Biodiversity Informatics, 17, 2022, pp. 50-58

DETECTING SIGNALS OF SPECIES' ECOLOGICAL NICHES IN RESULTS OF STUDIES WITH DEFINED SAMPLING PROTOCOLS: EXAMPLE APPLICATION TO PATHOGEN NICHES

Marlon E. Cobos*, A. Townsend Peterson

Department of Ecology and Evolutionary Biology & Biodiversity Institute, University of Kansas, Lawrence, Kansas 66045, USA

**Corresponding author: Marlon E. Cobos, Email: manubio13@gmail.com*

Abstract. Ecological niches are increasingly appreciated as a long-term stable constraint on the geographic and temporal distributions of species, including species involved in disease transmission cycles (pathogens, vectors, hosts). Although considerable research effort has used correlative methodologies for characterizing niches, sampling effort (and the biases that this effort may or may not carry with it) considerations have generally not been incorporated explicitly into ecological niche modeling. In some cases, however, the sampling effort can be characterized explicitly, such as when hosts are tested for pathogens, as well as comparable situations such as when traps are deployed to capture particular species, etc. Here, we present simple methods for testing the hypothesis that non-randomness in occurrence or detection exists with respect to environmental dimensions (= a detectable signal of ecological niche); i.e., whether a pathogen occurs nonrandomly with respect to environment, given the occurrence and sampling of its host. We have implemented a set of R functions that presents an overall test for nonrandom occurrence with respect to a set of environmental dimensions, and, a posteriori, a set of exploratory tests that identify in which dimension(s) and in which direction or form the nonrandom occurrence is manifested. Our tools correctly detected signals of niche in most of our example cases. Although such a signal may not be detectable in cases in which the niche of interest is broader than the universe sampled, such a possibility was correctly discarded in our analyses, preventing further interpretations. This kind of testing can constitute an initial step in a process that would conclude with development of a more typical ecological niche model. The particular advantage of the analyses proposed is that they consider the biases involved in sampling, testing, and reporting, in the context of nonrandom occurrence with respect to environment before proceeding to inferential and predictive steps.

Key words: ecological niche, host, niche position, niche breadth, non-parametric test, PERMANOVA

The ideas, tools, and methods used under the rubrics of "ecological niche modeling" and "species distribution modeling" (here referred to as ENM/SDM), have seen extensive application to understanding the geographic and environmental distributions of species (Franklin 2010; Peterson et al. 2011). Most popular have been correlative approaches, in which environmental characteristics of places of known occurrences of species are subjected to a variety of model-fitting approaches, to create a classification of different parts of environmental space into suitable and unsuitable sets of conditions (Peterson et al. 2011; Enriquez-Urzelai et al. 2019). A major challenge for these methods, however, has been the pervasive biases and gaps that characterize the sampling that produced the primary occurrence data, and how to avoid propagation of those biases through the analytical sequence to the results (Anderson and Gonzalez 2011; Acevedo et al. 2012; Araújo et al. 2019).

Some primary biodiversity occurrence data, however, may be connected to information that can characterize the sampling universe integrally. Such data may take the form of occurrences of pathogens detected by testing hosts (e.g., Eisen and Paddock 2021), disease case data that come

from active surveillance (e.g., M'ikanatha et al. 2008), biodiversity data that are accumulated by trapping where trap data are recorded (Meek et al. 2015), and biodiversity data that are accumulated by standardized sampling protocols (Manley et al. 2005). In each case, the geographic and temporal distribution of sampling can be characterized precisely, and all positive records of the species of interest must necessarily derive from one of the sampling events. This additional information offers considerable promise in informing the modeling process precisely about the sampling universe, rather than relying on assumptions of random sampling (Phillips et al. 2009) or an interpolated sampling bias surface (Warren et al. 2014).

To our knowledge, exploring signals of ecological niche differentiation in data for which the sampling universe is known has not been done before. Most studies use traditional approaches to characterize ecological niches of species and compare such niches without explicit consideration of the sampling universe. In this contribution, we present a logic for a suite of analyses designed to take advantage of this additional information (the sampling universe) available for occurrence data that come from such controlled sampling schemes. We provide a methodological protocol that first tests for any overall niche difference, and then characterizes these differences in terms of a spectrum of possible changes in niches in each environmental dimension. We present the protocol in the form of a set of R functions, to facilitate wide use and incorporation in many other analyses.

Protocol Description

We offer two complementary approaches to detect signals of niche: (1) a multivariate analysis based on a permutational multivariate analysis of variance (PERMANOVA; (Anderson 2017), and (2) a univariate non-parametric method based on descriptive statistics. To illustrate the utility of this approach, we use a suite of virtual species. For each, we created a sample of records that represent the universe of sampling (e.g., a host species that is to be tested for a particular pathogen), and then identify a subgroup of those records that may or may not be positive for the pathogen (see Example application).

That is, the data required to perform these analyses consist of a set of records representing the sampling universe, to which a test is applied that determines presence or absence of the species of interest (0 = negative and 1 = positive). Each of these records carries with it a vector of relevant environmental conditions (see example in Table S1). A typical such situation would be sampling a host species and testing for presence of a pathogen in each host, but many parallel applications exist. Each host record has a geographic reference and potentially also information about collection time—this place and time information can be used to extract environmental data that is place-specific or place-and-time-specific from diverse raster data layers (e.g., data on climate, remote-sensing information, etc.) that are relevant in niche characterization (Ingenloff and Peterson 2021).

Multivariate test

As a multivariate test to detect overall signals of niche, we propose an approach using a PERMANOVA. PERMANOVA is a non-parametric multivariate test that allows comparison of samples by testing a null hypothesis (H_0) that the position and dispersion of the sample are equivalent to those of the sampling universe. Rejecting H_0 indicates that either the centroid (position) or spread (dispersion) is different, which would be indicative of a niche in the pathogen distinct from that of the host. Similarity among groups is tested based on distances (e.g., Euclidean or Mahalanobis distances).

In this application, the groups to be compared are records of the host of which a few are infected (i.e., all host records vs records of infected hosts). We chose to base our PERMANOVA analyses on Mahalanobis distances, but other methods to calculate dissimilarities can be used to perform these processes (e.g., Euclidean distances, Bray-Curtis, and Jaccard indices; see code documentation). The PERMANOVA yields a result of significant or not, indicating rejection or acceptance of the null hypothesis of equivalency, but does not characterize the form of those differences. For this reason, the univariate tests described below are used to provide additional information about the form of these differences.

Univariate test

To characterize niches in individual environmental dimensions, we use a comparison of observed values of descriptive statistics summarizing characteristics of distributions of environmental conditions associated with known-infected hosts

against a null distribution of those statistics derived from many similar-sized random samples drawn from the set of all host records. The descriptive statistics explored and used in this approach are the mean, median, standard deviation (SD), and range. The mean and median of the environmental values are used as estimators of niche position; the SD and the range are descriptors of the spread of environmental conditions comprising the niche.

A null distribution of statistics is derived from many random samples (of size matching the number of positive tests) drawn from the set of all host records; this distribution informs about how common certain values of the descriptive statistics would be if the pathogen had no particular preference or bias from among the set of environmental conditions used by the host. Therefore, the null hypothesis (H_0) for this test is that the descriptive statistic calculated for the samples in which the pathogen was detected cannot be distinguished from the comparable statistic for the host (or the sampling universe). This H_0 is tested for the mean, median, SD, and range, as different measures of characteristics of the distribution.

The following is a sequential description of steps involved in running this analysis:

1. The number of infected host records is calculated (n_i).
2. The mean, median, SD, and range of environmental conditions for the infected hosts are calculated.
3. A random sample of n_i records is drawn from the entire set of host records.
4. The statistics of interest of environmental conditions are calculated for the sample in step 3.
5. Steps 3 and 4 are repeated n_t times (iterations; generally $n_t = 1000$).
6. The full distributions of the statistics of interest are compiled and characterized, particularly as regards the 2.5% and 97.5% levels of the distribution.
7. The observed value of the statistic of interest for infected hosts (step 2) is compared against the null distribution of values (step 6) to establish whether it falls in the central 95% of the null distribution.
8. Depending on the results from step 7, the statistic of interest for the niche of the pathogen is categorized as different or not from null expectations.
9. The direction of the difference is characterized by direct inspection to establish whether the pathogen's niche is shifted upward or downward in the values of the particular environmental dimension (mean or median), or whether it has broadened or narrowed (SD or range).

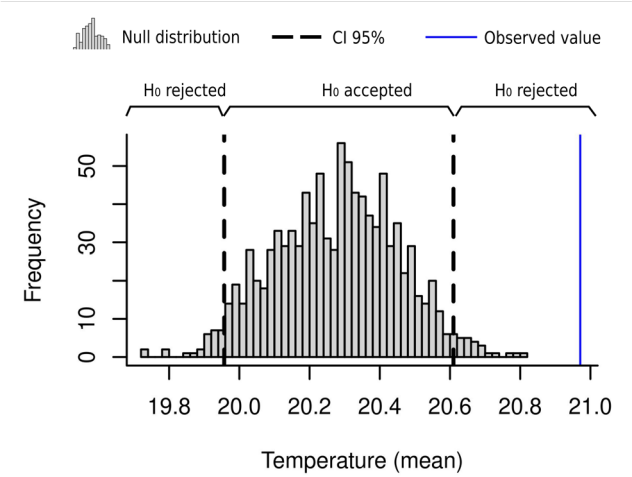

Figure 1. Representation of outcome and suggested interpretation of results from the univariate non-parametric test to detect signals of niche dissimilarity. In this case, we present mean temperature responses for Scenario 5; the null hypothesis would be rejected, in favor of an alternative hypothesis of higher-than-null mean temperature.

The results obtained from these steps allow us to accept or reject H_0 (Fig. 1). To reject H_0, the value of the statistic observed for the positive records must be as extreme or more extreme than the 2.5% or 97.5% of the null distribution. We conclude that the pathogen niche (in terms of the statistic under test) is not distinct if H_0 cannot be rejected. When H_0 is rejected, the statistic under test can be lower or higher depending on in which tail of the null distribution the observed value falls.

Software

We created a set of R functions to run the analyses described above. To aid interpretation, we also created functions to plot results from analyses. These functions are open-source tools that can be accessed following indications in Software Availability. Proper documentation describing the data required to run analyses and how parameter values can be established is provided with the R scripts.

Example Application

Example data

To explore and test the performance of the protocols described above, we generated virtual niches for a host and seven simulated pathogens (Figs. 2, S1) representing a distinct scenario of similarity of host and pathogen niches (Table 1). One case (Scenario 1) was designed to have a pathogen

Table 1. General description of parameters that define the host and pathogen virtual niches according to distinct scenarios of niche similarity. Values of resulting pathogen prevalence in the host are also shown in the table. T = temperature; P = precipitation.

Host / pathogen	Scenario description	Temperature range (°C)	Precipitation range (mm)	Covariance	Pathogen prevalence
Host	-	12–26	700–2800	T: 5.44 P: 122,500	-
Pathogen 1	Pathogen niche is equal to host niche	12–26	700–2800	T: 5.44 P: 122,500	0.52
Pathogen 2	Pathogen niche has the same size as host niche but with changed position	14–28	800–2900	T: 5.44 P: 122,500	0.38
Pathogen 3	Pathogen niche changed in position and size (smaller) compared to host niche	13–22	900–2600	T: 2.25 P: 8,277	0.61
Pathogen 4	Pathogen niche smaller than host niche	15–23	1000–2500	T: 1.78 P: 62,500	0.68
Pathogen 5	Pathogen niche smaller and changed in position	18–25	1000–4000	T: 1.36 P: 250,000	0.47
Pathogen 6	Pathogen niche larger than host niche, but overlaps most of it	14–29	600–4500	T: 6.25 P: 422,500	0.29
Pathogen 7	Pathogen niche larger than host niche but contains it completely	8–30	200–3200	T: 13.44 P: 250,000	0.30

with exactly the same niche as the host, in the other scenarios, the host and pathogen niches overlap in different ways (see Figs. 2, S1). Virtual niches were generated in R 4.1.1 (R Core Team 2021) using the package "evniche"[1], which uses ellipsoids to create niches based on user-defined limits (variable ranges) and covariance values.

We considered annual mean temperature and annual precipitation as the dimensions of our virtual niches. To make our simulations more realistic, when generating data from virtual niches, we considered suitability values derived from Mahalanobis distances (based on multivariate normal distributions) to the centroids of the ellipsoid-shaped niches, measured from points present in available environmental conditions in a region (for details see Etherington 2019; Nuñez-Penichet et al. 2021). Using ellipsoids and the multivariate normal transformations generates responses that are simple, symmetrical, and convex, which we consider appropriate to represent virtual fundamental niches; however, we emphasize that our methods are general, and do not depend on assumptions of normality, regardless of whether our example application makes such assumptions. As a result, the density of records generated from virtual niches increases towards the centroid of the ellipsoids, but it will also depend on the density of points representing available conditions across the area of analysis. Available environmental conditions were represented by values of annual mean temperature and annual precipitation present across South America. We used two of the so-called "bioclimatic" data layers from the WorldClim

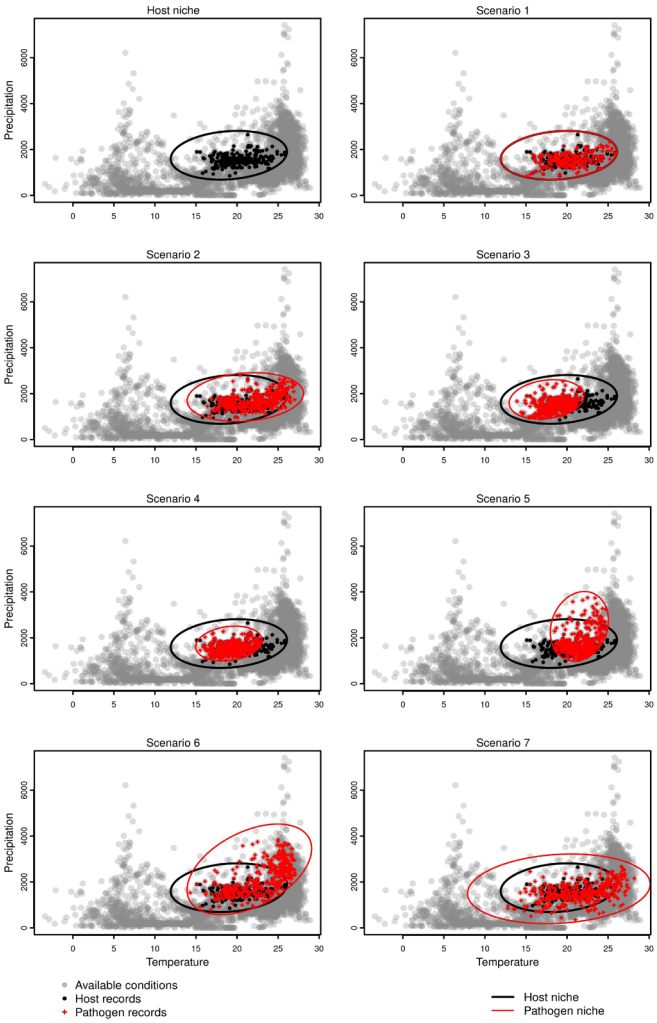

Figure 2. Virtual niches (ellipses) of a host and 7 pathogen scenarios used in the example application. Points in black and red represent records of host and pathogen, respectively, as if they were obtained from geographic records. Host and pathogen records were derived from ellipses, and are overlaid on environmental conditions across South America (gray points).

[1]https://github.com/marlonecobos/evniche.

database v1.4 (Hijmans et al. 2005) and masked them to South America to perform the analyses described above. Raster processing was done using the package "raster" (Hijmans 2019).

We generated populations of points via sampling from the centrality-weighted ellipsoids for the host and the pathogen niches separately, and then used as "pathogen-positive" records those host-niche points that coincided exactly with pathogen-niche points; for this purpose, we generated 200 host-niche points and 400 pathogen-niche points. Because in Scenario 1, host and pathogen niches were exactly the same, we simply subsampled 200 points from among the pathogen 400 points to exclude some of the host records from being considered as infected. Because records derived from ellipsoids with distinct sizes and positions in the cloud of available environmental conditions for host and pathogens, we were not able to control pathogen prevalences (see Table 1).

Niche comparisons

We compared host and pathogen ecological niches considering the 7 pathogen-niche scenarios using both the multivariate and univariate approaches. Multivariate comparisons were made using PERMANOVA analyses with 1000 iterations for calculation of statistical significance. For univariate comparisons, the mean, standard deviation, and range of values corresponding to infected hosts were compared to the distribution of the same statistics for 1000 random samples from the host records.

To aid with interpretation, we created ellipsoids for the environmental distributions of the host and all pathogens. For pathogens, we considered the data used in analyses (i.e, not all records generated using virtual niches of pathogens, but rather only those of pathogens that match the host). We plotted all ellipsoids derived from the data to explore and visualize the position and spread of host and pathogen niches.

Results

Final datasets prepared for analysis consisted of 200 records of the host, of which 57-136 matched virtual pathogen records, and thus were considered as infected hosts (see Table 1 for pathogen prevalences; see example dataset in Table S1). Environmental representations of datasets showed distinct levels of overlap between host records and infected ones, which helped us to understand the actual configurations of host and pathogen records that can be observed in real applications (Fig. 3).

No signal of a distinct pathogen niche was detected in 3 of the 7 scenarios using the PERMANOVA (scenarios 1, 6, and 7; Fig. 4). That is, based on the multivariate analysis, the centroid and dispersion of infected hosts can be considered as non-distinguishable from those of all hosts for scenarios 1, 6, and 7. Univariate analyses further indicated that host and pathogen in scenarios 1 and 7 were not distinct in any individual dimension. For all other scenarios, some signal of dissimilarity was detected (Table 2). For these cases, the observed mean, median, standard deviation, or range derived from individual variable values of infected hosts fell outside of the central 95% of the null distribution of values derived from 1000 random samples of all hosts for one or both of the environmental dimensions (Figs. S2-S5). Considering the qualities of the 7 scenarios, both methods could not reject H_0 when niches were exactly equal (Scenario 1), or when the pathogen niche contained completely that of the host (Scenario 7). However, the PERMANOVA did not detect niche dissimilarity for Scenario 6, even though the original host and pathogen niches were different. Both methods identified a signal of dissimilarity for scenario 2, yet the niche of the pathogen had only a slight change in position from that of the host.

Discussion

Both multivariate and univariate approaches performed well in detecting signals of niche dissimilarity in cases in which the pathogen niche represented a subgroup of that of the host (or the sampling universe). The two tests are complementary in the sense that they are based on different procedures and ideas, but both help to detect signals and interpret the type of signal detected. The PERMANOVA-based test seeks an overall signal of niche difference, and also considers covariation among variables, although a direct understanding of the sort of difference manifested does not derive directly from this test. The univariate analyses, in contrast, allow one to understand changes in niche position and breadth when an overall signal is detected, although it does not consider covariation among multiple environmental variables. Graphical representations of results help considerably with interpretation and complement further the understanding of the signals detected.

Apart from the obvious differences between the univariate and multivariate analyses, another

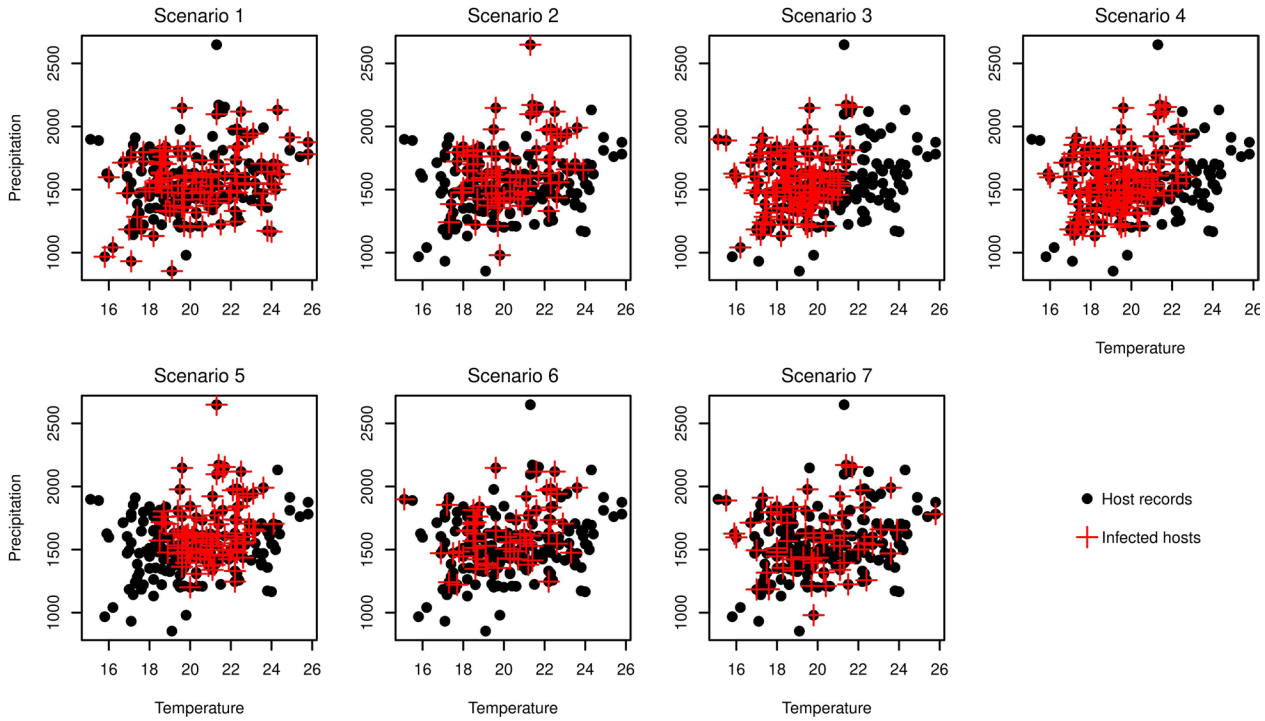

Figure 3. Visualization of data derived from virtual niches representing hosts infected under 7 pathogen-niche scenarios. This view shows how data would look if derived from sampling the host and testing for pathogens, with no previous knowledge of host or pathogen niches. Niche differences between host and pathogen can be noted as conditions under which hosts are not marked as infected, such as above 22°C in Scenario 3.

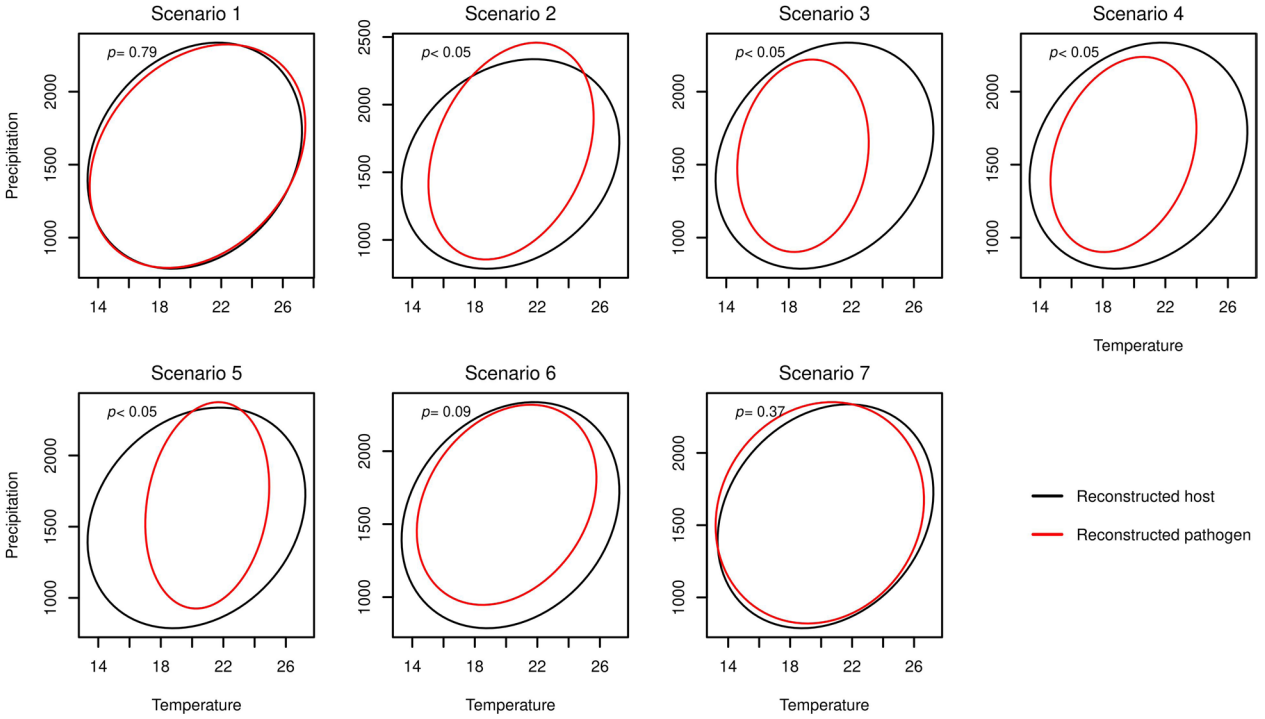

Figure 4. Results from niche comparisons using PERMANOVA analyses. Ellipses were reconstructed from the data created from virtual niches and the available background. Values of statistical significance are shown for each comparison.

Table 2. Summary of results derived from univariate niche comparisons. Comparisons identified as higher or lower were statistically significant (i.e., observed values from the pathogen were as extreme or more extreme than the central 95% of the null distribution).

Comparison	Variable	Pathogen niche mean vs null distribution	Pathogen niche median vs null distribution	Pathogen niche SD vs null distribution	Pathogen niche range vs null distribution
Scenario 1	Temperature	–	–	–	–
	Precipitation	–	–	–	–
Scenario 2	Temperature	–	–	lower	lower
	Precipitation	higher	higher	–	–
Scenario 3	Temperature	lower	lower	lower	lower
	Precipitation	–	–	lower	lower
Scenario 4	Temperature	lower	lower	lower	lower
	Precipitation	–	–	lower	lower
Scenario 5	Temperature	higher	higher	lower	lower
	Precipitation	higher	higher	–	–
Scenario 6	Temperature	–	–	lower	–
	Precipitation	higher	–	–	lower
Scenario 7	Temperature	–	–	–	–
	Precipitation	–	–	–	–

important difference should be noticed. In the univariate analyses, as summary statistics from pathogen testing data are compared against a null distribution derived from sampling host data, conditions analogous to accessible environments (Soberón and Peterson 2005) are considered in the univariate non-parametric approach. That is, the univariate approaches are considering the set of environmental conditions for all hosts as those to which the pathogen could have had access. If some conditions used by the host have no pathogen records, it may be because of niche-based limitations, and these methods assess how nonrandom environmentally those gaps in pathogen records are. These limitations may be related to pathogen tolerance of or preference for certain environments, or environmental conditions limiting pathogen transmission from one host to another. They could also relate to inappropriate delimitation of relevant hosts for analysis, such as if a pathogen were recently introduced in a region, and has not yet reached all areas inhabited by the host. This last detail is important because critical biases can be introduced in analyses if the data have not been filtered carefully based on ecological and biological considerations (Barve et al. 2011; Machado-Stredel et al. 2021).

Our protocols did not detect clear signals of dissimilarity in Scenarios 6 and 7, in which the pathogen niche was larger than and overlapped with or included the host niche. This outcome derives from the type of information available for analysis—a set of host records of which some are infected—which makes it difficult to detect signals of dissimilarity because pathogen niches will be characterized incompletely. A more comprehensive characterization of pathogen niches may require consideration of a larger group of host species, which could inform about the type of conditions that are suitable or unsuitable for a pathogen. However, the fact that this latter set of information will be scarce in real applications highlights the utility of our protocols in exploring signals of niche in pathogens. We note that a topic of current interest is that of co-infections of multiple pathogen species (Collinge and Ray 2006)—although the current implementation of our methods is in terms of single pathogen species, a clear potential extension is that of simultaneous evaluation of environmental bias in distributions of multiple pathogen species.

Although we have presented these protocols in the context of tests for pathogen infections in host organisms, as mentioned in the Introduction, these methods can be useful in any situation in which (1) the entire universe of sampling can be characterized, and (2) the set of positive records will be a strict subset of that universe of sampling. This situation

would be manifested in cases such as an analysis based on a single sampling protocol (e.g., data from the U.S. Breeding Bird Survey; Sauer et al. 2013), or deriving from a single-investigator sampling protocol (e.g., regional trapping of insects, such that trap positions are known completely; Sciarretta and Trematerra 2014). As such, this set of approaches can be considered as a precursor to formal ecological niche modeling, testing at the outset whether any nonrandom environmental use (= ecological niche) is manifested by that species, at that extent, at that resolution, and in those environmental dimensions.

Acknowledgments

We thank our colleagues Eric Ngeno, Ram Raghavan, and Abdelkafar Alkishe for the case studies that led to development of this paper. This research was supported by a grant from the National Science Foundation (OIA-1920946).

Competing Interests

The authors have declared that no competing interests exist.

Software Availability

R scripts containing the functions needed to run analyses and plot results are provided in a GitHub repository[2].

Data Availability

The data required to run the analyses presented in this work can be obtained using the script available at GitHub[3].

Supplementary Materials

All supplementary materials can be accessed openly via KU Scholarworks[4].

Literature Cited

Acevedo, P., A. Jiménez-Valverde, J. M. Lobo, and R. Real. 2012. Delimiting the geographical background in species distribution modelling. J. Biogeogr. 39:1383–1390.

Anderson, M. J. 2017. Permutational Multivariate Analysis of Variance (PERMANOVA). Wiley StatsRef Stat. Ref. Online 1–15.

Anderson, R. P., and I. Gonzalez. 2011. Species-specific tuning increases robustness to sampling bias in models of species distributions: An implementation with Maxent. Ecol. Model. 222:2796–2811.

Araújo, M. B., R. P. Anderson, A. M. Barbosa, C. M. Beale, C. F. Dormann, R. Early, R. A. Garcia, A. Guisan, L. Maiorano, B. Naimi, R. B. O'Hara, N. E. Zimmermann, and C. Rahbek. 2019. Standards for distribution models in biodiversity assessments. Sci. Adv. 5:eaat4858.

Barve, N., V. Barve, A. Jiménez-Valverde, A. Lira-Noriega, S. P. Maher, A. T. Peterson, J. Soberón, and F. Villalobos. 2011. The crucial role of the accessible area in ecological niche modeling and species distribution modeling. Ecol. Model. 222:1810–1819.

Collinge, S. K., and C. Ray. 2006. Disease Ecology: Community Structure and Pathogen Dynamics. Oxford University Press, New York.

Eisen, R. J., and C. D. Paddock. 2021. Tick and tickborne pathogen surveillance as a public health tool in the United States. J. Med. Entomol. 58:1490–1502.

Enriquez-Urzelai, U., M. R. Kearney, A. G. Nicieza, and R. Tingley. 2019. Integrating mechanistic and correlative niche models to unravel range-limiting processes in a temperate amphibian. Glob. Change Biol. 25:2633–2647.

Etherington, T. R. 2019. Mahalanobis distances and ecological niche modelling: Correcting a chi-squared probability error. PeerJ 7:e6678.

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

Hijmans, R. J. 2019. raster: Geographic data analysis and modeling. R package.[5]

Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol. 25:1965–1978.

Ingenloff, K., and A. T. Peterson. 2021. Incorporating time into the traditional correlational distributional modelling framework: A proof-of-concept using the Wood Thrush *Hylocichla mustelina*. Methods Ecol. Evol. 12:311–321.

Machado-Stredel, F., M. E. Cobos, and A. T. Peterson. 2021. A simulation-based method for identifying accessible areas as calibration areas for ecological niche models and species distribution models. Front. Biogeogr. 13:e48814.

Manley, P. N., M. D. Schlesinger, J. K. Roth, and B. Van Horne. 2005. A field-based evaluation of a presence-absence protocol for monitoring ecoregional-scale biodiversity. J. Wildl. Manag. 69:950–966.

Meek, P. D., G.-A. Ballard, K. Vernes, P. J. S. Fleming, P. D. Meek, G.-A. Ballard, K. Vernes, and P. J. S. Fleming. 2015. The history of wildlife camera trapping as a survey tool in Australia. Aust. Mammal. 37:1–12.

M'ikanatha, N. M., R. Lynfield, C. A. V. Beneden, and H. de Valk. 2008. Infectious Disease Surveillance. John Wiley & Sons.

Nuñez-Penichet, C., M. E. Cobos, and J. Soberon. 2021. Non-overlapping climatic niches and biogeographic barriers explain disjunct distributions of continental *Urania* moths. Front. Biogeogr. 13:e52142.

Peterson, A. T., J. Soberón, R. G. Pearson, R. P. Anderson, E. Martínez-Meyer, M. Nakamura, and M. B. Araújo. 2011.

[2] https://github.com/marlonecobos/host-pathogen.
[3] https://github.com/marlonecobos/host-pathogen.
[4] https://doi.org/10.6084/m9.figshare.16870159.v1.
[5] https://CRAN.R-project.org/package=raster.

Ecological Niches and Geographic Distributions. Princeton University Press, Princeton.

Phillips, S. J., M. Dudík, J. Elith, C. H. Graham, A. Lehmann, J. Leathwick, and S. Ferrier. 2009. Sample selection bias and presence-only distribution models: Implications for background and pseudo-absence data. Ecol. Appl. 19:181–197.

R Core Team. 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Sauer, J. R., W. A. Link, J. E. Fallon, K. L. Pardieck, and D. J. Ziolkowski Jr. 2013. The North American Breeding Bird Survey 1966–2011: Summary analysis and species accounts. North Am. Fauna 79:1–32.

Sciarretta, A., and P. Trematerra. 2014. Geostatistical tools for the study of insect spatial distribution: Practical implications in the integrated management of orchard and vineyard pests. Plant Prot. Sci. 50:97–110.

Soberón, J., and A. T. Peterson. 2005. Interpretation of models of fundamental ecological niches and species' distributional areas. Biodivers. Inform. 2:1–10.

Warren, D. L., A. N. Wright, S. N. Seifert, and H. B. Shaffer. 2014. Incorporating model complexity and spatial sampling bias into ecological niche models of climate change risks faced by 90 California vertebrate species of concern. Divers. Distrib. 20:334–343.