated environments*: p<0.001; mean absolute Schoener's D value minus true value= 0.106; mean absolute Warren's I value minus true value: 0.208; *real environments*: p<0.001; mean absolute Schoener's D value minus true value= 0.252; mean absolute Warren's I value minus true value: 0.208; Tables 4 and 5).

The effect of correcting niche comparisons by environment availability

Correcting niche quantifications by the abundance of available environments improved accuracy of niche similarity inference. Measured values of Schoener's D and Warren's I approximated the true values of niche similarity of the simulated species more closely whenever a correction was applied, relative to values measured in the absence of environmental correction (p<0.001 for both measurements using Schoener's D and Warren's I metrics; corrected niches: mean absolute Schoener's D value minus true value= 0.018; mean absolute Warren's I value minus true value: 0.206; uncorrected niches: mean absolute Schoener's D value minus true value= 0.260; mean absolute Warren's I value minus true value: 0.314; Tables 4 and 5).

The Niche E-space Correlation Index (NECI) was very high (0.782-0.899) in all comparisons that involved uncorrected environments, but was reduced whenever the E-space was corrected (NECI 0.178-0.462). Conversely, for comparisons within the same geographical area in uncorrected E-space, the NECI was low to moderate, 0.015-0.282, and was reduced to 0.011-0.126 in corrected E-space.

Effects of habitat heterogeneity on niche similarity inference

Our analyses demonstrate that the choice of niche similarity index (Schoener's D vs. Warren's I) directly impacts niche similarity inference under distinct scenarios of habitat heterogeneity. When Schoener's D was used, the estimates of niche similarity were significantly different among the three niche quantification methods (G-space, uncorrected E-space, and corrected E-space; ANOVA, F=7.53, p=0.076). In most spatial comparisons involving different scenarios of habitat heterogeneity, niche similarity values in corrected E-space were closest to the true values (mean absolute Schoener's D value minus true value= 0.018), followed by uncorrected E-space (mean absolute Schoener's D value minus true value= 0.260), and G-space being last (mean absolute Schoener's D value minus true value= 0.266; see Tables 4 and 5 for values). However, when using Warren's I, the estimates of niche similarity were not significantly different among the three niche quantification methods (G-space, uncorrected and corrected E-space; ANOVA, F=0.913, p=0.4274). Yet, niche similarity values in corrected E-space were closest to the true values in most spatial comparisons (mean absolute Warren's I value minus true value= 0.206), followed by G-space (mean absolute Warren's I value minus true value= 0.288), and then uncorrected E-space (mean absolute Warren's I value minus true value= 0.314; see Tables 4 and 5 for values).

Statistical tests in E-space using 'humboldt'. Measurements of the two simulated species' Niche Overlap and Niche Divergence from environment variables that were selected from all 19 Bioclim variables resulted in all six comparisons of the same simulated

Table 4. Niche similarity measurements: simulated species in different habitat heterogeneity levels. In each cell, the measured value and difference from the true value is depicted (in parentheses). To summarize performance, we summed the total differences (value is in the Sum of Error row). The lower the value, the closer measured values were to the true values. Tests were carried out in both geographic space (G-space) and environmental-space (E-space) within three different levels of habitat heterogeneity: equal, warm-biased, and cold-biased. Two niche similarity metrics were evaluated in all E-space and G-space metrics: Schoener's D and Warren's I. The environmental-space tests were carried out under two scenarios: an **uncorrected E-space scenario** where species' niches reflect the raw kernel density quantification of the E-space that they occupy, and a second **corrected E-space scenario** where species' niches reflect a standardized kernel density that correct the species' observed E-space densities by the frequency of E-space in the corresponding environment.

Scenario	I: *G-space*	D: *G-space*	I: *E- space corrected*	D: *E- space corrected*	I: *E- space uncorrected*	D: *E- space uncorrected*	True Value
Equal	0.65 (0.15)	0.50 (0)	0.72 (0.22)	0.52 (0.02)	0.72 (0.22)	0.52 (0.02)	0.5
Warm-Biased	0.60 (0.10)	0.36 (0.14)	0.71 (0.21)	0.49 (0.01)	0.47 (0.03)	0.23 (0.27)	0.5
Cold-Biased	0.80 (0.30)	0.65 (0.15)	0.74 (0.24)	0.54 (0.04)	0.90 (0.40)	0.80 (0.20)	0.5
Sum of Error	0.55	0.29	0.67	0.07	0.65	0.49	-

Table 5. Niche similarity measurements: simulated species in real environments. In each cell the measured value and difference from the true value is depicted (in parentheses). To summarize performance, we summed the total differences (value is in the Sum Delta row). The lower the value, the closer measured values were to the true values. Tests were carried out in both geographic space (G-space) and environmental-space (E-space) within two geographic regions: north-western South America (SA) and the Island of New Guinea (NG). Two niche similarity metrics were evaluated in all E-space and G-space metrics: Schoener's D and Warren's I. The environmental-space tests were carried out under two scenarios: an **uncorrected E-space scenario** where species' niches reflect the raw kernel density quantification of the E-space that they occupy, and a second **corrected E-space scenario** where species' niches reflect a standardized kernel density that correct the species' observed E-space densities by the frequency of E-space in the corresponding environment.

Scenario	I: *G-space*	D: *G-space*	I: *E- space corrected*	D: *E- space corrected*	I: *E- space Uncorrected*	D: *E- space uncorrected*	True Value
SA Sp1 vs. SA Sp2	0.98 (0.48)	0.95 (0.45)	0.70 (0.2)	0.49 (0.01)	0.95 (0.45)	0.88 (0.38)	0.5
NG Sp1 vs. NG Sp2	0.91 (0.41)	0.83 (0.23)	0.68 (0.18)	0.51 (0.01)	0.93 (0.43)	0.83 (0.33)	0.5
Sum of Error	0.89	0.68	0.38	0.02	0.88	0.71	-

species supporting the hypothesis that their measured niches are equivalent (Table 6). They also perfectly quantified niche similarity (average niche similarity= 1, Table 6, Supplementary Tables S2-4). However, when comparing the same species between environments, the measured overlap was lower than 1 (mean niche similarity in analogous environments= 0.547, total E-space=0.522). When comparing simulated species 1 to simulated species 2 in the same and different geographic regions, two (of four) NDT and NOT resulted in the rejection of the null hypothesis and recovered the divergent niches. In the case of non-significance in all inter-species comparisons, at least one of the Background statistics was also non-significant, suggesting that in that particular case there exists limited statistical power to actually recover a difference (see Table 6). Overall, eight (of ten) NDT and NOT matched the expected relationships (i.e., divergence between different species and equivalence niches within species comparisons; Fig. 8).

Statistical tests in E-space using 'ecospat'. Measurements of Niche Overlap between the two simulated species, using environmental variables selected from all 19 Bioclim variables, resulted in all but one tests supporting the hypothesis that their measured niches are equivalent; similarity measurements resulted in perfect quantification of niche similarity (average niche similarity= 1, Table 6, Supplementary Tables S2-S4). However, when comparing the same species between environments, the measured similarity was lower than 1 (mean niche similarity= 0.408). When comparing simulated species 1 to simulated species 2 in the same geographic region (e.g., within South America) and different geographic regions (e.g., between South America and the Island of New Guinea), all four Equivalence tests resulted in the acceptance of the null hypothesis that the two species are equivalent and did not infer that the niches are divergent. For all inter-species comparisons, at least one of the Background tests was also not significant, suggesting there exists limited statistical power to actually recover a difference (see Table 6). Overall, five (of ten) testEquivalence tests matched the expected relationships (i.e., divergence between

Table 6. Niche Overlap and Divergence of simulated species. All Bioclim variables. $\mathbf{D}_{obs}$ corresponds to the Niche Similarity Index quantified with Schoener's D measurement. $\mathbf{D}_{true}$ is the expected Niche Similarity. $\mathbf{E}_{obs}$ is the observed significance of the Equivalence statistic, and $\mathbf{E}_{true}$ is the expected relationship (S= significant and NS= non-significant). $\mathbf{B}_{2->1}$ and $\mathbf{B}_{1->2}$ corresponded to Background statistics comparing simulated species 2 to simulated species 1 or simulated species 1 to simulated species 2, respectively. In both cases, the first listed species is the one whose range was shifted. Tests significance: α =: * 0.01-0.05, ** at 0.01-.001, *** < 0.001

Scenario	Niche Test	D_{obs}	D_{true}	E_{obs}	E_{true}	$B_{2->1}$	$B_{1->2}$
***Method-ENMTOOLS:* G-space**							
SA1 x SA2	NOT	0.893	0.5	1	S	***	***
NG1 x NG2	NOT	0.853	0.5	1	S	***	***
***Method-ECOSPAT:* E-space**							
NG1 x SA1	NOT	0.525	1	0.010	NS	0.035	0.526
NG2 x SA2	NOT	0.291	1	1	NS	*	0.075
NG1 x SA2	NOT	0.291	0.5	0.995	S	0.085	0.099
NG1 x NG2	NOT	0.506	0.5	1	S	0.070	**
SA1 x SA2	NOT	0.486	0.5	0.995	S	0.065	*
SA1 x NG2	NOT	0.403	0.5	0.184	S	0.403	*
***Method-HUMBOLDT:* E-space**							
NG1 x SA1	NOT	0.543	1	1	NS	***	*
NG1 x SA1	NDT	0.565	1	1	NS	***	***
NG2 x SA2	NOT	0.513	1	1	NS	*	0.163
NG2 x SA2	NDT	0.498	1	0.991	NS	***	**
NG1 x SA2	NOT	0.434	0.5	0.980	S	***	0.148
NG1 x SA2	NDT	0.472	0.5	0.960	S	*	0.163
NG1 x NG2	NOT	0.541	0.5	***	S	***	***
NG1 x NG2	NDT	0.541	0.5	***	S	***	***
SA1 x SA2	NOT	0.586	0.5	*	S	***	***
SA1 x SA2	NDT	0.568	0.5	**	S	***	***
SA1 x NG2	NOT	0.464	0.5	1	S	*	0.089
SA1 x NG2	NDT	0.430	0.5	0.792	S	*	0.158

different species and equivalent niches within species comparisons; Fig. 8).

Statistical tests in G-space. All Equivalence tests in G-space that were based on the 19 Bioclim variables supported the null hypothesis of niche equivalence (P=1.000) in comparisons between species and within the same species— failing to recover significant differences in all tests (n=2) where divergence was expected (Table 6, Supplementary Tables S2-S4). Overall, four (of six) Equivalence tests matched the expected relationships (i.e., divergence between different species and equivalence niches within species comparisons; Fig. 8). However, all four of those were control tests, which assessed niche equivalence of a species to itself.

Potential niche truncation of simulated species in real environments

The two simulated species displayed varying levels of potential niche truncation (measured in both South America and the Island of New Guinea). Simulated species 1 displayed varying Potential Niche Truncation Index values between 0.124-0.365, corresponding with low to high levels of potential niche truncation in both the north-west South America and Island of New Guinea. In contrast, simulated species 2 consistently exhibited very high potential niche truncation index values (0.593-0.682) in the same regions.

Niches of a real species

The results of our analysis on *Conium maculatum* demonstrate that the occupied niches of native populations and invasive American populations are quite different (Niche Overlap test: D= 0.068, P< 0.001, Table 7). However, despite occupying considerably different E-space in their total ranges, in shared analogous environments, the species' niches are significantly equivalent (Niche Divergence test: D= 0.218, P=1.000). Thus, there is no evidence that the species' niches have diverged. The measured Potential Niche Truncation Index varied considerably between Europe and the Americas (0.408 and 0.146, respectively, and corresponds to 'high' and 'some' values of potential niche truncation). These results agree with the NOT and NDT results, with the range of accessible E-space in Europe being much smaller

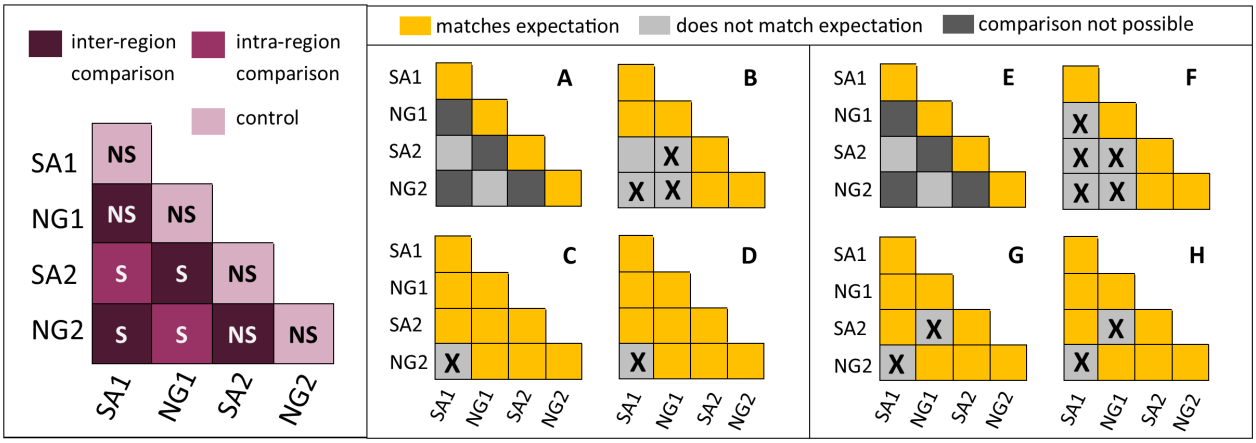


Fig. 8. Summary of results for Equivalence statistic among methods. Left. The triangle matrix characterizes the three comparisons considered: intra-region comparisons (i.e., within SA), inter-region comparisons (between SA and NG), and controls (e.g., SA1 vs. SA1). The Simulated Species 1 and Simulated Species 2 from the two regions (SA1, NG1, SA2 and NG2) are indicated in each half matrix. For example, SA1 corresponds to Simulated Species 1 in South America. The comparisons of the same simulated species in either region are expected to be non-significant, whereas, in comparisons among the two species, we expect Equivalence statistics to be significant. **Center** plots depict results from Equivalence statistics using only the maximum temperature of the warmest month and annual precipitation (see Tables S2 and S4), whereas the **right** group of plots depict results from Equivalence statistics using all 19 bioclim variables (Table 6 and Table S3). Yellow indicates the results of Equivalence statistics match expected significance and non-significance (summarized in left box). Light grey shows that the results do not match expectations, and 'X's within the grey squares represent non-significant Background statistics (suggesting that there is limited statistical power to detect differences). Dark grey shading depicts that the method cannot perform the comparison. The letters correspond to the following methods: **A & E.** ENMtools, **B & F.** 'ecospat', **C & G.** 'humboldt'- overlap test, **D & H.** 'humboldt'– divergence test. SA= North-Western South America, NG= Island of New Guinea.

Table 7. Niche Divergence and Niche Overlap tests on *Conium maculatum*. D corresponds to the Niche Similarity Index quantified with Schoener's D measurement. **E** is the observed significance of the Equivalence statistic. In all tests, the pairs of Background statistics (1->2 and 2->1) were significant at $\alpha < 0.05$.

Method	Test	D	E
'ecospat'	NOT	0.346	0.020
'humboldt'	NOT	0.068	<0.001
'humboldt'	NDT	0.218	1

than in the Americas. The values suggest E-space in Europe represents a sub-portion of the species' fundamental niche.

Discussion

The new methods introduced here translate several important theoretical advances into tests of niche divergence that allow researchers to more accurately estimate whether species have actually evolved different niches or if they occupy different environmental-spaces as the result of differences in life history, their biological interactors, or in the variety and configuration of accessible environments (e.g., Hardin 1960, Gaulin and Fitzgerald 1988, Garcia-Barros and Benito 2010, Grossenbacher et al. 2015, Estrada et al. 2015, Borda-de-Água et al. 2017). The foundation of these improvements is based on the underlying assumption that most species' contemporary distributions are in non-equilibrium state (Cheng et al. 2013, Calatayud et al. 2019) and, because of this, the geographic manifestations of their niches (occupied, potential, and available fundamental) are dynamic through time (Araújo and Pearson 2005, Peterson and Soberón 2012). The new methods provide several quantitative advances that characterize the accessible climates in both species' distributions and the corresponding relationship between non-analogous and analogous climates. Overall, the new methods improve the accuracy of niche similarity quantifications and corresponding statistical tests, consistently outperforming similar tests in correctly quantifying niche equivalence and divergence in simulated data with known truths (Tables 4-6, Supplementary Tables S2-S4).

Quantifying niche similarity

In estimates of simulated species' niches in both G-space and E-space, Schoener's D (Schoener 1968, Schoener and Gorman, 1968) consistently outperformed a measure derived from Hellinger (Hellinger, 1909) and Warren et al. (Warren et al. 2008), commonly called Warren's I. Thus, for general use, we recommend using Schoener's D for measuring niche similarity in

'humboldt' and other methods. These results agree with a similar study by Rödder and Engler (2011).

Quantifying and comparing niches in E-space and G-space

Our results demonstrate that measurements in geographic space are only accurate when important environmental variables are equally represented in geographic space (Fig. 7, Tables 4 and 5). In natural systems, environmental variables are rarely equally disturbed, and researchers are likely over- or under-estimating niche similarity based on the distribution of key environmental-space across geography. A second major limitation to assessment of niches in G-space is the requirement that species occupy the same geographic area before you can assess niche similarity. This is despite the fact that many analogous environments might occur elsewhere.

The final limitation is the lack of environmental context provided by analyses in G-space. Analyses in E-space provide explicit context for how the species' niche is characterized by the available environments. Based on the relationship between the species' E-space and that available in adjacent environments, we can assess the likelihood that the observed E-space is truncated and how likely we are to underestimate the species' fundamental niche. Further, unlike niche quantification in G-space, niche quantification in E-space can be rescaled by the abundance of environments throughout the landscape. In our study this worked very well; we recommend doing this when the species occur in two distinct environments or regions of the same larger environment.

Detailed comparisons of the niche Equivalence tests of 'ecospat' to 'humboldt'

Our methods were heavily influenced by those of Broennimann et al. (2012), now implemented in the R-package 'ecospat' (Di Cola et al. 2017). It is important to clarify the differences between 'ecospat' and 'humboldt' not discussed in Table 3. The first distinctions regard how both packages incorporate the abundance of environments into niche quantification. Both 'ecospat' and 'humboldt' incorporate the E-space adjustments of Broennimann et al. (2012) by calculating a standardized kernel density that corrects the species' observed densities by the frequency of E-space of input environments. They differ in how they incorporate non-analogous environments and how accessible e-space is defined. The R package 'humboldt' provides a user-friendly way to integrate both factors directly into niche quantifications and associated statistics, whereas 'ecospat' provides no incorporated methods. Certainly, users of 'ecospat' can curate their data to carefully define accessible E-space (i.e., using the methods of Petitpierre et al. 2012 or those in 'humboldt') and remove non-analogous climates in such tests before use. However, the issue of removing the non-analogous climates is non-trivial and likely beyond most users. The R package 'humboldt' provides several methods to determine accessible E-space, as it uses the input localities to calculate either a buffered minimum-convex-polygon, a radial method that buffers each point, or allows users to input their own shapefile. The R package 'humboldt' also provides a power test (humboldt: humboldt.accessible.e.distance) that measures the effects of the input distance parameter associated with buffers used to quantify accessible environments. This function assesses a range of buffer distances and performs NOT and NDT at each buffer distance, assessing how NOT and NDT significance changes as a result of the distance parameter input.

The estimation of the kernel densities is central for quantifying a species' niche and directly dictates how the point occurrence data are converted to a continuous E-space surface. The R package 'ecospat' uses a single method, h_{ref} (Worton 1989), to estimate the kernel smooth density parameter by calculating the standard deviation of rescaled PC1 and PC2 coordinates, divided by the sixth root of the number of locations (Worton 1989, Benhamou and Cornélis, 2010). This method can be unreliable when used on multimodal E-space distributions, as it often results in over-smoothing (Worton 1995, Seaman et al. 1999). Multimodal E-space occupancy can be somewhat common when a species occupies an extreme aspect of habitats (i.e., mountains), where habitats are not equally represented in geography or there are strong imposing biotic interactions (i.e., competitive exclusion), all of which can cause E-space to not be uniformly distributed in E-space dimensions. The independent calculation of h_{ref} for each species in 'ecospat' may also result in an undesirable situation where one species' niche has fine-scale detail, while the other species' niche is coarse. The default setting of the 'humboldt' method implements a fixed smoothing value, h, that is the same for both species (kern.smooth=1). The h value in 'humboldt' can be easily adjusted to allow fine-scale tuning of the area that kernel smoothing occurs across. Larger values (i.e., 2) increase scale, making E-space transitions smoother and typically larger, whereas smaller values (i.e., 0.5) decrease scale, making occupied E-space clusters denser and more irregular (see Appendices 1 and 2).

The Background statistics in 'humboldt' and 'ecospat' differ in their explicit connections between the estimated E-space and the observed G-space. The Background statistic in 'ecospat' shifts the quantified E-space of one species randomly in the two dimensions of E-space. This test then measures the Niche Similarity between the randomly shifted E-space and the original E-space of the other species. Conversely, 'humboldt' shifts the raw occurrence localities of one species randomly in latitude and longitude (in G-space) and then the new distribution is converted to E-space. The shifting of E-space density grids (vs. shifting in G-space) is problematic for several reasons. First, because E-space is shifted, some portion of the original E-space is often shifted off the analysis grid. When this occurs, it reduces the total number of E-space pixels occupied and can increase the likelihood that similarity scores are lower simply due to the number of pixels occupied (vs. the original grid). Second, the random shifting in E-space frequently moves the species' E-space into areas with no corresponding

E-space present in the original G-space. Alternatively, the shifting of E-space can create a situation that does not characterize abundances of the E-space in a way that matches the original G-space. Thus, the shifted E-space can either not be present, or the inferred E-space abundances do not exist in the original environment. Lastly, the shifting of E-space in 'ecospat' (vs. G-space in 'humboldt') does not maintain the nuances of the original dataset in terms of spatial autocorrelation, which can dramatically affect how the spatial distribution patterns are translated into E-space. In contrast, the methods in 'humboldt' that shift geographic space always result in plausible combinations of E-space that could exist in the environment, and species niche density values are always rescaled to sum to 1 (Box 2C). These changes resulted in an increased performance of the Equivalence statistic in 'humboldt' and in a higher proportion of correctly classified niches (Fig. 8).

Caveats of the simulations employed

For future researchers, it is important to clarify why we used the maximum temperatures of the warmest month over the minimum temperatures of the coldest month, which is considered by many a more intuitive indicator of the physiological limits of lowland and montane species (Araújo et al. 2013, Cunningham et al. 2016). In the extent of this study, the two variables were highly correlated (r^2=0.95), making them functionally redundant. However, that is not always the case. The sampling of the occupied altitudes of both species reaffirm their intended montane and lowland distributions (Fig. 6F). In these environments, the species' fundamental niches are not equivalent to their potential niches, as landscapes encompassing all suitable habitats do not exist (Fig. 6F). However, in contrast to most real species, here all habitats possess favorable biotic factors for both species and our measurements are not constrained by biotic interactions.

Understanding low niche similarity values of the simulated species among real landscapes

Our comparisons of the niches of the two simulated species, one in NW South America (SA) and another in the Island of New Guinea (NG), aptly demonstrate the challenges of measuring niche similarity among different landscapes. When comparing the same species' niche similarity between the two regions and using our methods, we observed niche similarity values between 0.469-0.557 (Niche Overlap tests) and not 1 (a similar pattern was observed in the 'ecospat' results). The deviation from 1 is due to the presence of many non-analogous environments and different abundances of key analogous environments between both regions (see Fig. 6G). In contrast, when performing comparisons within the same region, 'humboldt' and 'ecospat' perfectly recovered an overlap value of 1, as expected. Thus, despite possessing identical fundamental niches, the actual potential niches available in each region were quite different between the two regions (see analogous vs. non-analogous environments, Fig. 6G). When comparing the fundamental niches (Fig. 6A) to the E-space available in the landscapes (Fig. 6G), the precipitation E-space appears to be truncated in both SA and NG, and the broad range of precipitation that is habitable is occupied to the edges of what is available. In several instances, our methods failed to recover significant differences between the two species, when simulated in different landscapes, supporting the null hypothesis of niche equivalence. Again, when looking at the distributions of analogous and non-analogous environments between the simulated species in either region (Fig. 6G), as well as the high potential niche truncation values, the failure to detect differences is not surprising. The portion of the niche of simulated species 1 that extends into cooler temperatures (which are distributed in higher elevations in both regions), represents a much smaller combination of climate space compared to broader combinations of climate space in the lowland climates that are shared with simulated species 2 (Fig. 6G). This difference is most pronounced in the comparison between simulated species 1 and species 2 where one species occupies SA and the other NG. In these comparisons, little unique climate space is occupied by simulated species 1 relative to simulated species 2.Therefore, based on the environmental-space available, the occupied niches measured in shared accessible environments (and their potential niche by extension) appear to be, in fact, mostly equivalent.

Testing for niche divergence when there is no shared accessible environmental-space

The NDT is only possible if some portion of the distributions of the two species are in shared environmental-space. If the two species' shared available E-space does not overlap where the two species exist, then there are no analogous climates. In this situation, we recommend reporting that no analogous environments exist and performing only the NOT (Table 2). In such cases, the absence of shared accessible analogous climates is strong evidence of niche divergence, which would be strongly supported by a low niche similarity values and a highly significant NOT.

Assessing niche evolution in shared environmental-space

At the time of writing, most published tests of niche divergence have been performed in each species' total distribution (but see Peterson and Holt 2003, Petitpierre et al. 2012, Qiao et al. 2017). As demonstrated previously (Petitpierre et al. 2012, Qiao et al. 2017) and here, this provides a good metric for total niche similarity between two distributions, but a poor measure of niche evolution in non-equilibrium distributions (Soberón and Nakamura 2009). Species' access to environments can differ due to distinct natural and biogeographic histories, which could have little to do with their underlying fundamental niches (Alexander et al. 2015). Even when two species that share the same ecological tolerances occur in the same geographic area, their life history or distpersal limitations and biological interactions may result in largely different contemporary distributions (e.g., Hardin

1960, Gaulin and Fitzgerald 1988, Garcia-Barros and Benito 2010, Grossenbacher et al. 2015, Estrada et al. 2015, Borda-de-Água et al. 2017). To avoid issues associated with dispersal and life history limitations between taxa, we strongly recommend performing analyses of niche evolution only in shared accessible E-space.

Another important argument for performing analyses of niche divergence in only shared accessible environments is to minimize type I statistical errors. In general, all comparisons of two niches can be classified into three scenarios, where fundamental niches are: [1] identical, [2] different but overlapping, and [3] non-overlapping. In the case of identical fundamental niches, often the NOT will result in the Equivalence statistics being significant (= non-equivalent niches) due to the two species occupying non-analogous climates (Fig. S4). However, if the analysis had occurred only in analogous shared environments (i.e., in the NDT), in most cases (see Fig. 8), the Equivalence statistic would be non-significant, correctly supporting the hypothesis that niches are identical. We also demonstrate that the NDT does not impede detection of divergent, but partically overlapping niches in only analogous environments (Fig. 8).Ssimilarly, it should not affect tests between non-overlapping niches (Supplementary Figs. S5 and S6).

Disentangling biotic interactions, dispersal differences, and niche evolution

Without detailed physiological studies, it may be impossible to conclude whether the fundamental niches of two species are different due to divergence, or if their fundamental niches are identical but their potential niches are different as a result of imposing biotic factors. The latter scenario is less likely in shared environments among similar species (i.e., sister taxa), unless one or both focal species themselves are the main causes of unfavorable biotic factors. As geographic distance increases between the ranges of the two species, shared biotic factors are expected to decrease, increasing the likelihood that biotic factors play different roles in the distribution of the two taxa. For example, biotic factors limiting species' distributions are presumed to be more constraining in a species' native distribution vs. its invasive range. Because the co-evolutionary dynamics of species interactions may be more tightly linked in the native range of a species, it should be expected that its potential niche in the invasive range will be larger than in its native range.

These realizations mean that, for most of the taxa we study, exhaustive proof of niche divergence via casual correlations is not possible. However, it does mean that we can place a high degree of confidence in the assessment of niche equivalency. When NDTs are non-significant (fail to reject the null hypothesis of equivalence), researchers have strong evidence that the portion of the fundamental niche in shared analogous environments is similar, regardless of differences in favorable or unfavorable biotic factors. Though this may be disparaging for many, significant NDTs do provide considerable insight into niche evolution, particularly when researchers strongly consider the relationships between a species' occupied niche, potential niche, and fundamental niche in the context of the BAM diagram. Further, for many species, there may not exist apparent biological factors or dispersal limitations driving differences in distributions, leaving differences in a species' fundamental niche as the most probable cause of differences.

Regarding Alexander von Humboldt

The new methods presented here are implemented in an R package named after Alexander von Humboldt (1769-1859), who is widely recognized for his work on botanical geography, and whose drive to understand nature as a whole laid the foundations for the fields of biogeography and ecology (Nocolson 1987, Wulf 2015, Schrodt et al. 2019). However, to us, Humboldt's greatest legacy regards his ideas about the explicit interconnectedness of the world, which were derived from firsthand experiences filtered through a strong quantitative and scientific perspective (Keppel and Kreft 2019).

Acknowledgements

We are grateful to Morgan R. Muell, Wilson X. Guillory, Connor M. French, Robert Boria, Jaime Kass, Robert P. Anderson, Gonzalo Pinilla-Buitrago, Kathryn Mercier, Layla Hernandez, Diego Serrano-Alvarado, Andrew Townsend Peterson, and Marlon E. Cobos for insightful comments and constructive criticism over the long development and implementation of this project. A. Carnaval acknowledges support by the National Science Foundation (DEB 1120487, 1343578), FAPESP (BIOTA 2013/50297-0) and NASA.

Author contributions: J.L.B. and A.C.C. conceived the ideas and wrote the manuscript; J.L.B. wrote the software, performed analyses, and created figures.

Supplementary Materials

The following materials are available as part of the online article from https://escholarship.org/uc/fb

Table S1. Key 'humboldt' parameters for Niche Overlap and Niche Divergence tests.

Table S2. Niche overlap and divergence of simulated species: control self-comparison. Climate: only Bio5 and Bio12.

Table S3. Niche overlap and divergence of simulated species, control self-comparison. Climate: all bioclims.

Table S4. Niche overlap and divergence of simulated species. Climate: only Bio5 and Bio12.

Figure S1. Quantifying shared accessible E-space in 'humboldt'.

Figure S2. Interpreting shared accessible E-space projected into G-space in 'humboldt'.

Figure S3. Interpreting shared accessible E-space in 'humboldt'. E-space

Figure S4. Hypothetical distributions of environmental-space of two species with identical

niches and outcomes of the Niche Overlap and Niche Divergence tests

Figure S5. Hypothetical distributions of environmental-space of two species with different overlapping niches and outcomes of the Niche Overlap and Niche Divergence tests

Figure S6. Hypothetical distributions of environmental-space of two species with different non-overlapping niches and outcomes of the Niche Overlap and Niche Divergence tests.

Appendix S1. An overview of syntax in 'humboldt'.

Appendix S2. A visual guide to parameter options in 'humboldt'

Appendix S3. A visual guide and interpretation of results output from 'humboldt'

References

Alexander, P.R., Costa, G.C., Patton, M. A. & Burbrink F. (2015) Phylogenetic niche conservatism and the evolutionary basis of ecological speciation. Biological Reviews, 90, 1248-1262.

Araújo, M.B. & Pearson, R.G. (2005) Equilibrium of species' distributions with climate. Ecography, 28.5, 693-695.

Araújo, M.B., Ferri-Yáñez, F., Bozinovic, F., Marquet, P.A., Valladares, F. & Chown, S.L. (2013) Heat freezes niche evolution. Ecology Letters, 16, 1206-1219.

Austin, M.P. & Van Niel, K.P. (2011) Improving species distribution models for climate change studies: variable selection and scale. Journal of Biogeography, 38, 1-8.

Beale, C. M., Lennon, J. J. & Gimona, A. (2008) Opening the climate envelope reveals no macroscale associations with climate in European birds. Proceedings of the National Academy of Sciences USA, 105, 14908-14912.

Benhamou, S. & Cornélis, D. (2010). Incorporating movement behavior and barriers to improve kernel home range space use estimates. The Journal of Wildlife Management, 74, 1353-1360.

Bonetti, M.F. & Wiens, J.J. (2014) Evolution of climatic niche specialization: a phylogenetic analysis in amphibians. Proceedings of the Royal Society of London B, 281, 20133229.

Borda-de-Água, L., Whittaker, R.J., Cardoso, P., Rigal, F., Santos, A.M., Amorim, I.R. & Borges, P.A. (2017). Dispersal ability determines the scaling properties of species abundance distributions: a case study using arthropods from the Azores. Scientific Reports, 7, 3899.

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. (2012) Measuring ecological niche overlap from occurrence and spatial environmental data. Global Ecology and Biogeography, 21, 481-497.

Brown, J.L., Bennett, J.R. & French, C.M. (2017) SDMtoolbox 2.0: the next generation Python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. PeerJ, 5, e4095.

Calatayud, J., Rodriguez, M. A., Molina-Venegas, R., Leo, M., Horreo, J. L. & Hortal, J. (2019). Pleistocene climate change and the formation of regional species pools. Proceedings of the Royal Society of London B, 286, 20190291.

Cheng, H., Sinha, A., Cruz, F.W., Wang, X., Edwards, R.L., d'Horta, F.M., Ribas, C.C., Vuille, M., Stott, L.D. & Auler, A.S. (2013) Climate change patterns in Amazonia and biodiversity. Nature Communications, 4, 1411.

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.

Cunningham, H.R., Rissler, L.J., Buckley, L.B. & Urban, M.C. (2016) Abiotic and biotic constraints across reptile and amphibian ranges. Ecography, 39, 1-8.

Di Cola, V., Broennimann, O., Petitpierre, B., Breiner, F.T., D'amen, M., Randin, C., Engler, R., Pottier, J., Pio, D. & Dubuis, A. (2017) ecospat: an R package to support spatial analyses and modeling of species niches and distributions. Ecography, 40, 774-787.

Elton, C. (1927) Animal ecology. Sidgwick and Jackson, Ltd, London.

Estrada, A., Meireles, C., Morales-Castilla, I., Poschlod, P., Vieites, D., Araújo, M.B., et al. (2015). Species' intrinsic traits inform their range limitations and vulnerability under environmental change. Global Ecology and Biogeography, 24, 849–858

Fick, S.E. & Hijmans, R.J. (2017) WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. International Journal of Climatology, 37, 4302-4315.

Galbreath, K.E., Hafner, D.J. & Zamudio, K.R. (2009) When cold is better: climate-driven elevation shifts yield complex patterns of diversification and demography in an alpine specialist (American pika, Ochotona princeps). Evolution, 63, 2848-2863.

Gaulin, S. J. & FitzGerald, R. W. (1988). Home-range size as a predictor of mating systems in Microtus. Journal of Mammalogy, 69, 311-319.

Garcia-Barros, E. & Romo Benito, H. (2010). The relationship between geographic range size and life history traits: is biogeographic history uncovered? A test using the Iberian butterflies. Ecography, 33, 392-401.

Gaston K.J. (2003) The Structure and Dynamics of Geographic Ranges. Oxford University Press, Oxford, UK.

GBIF (2017) *Conium maculatum* in GBIF Secretariat. GBIF Backbone Taxonomy. Checklist dataset https://doi.org/10.15468/39omei, accessed via GBIF.org on 2017-06-07.

Grinnell, J. (1914) Barriers to distribution as regards birds and mammals. The American Naturalist, 48, 248-254.

Grinnell, J. (1917) Field tests of theories concerning distributional control. The American Naturalist, 51, 115-128

Grossenbacher, D., Briscoe Runquist, R., Goldberg, E. E. & Brandvain, Y. (2015). Geographic range size is predicted by plant mating system. Ecology letters, 18, 706-713.

Guisan, A., Petitpierre, B., Broennimann, O., Daehler, C. & Kueffer, C. (2014) Unifying niche shift studies: insights from biological invasions. Trends in Ecology & Evolution, 29, 260-269.

Hardin, G. (1960). The competitive exclusion principle. Science, 131, 1292-1297.

Hellinger, E. (1909) Neue begründung der theorie quadratischer formen von unendlichvielen veränderlichen. Journal für die reine und angewandte Mathematik, 136, 210-271.

Hortal, J., Jiménez-Valverde, A., Gómez, J. F., Lobo, J. M. & Baselga, A. (2008). Historical bias in biodiversity inventories affects the observed environmental niche of the species. Oikos, 117, 847-858.

Hutchinson, M., Xu, T., Houlder, D., Nix, H. & McMahon, J. (2009) ANUCLIM 6.0 user's guide. Australian National University, Fenner School of Environment and Society, Canberra.

Jackson, S.T. & Overpeck, J.T. (2000) Responses of plant populations and communities to environmental changes of the late Quaternary. Paleobiology, 26, 194-220.

James, F.C., Johnston, R.F., Wamer, N.O., Niemi, G.J. & Boecklen, W.J. (1984) The Grinnellian niche of the wood thrush. The American Naturalist, 124, 17-47.

Keppel, G. & Kreft, H. (2019) Integration and synthesis of quantitative data: Alexander von Humboldt's renewed relevance in modern biogeography and ecology. Frontiers of Biogeography, 11, e43187.

Kozak, K.H. & Wiens, J.J. (2006) Does niche conservatism promote speciation? A case study in North American salamanders. Evolution, 60, 2604-2621.

Leroy, B., Meynard, C.N., Bellard, C. & Courchamp, F. (2016) virtualspecies, an R package to generate simulated species distributions. Ecography, 39, 599-607.

Lomolino, M.V., Riddle, B.R. & Whittaker, R.J. (2017) Biogeography: Biological diversity across space and time. Fifth edition. Sinauer Associates, Sunderland.

Losos, J.B. (2008) Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. Ecology Letters, 11, 995-1003.

McCormack, J.E., Zellmer, A.J. & Knowles, L.L. (2010) Does niche divergence accompany allopatric divergence in *Aphelocoma* jays as predicted under ecological speciation? Insights from tests with niche models, Evolution, 64, 1231-1244.

Nicolson, M. (1987) Alexander von Humboldt, Humboldtian Science and the origins of the study of vegetation. History of Science, 25, 167-194.

Nunes, L.A. & Pearson, R.G. (2017) A null biogeographical test for assessing ecological niche evolution, Journal of Biogeography, 44, 1331-1343.

Pearman, P.B., Guisan, A., Broennimann, O. & Randin, C.F. (2008) Niche dynamics in space and time. Trends in Ecology & Evolution, 23, 149-158.

Petitpierre, B., Kueffer, C., Broennimann, O., Randin, C., Daehler, C. & Guisan, A. (2012) Climatic niche shifts are rare among terrestrial plant invaders. Science, 335, 1344-1347.

Pellissier, L., Bråthen, K.A., Vittoz, P., Yoccoz, N.G., Dubuis, A., Meier, E.S., Zimmermann, N.E., Randin, C.F., Thuiller, W. & Garraud, L. (2013) Thermal niches are more conserved at cold than warm limits in arctic-alpine plant species. Global Ecology and Biogeography, 22, 933-941.

Peterson, A. T. & Holt, R. D. (2003) Niche differentiation in Mexican birds: Using point occurrences to

detect ecological innovation. Ecology Letters, 6, 774-782

Peterson, A., Soberón, J. & Sánchez-Cordero, V. (1999) Conservatism of ecological niches in evolutionary time. Science, 285, 1265-1267.

Peterson, A.T. & Soberón, J. (2012) Species distribution modeling and ecological niche modeling: getting the concepts right. Natureza & Conservação, 10, 102-107.

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.

Qiao, H., Escobar, L.E. & Peterson, A.T. (2017) Accessible areas in ecological niche comparisons of invasive species: Recognized but still overlooked. Scientific Reports, 7, 1213.

Qiao, H., Peterson, A.T., Campbell, Li., Soberón, J., Ji, L. & Escobar, L.E. (2015) NicheA: Creating simulated species and ecological niches in multivariate environmental scenarios, Ecography.

Rödder, D. & Engler J.O. (2011) Quantitative metrics of overlaps in Grinnellian niches: advances and possible drawbacks. Global Ecology and Biogeography, 20, 915-927.

Rödder, D. & Stefan, L. (2009) Niche shift versus niche conservatism? Climatic characteristics of the native and invasive ranges of the Mediterranean house gecko (*Hemidactylus turcicus*). Global Ecology and Biogeography, 18, 674-687.

Saupe, E. E., Barve, N., Owens, H. L., Cooper, J. C., Hosner, P. A. & Peterson, A. T. (2017). Reconstructing ecological niche evolution when niches are incompletely characterized. Systematic biology, 67, 428-438.

Schoener, T.W. (1968) The Anolis lizards of Bimini: Resource partitioning in a complex fauna. Ecology, 49, 704-726.

Schoener, T.W. & Gorman, G.C. (1968) Some niche differences in three Lesser Antillean lizards of the genus Anolis. Ecology, 49, 819-830.

Schrodt, F., Bailey, J.J., Santos, M.J. & Field, R. (2019) Challenges and opportunities for biogeography - what can we still learn from von Humboldt? Journal of Biogeography, doi:10.1111/jbi.13616.

Seaman, D.E., Millspaugh, J.J., Kernohan, B.J., Brundige, G.C., Raedeke, K.J. & Gitzen, R.A. (1999) Effects of sample size on kernel home range estimates. The Journal of Wildlife Management, 63, 739-747.

Soberón, J. (2007) Grinnellian and Eltonian niches and geographic distributions of species. Ecology Letters, 10, 1115-1123.

Soberón, J. & Nakamura, M. (2009) Niches and distributional areas: Concepts, methods, and assumptions. Proceedings of the National Academy of Sciences USA, 106, 19644-19650.

Vetter, J. (2004) Poison hemlock (Conium maculatum L.). Food and Chemical Toxicology, 42, 1373-1382.

Warren, D.L., Glor, R.E. & Turelli, M. (2008) Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. Evolution, 62, 2868-2883.

Warren, D.L., Glor, R.E. & Turelli, M. (2010) ENMTools: a toolbox for comparative studies of environmental niche models. Ecography, 33, 607-611.

Wiens, J.J. & Graham, C.H. (2005) Niche conservatism: integrating evolution, ecology, and conservation biology. Annual Review of Ecology, Evolution and Systematics, 36, 519-539.

Worton, B.J. (1995) Using Monte Carlo simulation to evaluate kernel-based home range estimators. The Journal of Wildlife Management, 59, 794-800.

Wulf, A. (2015). *The invention of nature – Alexander von Humboldt's New World.* Alfred Knoopf, New York.

Submitted: 17 June 2019
First decision: 19 July 2019
Accepted: 12 August 2019

Edited by Alycia Stigall, Michael N Dawson, Joaquín Hortal, Cascade Sorte, Mark Lomolino and Robert J. Whittaker

Glossary of Terms

Glossary	Definition
Accessible Environmental-space	The E-space estimated to be available to the species based on its contemporary distribution and access to adjacent climates.
Background Statistic	A spatial statistic used in the Background, Niche Overlap, and Niche Divergence tests. It asks if the two distributed species are more different than would be expected given the underlying environmental differences between the landscapes in which they occur.
Background Test	A statistical test based on the Background statistic. If it is non-significant and an Equivalence statistic is non-significant, there is limited power for the Equivalence statistic to actually detect significant differences among taxa.
Buffer	A buffer is an area defined by the bounding region determined by a set of points or a polygon. Users input a maximum distance from all segments of an object (typically a minimum convex polygon of all occurrence localities for a species or each individual locality).
Corrected E-space	A scenario where species' niches reflect a standardized kernel density that corrects the species' observed E-space densities by the frequency of E-space in the corresponding environment.
Environmental-space or E-space	Multidimensional spaces of environmental variables that exist in a given region at a given time. Here we visualize and analyze E-space in two dimensions, usually characterized by the first two Principal Components from a Principal Component Analysis.
Geographic space or G-space	The two dimensional (latitude and longitude) geographic distribution of species, populations, and environmental variables that exist in a given region at a given time.
Niche Divergence Test (NDT)	A statistical test using the Niche Equivalence and Background statistics where the E-space of both species are reduced to areas of analogous environments. If significant, the two occupied niches are statistically different and likely the result of divergence. Also see complimentary Niche Overlap test.
Niche E-space Correlation Index (NECI)	If key habitats associated with different environmental-space are not equally represented, biases can also occur towards the most abundant suitable habitats. This can lead to the appearance that two taxa occupy different environmental-space, whereas it actually is only an artefact of the differential abundance of habitats between the two distributions. The Niche E-space Correlation Index (NECI)determines if the species occupied environmental-space should be standardized by the abundance of environmental-space throughout the species' accessible environments. If the NECI is high (e.g., > 0,5), species occupied niches should be corrected by the frequency of E-space in accessible environments to reduce type I errors.
Niche Equivalence statistic	A resampling statistic that compares the observed niche similarity values to resampling the observed dataset. It is used in the Equivalence, Niche Overlap, Niche Divergence tests. Occurrence localities are pooled and resampled in two groups equaling the number of localities in the original dataset. Niche similarity is then assessed and compared to the niche similarity of the observed data.
Niche Equivalence test	A statistical test of the null hypothesis that two occupied niches are equal. If significant, the two niches are statistically different.
Niche Identity Test	The statistical test of the null hypothesis that occupied niches are equivalent from Warren et al. (2010). This test is synonymous with the Niche Equivalence test. If significant, the two occupied niches are statistically different.
Niche Overlap Test (NOT)	A statistical test using the Niche Equivalence and Background statistics where the full accessible E-space of both species is incorporated. If significant, the two occupied niches are statistically different in their total distribution. Note, this test cannot state much about species' niche divergence. Also see complimentary Niche Divergence test.

Glossary of Terms

Glossary	Definition
Niche Similarity Indices, Metrics, or Measurements	A numeric value from 0-1 reflecting the niche similarity among two occupied niches. A value of 1 equals identical niches and a value 0 is completely different niches. Two common niche measurements are Warren's I & Schoener's D.
Potential Niche Truncation Index (PNTI)	This value describes the amount of observed E-space of the species that is truncated by available E-space. The value is the portion of the species' niche perimeter that falls outside of the perimeter of environmental-space. This value physically measures how much of the perimeter of the species' E-space abuts or overlaps with the margins of the environment's E-space. If this index is moderate or high, the occupied niche likely does not adequately reflect the species' fundamental niche. See translation of empirical of values to levels of truncation within software.
Schoener's D	The niche similarity measurement *D* of Schoener (Schoener 1968). Equals one minus the total variation distance between two occupied niches.
Similarity Test	A statistical test to assess to power to detect significant differences in the Equivalence test from Warren et al. (2008) and Broenniman et al. (2012). This test is synonymous with the Background test.
Uncorrected E-space	A scenario where a species' niches reflect the raw kernel density quantification of environmental-space occupied (the abundance of E-space in the environment is not accounted for).
Warren's I	The niche similarity measurement *I* of Warren et al. (2008). Equals one minus Hellinger's distance (measured between two occupied niches).

