University of New Mexico
UNM Digital Repository

Earth and Planetary Sciences ETDs | Electronic Theses and Dissertations

Summer 8-1-2023

Testing the Niche Center Hypothesis in the Fossil Record of Atlantic Bivalves

Rhiannon Z. Nolan
University of New Mexico

Follow this and additional works at: https://digitalrepository.unm.edu/eps_etds

 Part of the Paleobiology Commons

Recommended Citation

Nolan, Rhiannon Z.. "Testing the Niche Center Hypothesis in the Fossil Record of Atlantic Bivalves." (2023). https://digitalrepository.unm.edu/eps_etds/347

This Thesis is brought to you for free and open access by the Electronic Theses and Dissertations at UNM Digital Repository. It has been accepted for inclusion in Earth and Planetary Sciences ETDs by an authorized administrator of UNM Digital Repository. For more information, please contact disc@unm.edu.

Rhiannon Z. Nolan
Candidate

Earth & Planetary Sciences
Department

This thesis is approved, and it is acceptable in quality and form for publication:

Approved by the Thesis Committee:

Dr. Corinne Myers

Dr. Jason Moore

Dr. Louis Scuderi

TESTING THE NICHE CENTER HYPOTHESIS IN THE FOSSIL RECORD OF ATLANTIC BIVALVES

by

Rhiannon Z. Nolan

B.A. GEOSCIENCES,
SMITH COLLEGE, 2019

THESIS

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Master of Science
Earth and Planetary Sciences

The University of New Mexico
Albuquerque, New Mexico

August, 2023

ACKNOWLEDGEMENTS

I would first like to acknowledge the support of the Museum Research Traineeship throughout this project. This work was supported by the National Science Foundation under Grant No. 2021744. Support from this grant facilitated both field experience and the opportunity to present this material at the annual meeting of the Geological Society of America.

Thank you also to the many people without whom this project would not have been possible. Many thanks to my advisor, Cori Myers, who brought up the idea for this project and helped me worth through the theory. Thanks to my other committee members, Jason Moore – for talking through the taphonomic biases with me – and Louis Scuderi – for spending time dissecting how best to tackle geographic problems with me and developing the methodologies. Thanks to Fernando Machado-Stredel, Nick Freymueller, Jonė Naujokaityte and Ceara Purcell for their help working through the R code, statistics, and all other details of the data analysis. Thanks to Alexander Farnsworth for finding and providing paleoenvironmental climate models used in this analysis.

Finally, thanks to my friends and family, without whom I would never have gotten this far.

TESTING THE NICHE CENTER HYPOTHESIS IN THE FOSSIL RECORD OF ATLANTIC BIVALVES

by

Rhiannon Z. Nolan

B.A., Geosciences, Smith College, 2019
M.S., Earth and Planetary Sciences, The University of New Mexico, 2023

ABSTRACT

Paleoecological analyses of six shallow marine bivalves were conducted to test the Abundant Center Hypothesis using data from the fossil record of the Pleistocene through modern day. This hypothesis predicts the highest abundance of a species is at the center of the geographic or environmental range, decreasing toward the edges. In geographic space, distances to a centerline within a geographic range were variably correlated with population abundances, and some species displayed a sharp drop-off in abundance as distance increased. In environmental space, bivalve species showed moderate correlations between abundance and centrality when measured using cumulative data across the last 2.8 Ma. Shorter time bins across that duration show no consistent patterns, potentially indicative of an abundance-centrality pattern apparent only in the species' fundamental niche, which is best measured cumulatively over geologic time. These results suggest caution should be taken when interpreting modeled environmental preferences, particularly over short durations.

TABLE OF CONTENTS

LIST OF FIGURES

INTRODUCTION

One of the core motivations of ecology has always been characterizing the distribution of organisms across the environment and the mechanisms involved in these patterns. From the pioneering studies of the effects of climatic change on invasive species (Johnston, 1924), and the regionality of agricultural crop success (Visher, 1915), a rich study of species distribution modeling has grown. All species distribution models (SDMs) attempt to characterize the relationships between the presence of a species and the geographic context of the environment it inhabits, generally with the goal of determining how the confluence of geographic and environmental factors predict occurrence or elicit distributional change (Guisan and Zimmermann, 2000). One such model is the Abundant Center Hypothesis (ACH), which was formalized by Brown (1984) as a potential model for the abundance distribution within a species' geographic range based on the assumption that geographic range could be used as a proxy for environmental suitability. According to the ACH, the abundance of a species is greatest at the geographic center of that species' distributional range, where the environment is presumed to be most suitable for that species (Brown, 1984). If the ACH holds true, the center of the distributional range would then likely be associated with net population increase, while the peripheries of the range, containing less environmentally favorable conditions, would lead to an insuperable decline in fitness resulting in net population loss at the periphery (Brown, 1995). This pattern has not been found in a majority of studies performed in geographic space, likely due to the heterogeneous distribution of environments mapped onto geography (i.e., environments do not smoothly change across geographic barriers such as rivers, mountain ranges or steep marine bathymetric gradients) (Myers and Saupe, 2013; Sagarin and Gaines, 2002). However, there

is mixed evidence of an abundance-centrality pattern when species' niche characteristics are modeled in environmental space - dubbed a Niche Center Hypothesis (NCH) (Dallas et al., 2017; Dallas and Hastings, 2018; Martínez-Meyer et al., 2013; Osorio-Olvera et al., 2020). This environmentally delimited distribution modeling relies on Ecological Niche Modeling (ENM), which predicts species' distributions based on abiotic habitat suitability (e.g., Peterson and Soberón, 2012). While generally ENM is used to map habitat suitability onto geographic maps, it can also plot distributions of a species entirely within n-dimensional niche-space. NCH studies of present-day species are mixed and consequently any generalizable pattern that holds across many species has yet to be identified. However, this may be a result of the short timescale over which species' environmental niches are often characterized. Modern ecological studies may incorporate years, or even decades, of species-environment data. However, species are known to survive in the fossil record for over a million years, and average 11 million years within marine invertebrates (Lawton and May, 1995). Studying ACH/NCH patterns in deep time expands the breadth of geographic space and environmental conditions that individuals from a species experienced – including environmental conditions that no longer exist in a snapshot of the modern Earth – and allows for the full dispersal potential of a species to be realized geographically. A time-transgressive collection of species-environment data also provides a closer approximation of the fundamental environmental niche of the species than can be measured using modern data alone, which improves the potential to observe any generalized NCH patterns characterizing the relationship between species and their environment over species' lifetimes.

Understanding species-environment interactions and the ways they drive survivorship is critical to predicting the indicators of extinction risk imposed by modern climate change.

Even in the fossil record, the concept of a range-based center-periphery abundance change has been most often studied in the context of speciation and extinction. The edges of geographic ranges are areas of particular importance in the origination of biodiversity and determination of extinction (Mayr, 1963). This concept is supported by some modern studies of center-periphery dynamics within the geographic range, which show the highest rates of gene flow and greatest genetic diversity in the centers of many species geographic ranges (e.g., Jin et al., 2020). Although in geographic space the correlation between distance from a range centroid and genetic diversity is not ubiquitous across species, there is strong support for an abundant center of genetic diversity in environmental space (e.g Lira-Noriega and Manthey, 2014). The underlying mechanism of extinction (excluding mass extinctions) is generally considered to be habitat loss (Wiens and Slaton, 2012) and this loss is frequently driven by changes in environmental conditions within that space (e.g., marine anoxia or increased temperature). Thus, size and distribution of suitable environmental space is a crucial factor in assessing extinction risk, particularly as modern climate change drives rapidly changing availability in environmental space, while geographic landscapes are not altered on the same temporal scale. Range shifts in response to climate change have been well documented among a wide variety of both extant and extinct species, where species with the most restricted ranges and/or experiencing severe range contraction are more susceptible to elimination of suitable habitat, leading to extinctions (Parmesan, 2006).

Taxonomic Selection

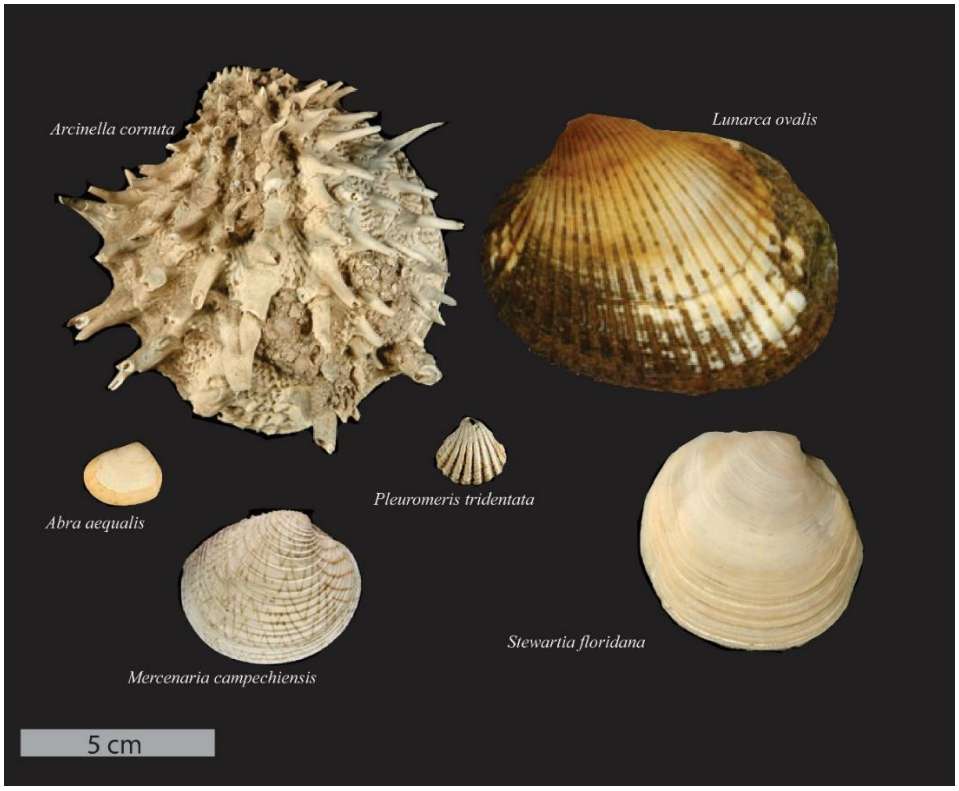

Figure 1: Images of the bivalve species used in this study. (Image sources, top to bottom, left to right: 1. NeogeneAtlas UF 217954; 2.J. Trausel & F. Slieker; 3. NeogeneAtlas UF 214314, 4. NeogeneAtlas UF 224948, 5. J. Trausel & F. Slieker, 6. NeogeneAtlas UF 61957)

The use of modern bivalve species (see Figure 1 for visual reference) with roots in or before the Pleistocene in this study permits analysis using known biotic characteristics observed in both modern populations and in the fossil record. Bivalve larval dispersal rates in particular have a strong control on observed distribution and abundance patterns because dispersal of these taxa occurs primarily during the free-swimming larval stage, during which larvae use ocean currents to transport themselves to suitable habitats and thereby ensure population mixing and range expansion (Blackwelder, 1981). Some studies suggest most bivalve species are extremely sensitive to climatic shifts, particularly thermal change (Thomas and Bacher, 2018). Through larval dispersal, the distribution of bivalve species

likely rapidly reaches environmental equilibrium during periods of climatic change, which occurred geologically frequently throughout the Pleistocene.

Ecological Niches and Niche Modeling

The fundamental ecological niche, an n-dimensional hypervolume in environmental space (e-space) in which a species can experience net population growth based on its abiotic tolerance, defines the axes along which the NCH patterns are expected to occur (Brown, 1984; Hutchinson, 1957; Martínez-Meyer et al., 2013). In contrast, the realized ecological niche, which is the e-space in which the species actually occurs, is the intersection of the fundamental niche with other factors influencing species survival across geography, including biotic factors and movement (dispersal) capacity (Soberón and Peterson, 2005). Biotic factors include all interactions with other species in an ecosystem that affect a species distribution, growth, or abundance, which is notoriously difficult to capture in the fossil record. Even if biotic interactions could be identified precisely, changes in the intensity and direction of biotic interactions can also be driven by climatic changes, further complicating system of factors influencing species distribution (Blois et al., 2013). Examples of biotic factors that drive environmental patterns might include symbiotic relationships, such as in the case of chemosymbiotic bivalves which necessarily must be geographically limited by the environmental tolerances and requirements of their symbiotes, or competition for resources, such as the competition for light among trees in densely populated forests (Grams and Andersen, 2007; Taylor and Glover, 2010). The movement aspect of the realized niche is determined by the ability of the taxon to disperse in search of suitable abiotic and biotic habitat within an individual lifetime. In highly motile taxa, dispersal potential is high, so

movement is not a limiting factor distinguishing the realized and fundamental niche, but in predominantly sessile organisms that may disperse only in the seed or larval stage (e.g., bivalves, corals, plants, etc.), movement may play a dominant role in delimiting the realized niche, and thereby geographic range location and size. Direct measurement of the fundamental niche of a taxon is not possible, because at any given time individuals can only occupy the current realized niche of the species and may not reach the full potential of the fundamental niche due to dispersal limitations, prohibitive biotic interactions, or a simple lack of the full range of potentially favorable environmental conditions in the current climatic regime. Thus, the difference between the realized niche and the fundamental niche of a species involves a difference in the quantity of time available, the environmental conditions accessible, and the complexity of the biotic interactions inherent in ecosystems. Although biotic factors are difficult to interpret, using modern species with a robust record through the Pleistocene allows a deep-time approach to observe species at their maximum dispersal capability, and encompasses a wide range of climatic regimes, which results in a closer approximation of the fundamental niche than is possible to achieve in modern studies.

All predictive ENM models (e.g., Maxent models, shown Figure 2) require an assumption of niche stability over time – that is the expectation that a species will retain its ecological niche across that species' lifetime (Peterson et al., 1999; Stigall, 2014). Niche stability within higher taxonomic groups or throughout speciation is not a requirement on the scale of this study, but any niche change of a single species across the lifetime of the species would render ENM analyses uninterpretable and any test of NCH intractable. Studies at the genus or family level have demonstrated broad niche stability across evolutionary turnover events, but the nature of the fossil record is such that determination of species-level niche

stability is difficult to quantify (e.g., Brame and Stigall, 2013; Di Marco et al., 2021). Predictions of future occurrence distributions hinge on the assumption that future individuals of the species exhibit identical environmental tolerances to modern specimens, so a species with changing niche parameters cannot be modeled using ENM.

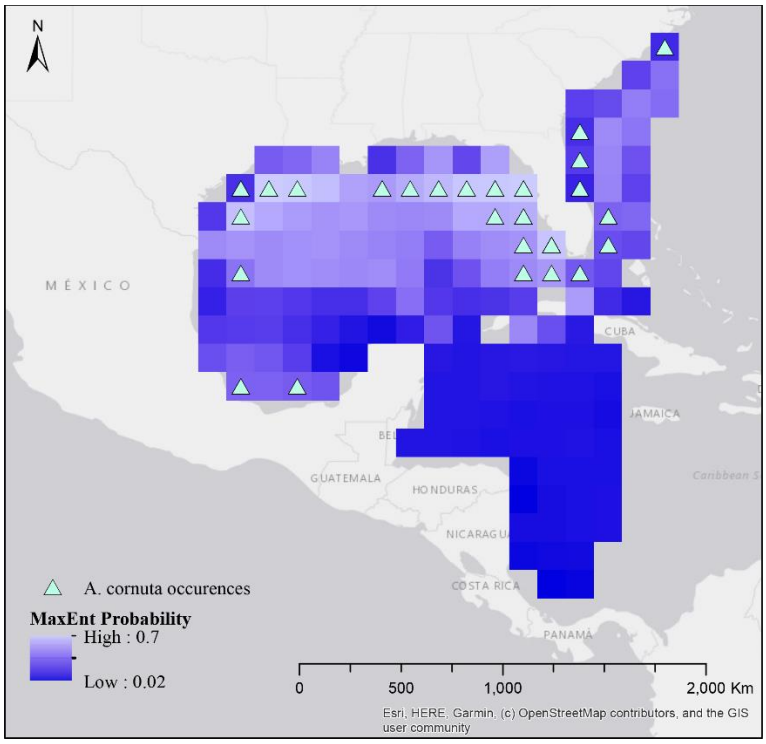

Figure 2: Map showing Maxent model derived logistic probability of occurrence of Arcinella cornuta *in the Holocene, across its geographic range determined using an alpha hull. White triangles show occurence points from the Holocene, adjusted geographically to the nearest available environmental data value. Darker tiles indicate less environmental suitability, and lighter tiles indicate greater environmental suitability. Suitability is interpreted from the environmental variable (see Appendix A) values present at the occurrence points and show a pattern concordant with the hypothesis that ocean depth is a driving factor in determining environmental suitability of* A. cornuta.

Geologic and Environmental Background

Significant variation in sea level, global ice volume, and temperature have been recorded across the Pleistocene epoch, in conjunction with Milankovitch cyclicity and concurrent glacial-interglacial cycles (Pillans et al., 1998; Pisias and Moore, 1981). Despite up to ~130 meter swings in global sea level, there is no evidence that tectonic uplifting or subsidence contributed significantly to sea level changes, particularly in the Gulf Coastal Plain (GCP) and Caribbean Sea (Fairbanks, 1989; Ludt and Rocha, 2015). The northern section of the Gulf Coast margin is built on the transform margin created during the opening of the Gulf of Mexico in the Jurassic, and along with the Atlantic Coastal Plain (ACP) has experienced little tectonic change during the Pleistocene (Bird et al., 2005; Blackwelder, 1981; Davis and Mitrovica, 1996). Ocean circulation in the Gulf of Mexico is driven by the Loop Current System (LCS), and exhibits a significant salinity gradient and seasonal variability mediated by freshwater inputs from North American rivers and saline concentration due to the semi-enclosed nature of the Gulf (Brokaw et al., 2019). The LCS was established in the late Pliocene, with the closure of the Isthmus of Panama, and varied in strength across the Pleistocene, driven by climatic temperature changes (Hübscher and Nürnberg, 2023). The varying salinity, temperatures, and circulation-driven mixing in the GCP and lower ACP across the duration of this study widens the range of climatic extremes experienced by fauna living in those areas, while relative tectonic stability avoids conflation with uplift or subsidence-driven habitat changes.

For the purposes of this study, the Pleistocene was divided into three substages - early, middle, and late Pleistocene, with durations 2.588 - 0.781 Ma, 0.781 - 0.126 Ma, and 0.126 - 0.012 Ma, respectively, as designated by International Union of Geological Sciences

(Gibbard et al., 2010). The Holocene was treated as a time bin including data dated from 0.012 Ma to present. Analyses were conducted using each time bin individually as well as a cumulative set of all data. The early and middle Pleistocene, as defined here, were characterized by ~ 41 kyr oscillations in global ice volume, while the late Pleistocene was dominated by more extreme but lower frequency variability, with sea surface temperatures fluctuating by as much as ~2 °C and surface temperatures varying by as much as 10 °C between the Last Glacial Maximum and the pre-industrial Holocene (Malakhova and Eliseev, 2020). The glacial-interglacial cyclicity during the Pleistocene also drove rapid and extreme variability in sea-level, which would have significant effects on the distribution of shallow marine bivalves (Shelley et al., 2020). Rapid sea level change would have forced range shifts and encouraged dispersal to the maximum extent of the environmental suitability of each species, further reinforcing the filling and re-filling of each new addition to the realized niche.

METHODS

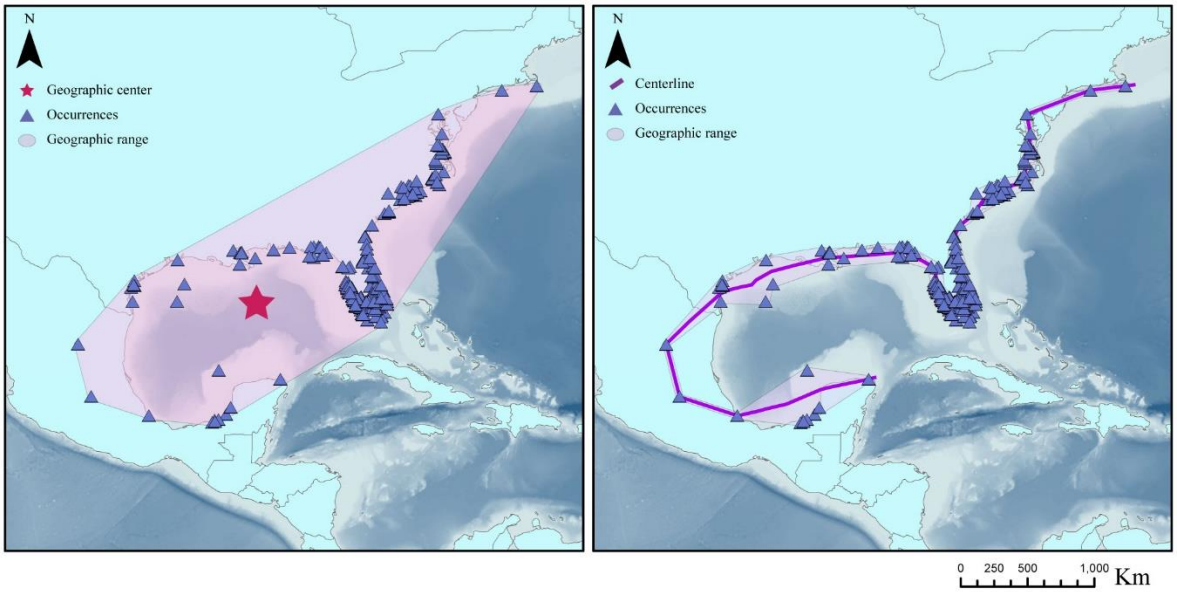

Figure 3: Map showing the geographic range of Mercenaria campechiensis *(pink enclosing polygons) as determined by each method used in this study; supporting taxon occurrences are represented by purple triangles. In species whose range is strongly constrained by environmental features, such as water depth, the geographic center of a convex hull surrounding the occurrence points (left) can be erroneously measured in an environmentally unsuitable location for the species, such as the middle of the Gulf of Mexico. A polygon buffer surrounding each occurrence, but following the known coastline (right), has a centerline that better describes a geographic center habitable by the species and reflecting its geographic range.*

Taxa and Model Setup

Occurrence data, with latitude, longitude, and geologic time designation to the early, middle or late Pleistocene, or Holocene were downloaded from the Paleobiology Database, and recent data for the same six species was downloaded from the Global Biodiversity Information Facility. The two datasets were filtered for duplicate records, and within each species occurrence records with duplicate coordinates were removed. Species occurrences were also limited to specimens from published literature or established museum specimens and filtered to include only occurrences found in the Americas. This rigorous vetting resulted

in six well-sampled bivalve species from the Gulf of Mexico and Atlantic Coastal Plains (GCP and ACP, respectively) as the test subjects for this case study (see Table 1). All species have well-sampled modern distributions and a robust fossil record stretching back to at least the early Pleistocene.

Table 1: Bivalve species used for ACH testing with Life Mode used to interpret dispersal capacity, used in determination of geographic range buffer distances.

Species	**Common Name**	**Age Range**	**Life Mode (Dispersal Capacity)**
Abra aequalis	Atlantic abra	Late Pliocene-Present	Facultatively mobile (25 km)
Arcinella cornuta	N/A	Late Miocene-Present	Sessile (10 km)
Lunarca ovalis	Blood ark	Pliocene-Present	Facultatively mobile (25 km)
Mercenaria campechiensis	Südliche Quahog-Muschel	Middle Miocene-Present	Facultatively mobile (25 km)
Pleuromeris tridentata	Threetooth carditid	Pliocene-Present	Facultatively mobile (25 km)
Stewartia floridana	N/A	Early Pleistocene-Present	Facultatively mobile (25 km)

In both geographic and environmental-space (e-space) tests, the Abundant Center Hypothesis (ACH) and Niche Center Hypothesis (NCH) were tested as linear phenomena, where abundance and distance (from central feature) data were fit to linear models. The ACH/NCH has commonly been tested as a linear phenomenon in modern studies (e.g., Adhikari et al., 2018; Sagarin and Gaines, 2002), so this first test is directly comparable to those studies. The tests were also performed using log-transformed abundance data, to test whether species' response to environmental differences can be modeled as a threshold phenomenon. If species tolerance appears as a threshold rather than a gradient, a log transformation will decrease the skew in the abundance values, making the abrupt transition from presence to absence graphically visible.

Geographic Range Reconstruction

Using ArcGIS, geographic ranges were determined for each species by first creating a buffer surrounding each occurrence point, the size of which depended on its dispersal ability as interpreted from its recorded life mode (Table 1). Geographic range size was measured by calculating convex hulls surrounding all buffered occurrence points, and then editing the resultant polygons to avoid the deeper water offshore, and thereby better represent the realized geographic range given the strong bathymetric control over shallow marine bivalve ranges. The traditional method of calculating geographic ranges based on simple convex hulls was also tested (see Supplementary Information for model results), but this method presented an issue wherein shallow bivalve occurrences tended to follow coastlines around the Gulf of Mexico, resulting in range centroids occurring in waters too deep for the species to survive (example in Figure 3). For ACH testing, the center of the edited polygon was taken as the centerline of the polygon, generated by calculating multiple ring buffers using ArcGIS at 10 km intervals and joining central meeting points of each buffer with a line. Geographic distances from each occurrence point to the centerline were measured using the Near tool in ArcGIS. Separate polygons and centerlines were measured for each species' cumulative occurrences across the Pleistocene and Holocene, and then within each time bin individually for each species.

Environmental Distance Reconstruction

Modern environmental layers used for Holocene samples were downloaded from MARSPEC (Ocean Climate Layers for Marine Spatial Ecology) and interpolated using the 'Resample' tool in ArcGIS to match the resolution of paleoclimatic layers (Sbrocco and

Barber, 2013). Environmental layers for the Pleistocene time bins were sourced from simulations run by the Bristol Research Initiative for the Dynamic Global Environment (BRIDGE) group (Huntley et al., 2023). For the early Pleistocene, the teIXG run was used, from 800 Ka; for the middle Pleistocene, the teiXN run was used, at 332 Ka, and for the late Pleistocene the teixH run was used, at 60 Ka (Huntley et al., 2023). Run ages for each time bin were chosen from midpoints within the time encompassed by each time bin.

To calculate environmental distances, the *princomp* function of the 'stats' package in R was used to perform a principal components analysis (PCA) of the environmental data present at each occurrence point. The first three components of the analysis (for this data, generally describing ~ 90% of the variability in the data) were extracted as coordinates in e-space as a cumulative PCA of all the data across the Pleistocene and Holocene. Separately, the original data was subset by time bin and individual PCAs were performed for each time bin, creating the e-space defined for only that time bin. Minimum volume ellipsoids (MVEs) define the smallest volume ellipsoid enclosing all occurrence points in 3-dimensional space, in this case using coordinates in e-space defined by environmental niche PCA axes. The e-space MVE centroids were calculated using the *cov.rob* function of the 'MASS' package in R, and Mahalanobis distances for each point were calculated using the *Mahalanobis* function from the 'stats' package. Mahalanobis distance is a multivariate distance metric that reports the distance between each point in e-space and the distribution of points in the time bin, effectively scaling the PCA axes such that their variance is 1, and then calculating the distance, providing a measure of distance from the ellipsoid centroid. Each resultant dataset was displayed as a histogram of distance frequencies, exported as frequency counts, and a linear model was fitted to the counts using the *lm* function of the 'stats' package. The

frequencies were also transformed logarithmically and fitted to linear models with distance data. Spearman rank correlation coefficient tests were performed on these same variables to estimate statistical dependence between the two variables and effect size. The Spearman rank correlation measures how closely correlated the two variables are, regardless of linearity, so highly correlated data (defined as $\rho > 0.75$ in this study) would suggest multivariate distance from a center may have some predictive power of abundance in e-space, and thereby support the NCH.

Ecological niche modelling (ENM) uses multivariate and machine-learning algorithms based on known environmental variables and known presence points of a taxon to estimate environmental suitability across abiotic variable ranges, and using this, predicted distribution across a geographic range (e.g., Barve et al., 2011; Peterson, 2001; Peterson et al., 1999). Forecasting distributional and occurrence changes across geographic space is a powerful tool in conservation biology, and ENM is generally used in this field as a tool to estimate current occupation, measure environmental tolerance, and predict the response of a species to potential environmental changes (e.g., Qiao et al., 2015).

In a final test, ecological niche models were generated using the Maxent (version 3.4.4) algorithm (Phillips et al., 2006, 2004), under default parameters with logistic output (Phillips et al., 2017). Maximum entropy (Maxent) niche modeling requires only presence information (no verified absence data) and defines probability densities from the presence data and environmental rasters, minimizing the relative entropy between them to produce a relative likelihood surface that can be projected across geographic space (Elith et al., 2011a). An example of such a projected surface can be seen in Figure 2, which shows the logistic prediction for the bivalve *Arcinella cornuta* in the Holocene, and demonstrates the higher

environmental suitability calculated where the most occurrence points cluster. A Maxent niche model was run for each species, separated by time bin. Maxent-derived logistic predictions of environmental suitability (ranging from unsuitable "0" to perfectly suitable "1") for each training point input into the algorithm were exported. Maxent model significance was determined based on the Area Under the Curve (AUC) and maximum training sensitivity plus specificity (MSS) values reported by the algorithm. MSS calculation includes sensitivity and specificity values, where sensitivity refers to the proportion of presence points correctly predicted by the model, and specificity refers to the proportion of absence points correctly identified by the model (Allouche et al., 2006). The AUC is a direct measure of the model sensitivity and specificity plot, also called the receiver-operating characteristic (ROC) plot, where a value of 0.5 indicates the model predicted by random chance, so a value greater than 0.5 indicates the maxent model can predict more accurate occurrences than random chance (Phillips et al., 2006).

Testing the ACH/NCH

Each set of distances was exported as histogram frequencies per distance bin using the *hist* function in the 'raster' package in R. Linear models were fit to those exported distance frequencies to determine whether any range centrality relationships could be observed. Logarithmic transformations of data can be used to make evident patterns in data that are highly skewed, so linear models were also fit to the log of the frequency to verify if this transformation revealed any ACH/NCH patterns driven by threshold phenomena. If a strong abundance-centrality pattern were to be predictable in geographic space, the linear

model R^2 statistic should be high (for this study, $R^2 > 0.5$) with statistical significance ($p < 0.05$).

As a control used to determine the relationship between abundance and geographic range centrality as measured in studies of most other taxa (e.g., Attorre et al., 2013), geographic minimum volume ellipsoids and centroids were also calculated using the *ellipsoidhull* function in the 'cluster' package in R, and a 2-dimensional distance calculating function was defined to measure the distance from each point to that centroid (see Appendix A). Frequency histograms of the resulting distances were exported as frequency values and linear models were fit to the resultant values using the *lm* function from the 'stats' package. As before, a strong abundance-centrality pattern would be observable as a high linear model R^2 statistic (for this study set as $R^2 > 0.5$) with statistical significance ($p < 0.05$).

Finally, using Maxent derived niche models, the correlation between Mahalanobis distances for each occurrence point measured in e-space defined by PCAs and logistic predictions estimated from Maxent for each occurrence point was estimated by linear models fit using the *lm* function of the 'stats' package of R. A strong correlation coefficient would be indicative of a pattern wherein specimens that occur near the center of the ellipsoid representing their ecological niche in e-space are also occurring where their Maxent-defined environmental parameters are predicted to be the most suitable. That is, a strong correlation between these variables suggests the e-space niche centroid is in fact where the species-specific optimal environment occurs, which would support the Niche Center Hypothesis.

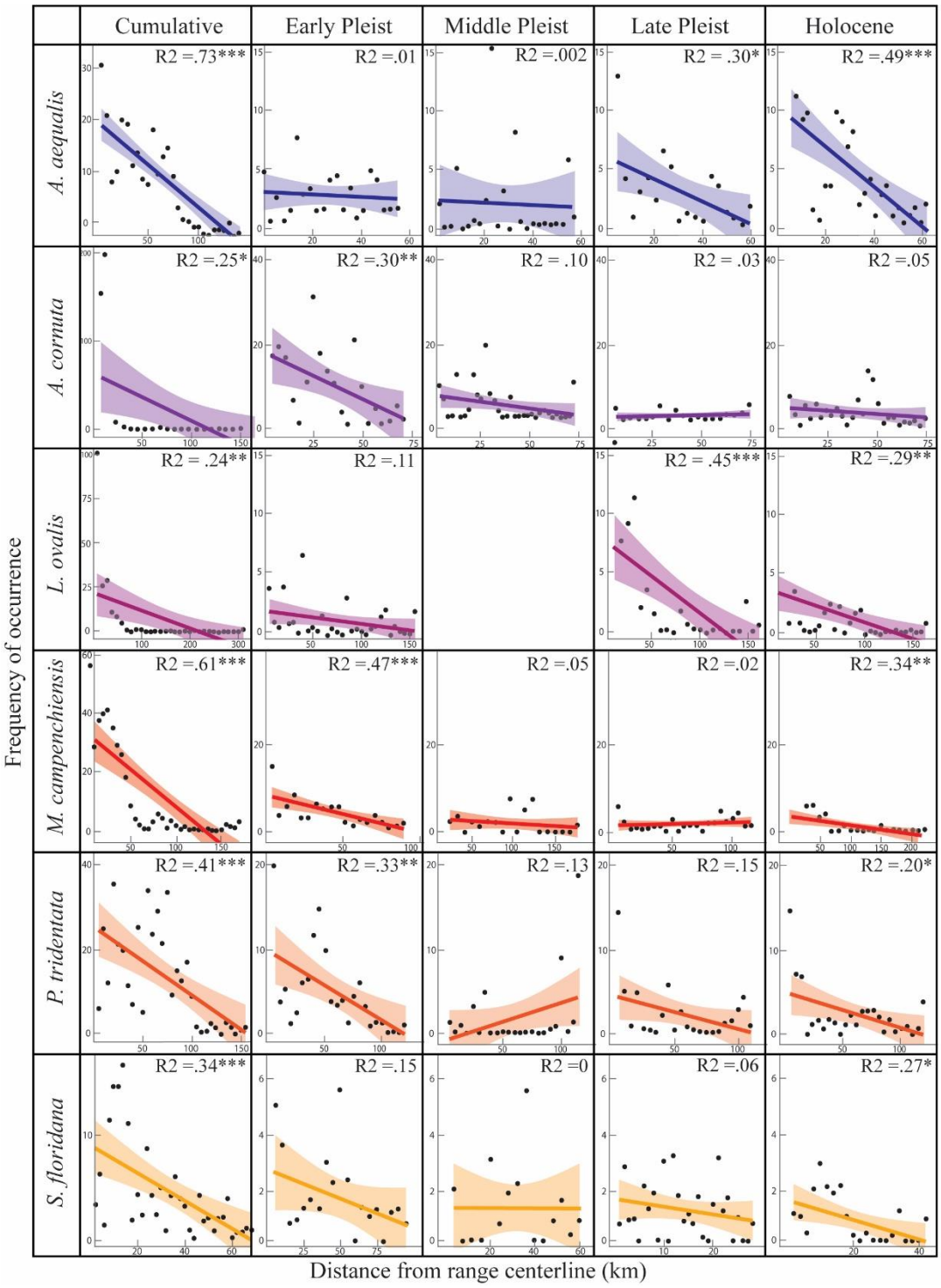

*Figure 4: Relationships between distance from the geographic range centerline and frequency of occurrence. Colored lines show linear model fit; lightly colored bands show 95% confidence interval of the linear model. R^2 values correspond to each linear model, with p values indicated by *p<0.05, **P<0.01 and ***P<0.001. Empty box indicates too few datapoints in time bin to fit model.*

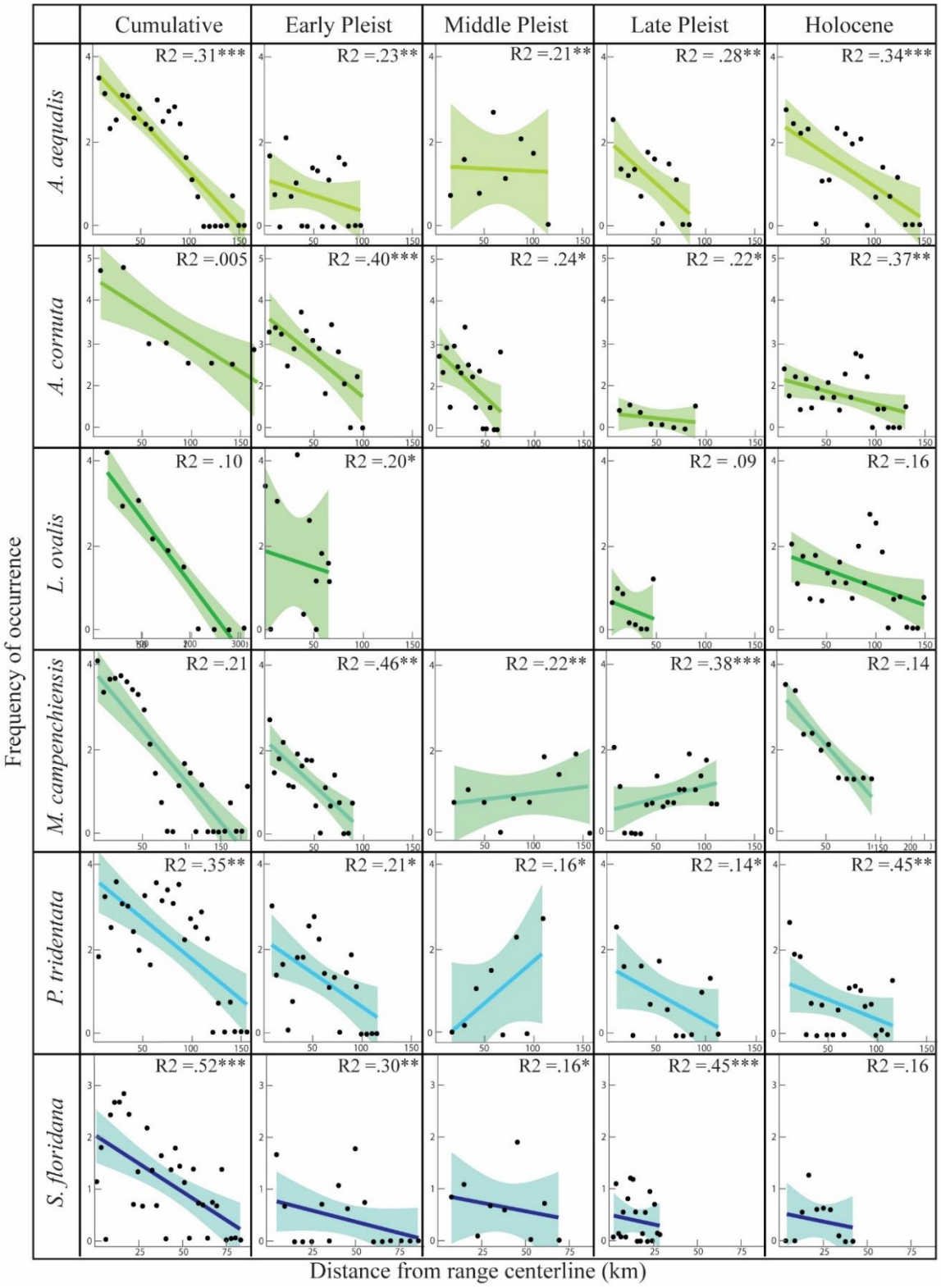

Figure 5: Geographic space log transformations of frequency data, with linear models fitted with distance from centerlines.

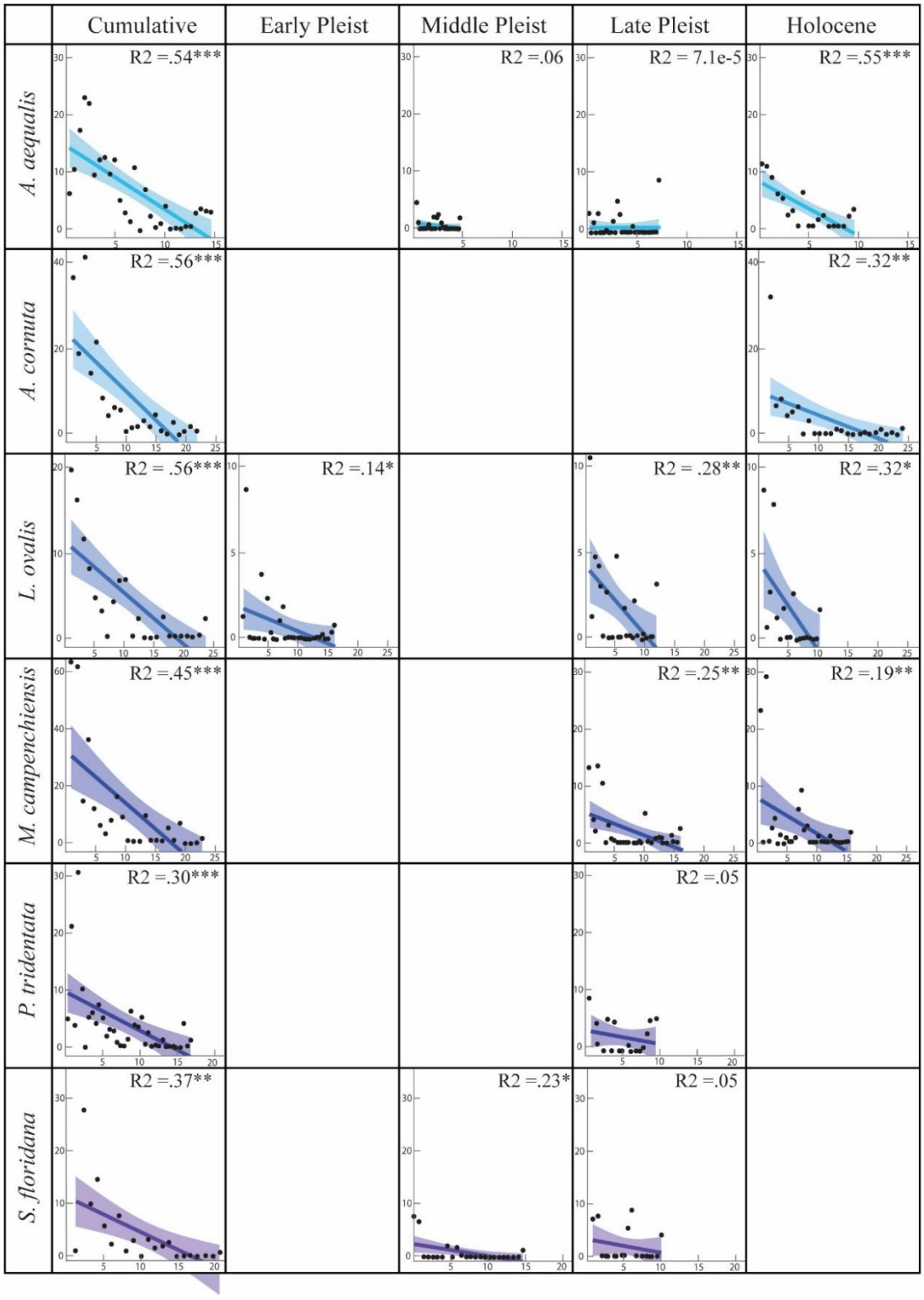

Figure 6: Linear model fits to Mahalanobis distance data calculated in e-space plotted against frequency of distance data. Empty boxes indicate linear models could not be fit, either because data was collinear, or the time bin contained too few datapoints.

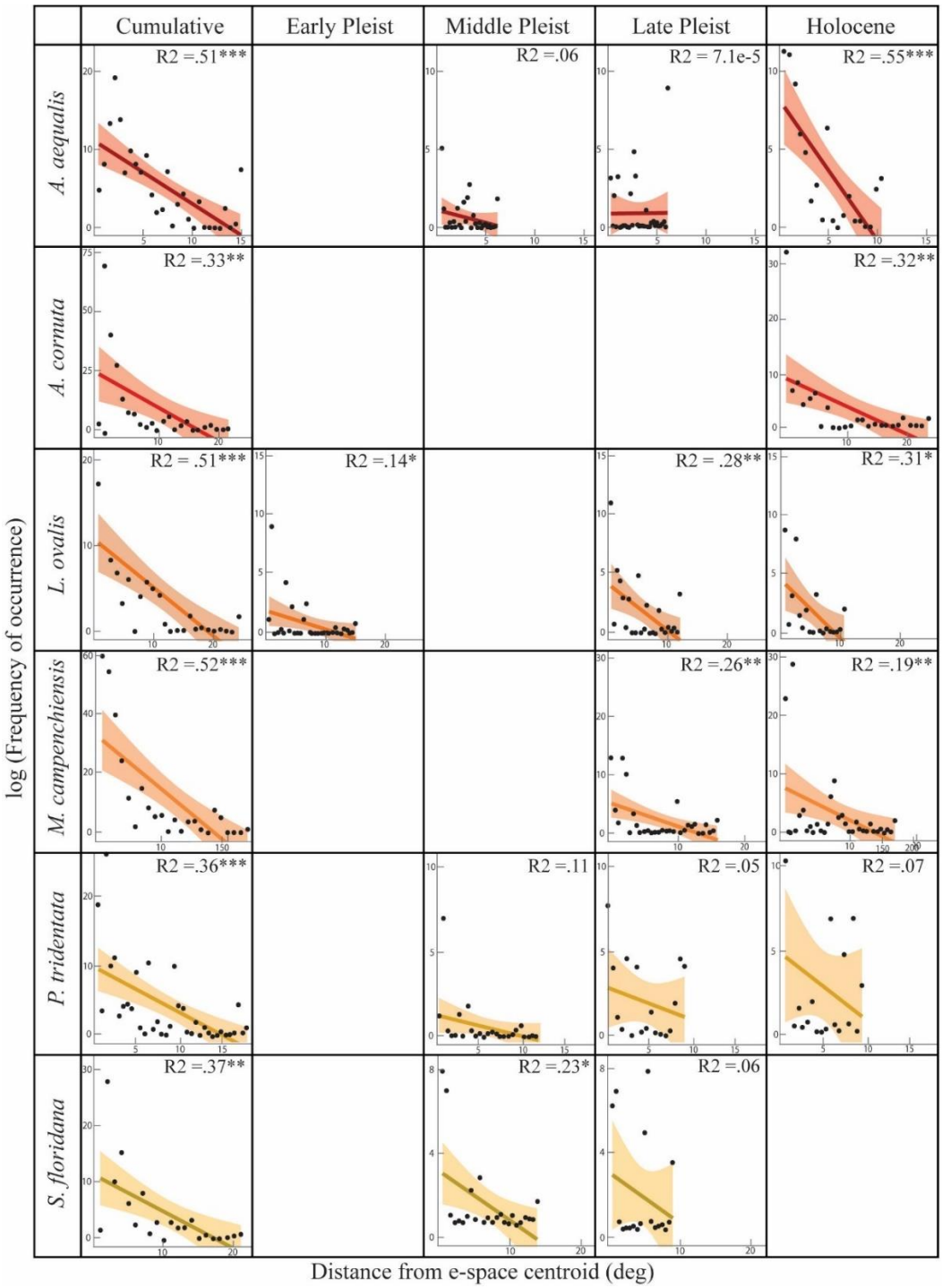

*Figure 7: Environmental space distance frequencies, with frequencies transformed logarithmically. Colored lines show linear model fit; lightly colored bands show 95% confidence interval of the linear model. R^2 values correspond to each linear model, with p values indicated by *p<0.05, **P<0.01 and ***P<0.001. Empty boxes indicate linear models could not be fit, either because data was collinear, or the time bin contained too few datapoints to fit a linear model.*

RESULTS

ACH in Geographic Space

There is some support for a geographic range-based abundance center pattern in these shallow marine bivalves when calculated using cumulative deep-time occurrence data, and when the center is calculated as a centerline within a range polygon drawn avoiding deep, uninhabitable water. This pattern is not generally observed in species studied in individual time slices across the Pleistocene through Holocene, consistent with previous studies showing no apparent pattern using the single time slice of modern data alone. There is no support for an abundant centroid of the geographic range of these species when calculated as a minimum volume ellipse in geographic space in cumulative or time series data, consistent with studies of many other taxa (Dallas et al., 2017; Martínez-Meyer et al., 2013).

In geographic space, using distance from a centerline as a measure of range centrality (Figure 4), all six bivalve species displayed statistically significant negative correlations between frequency of occurrence and distance from the range centerline when the cumulative data from the entire Pleistocene and Holocene is combined. Although statistically significant, in four out of six bivalve species a linear model displayed an $R^2 < 0.5$, suggesting a linear model does not explain a substantial amount of variance in the data. For all six species, relationships within individual time bins, including the Holocene time bin that includes all modern data, were less consistent, and generally displayed significant variability between time bins. The predictor effect plots shown in Figure 4 demonstrate visually that for many individual time bin analyses, the 95% confidence interval of the simple linear model is not significantly distinguishable from a slope, $m = 0$.

In the logarithmic transformation of g-space frequency data (Figure 5), when plotted against the distance from the centerline, more variable patterns are revealed across the six bivalve species, and within individual time bins. Again, all species have highly correlated patterns in the time-averaged, cumulative dataset. In particular, a linear model for *L. ovalis* explains 92% of the observed variance in the data. This pattern is also not ubiquitous across species however, and several species, such as *Abra aequalis, Mercenaria campechiensis*, and potentially *Pleuromeris tridentata*, show a visual drop-off in log-frequencies at a specific distance from the range centerline. Additionally, both *M. campechiensis* and *P. tridentata* fit a negative linear model when taken cumulatively, but also fit statistically significant positively sloped linear models in some time bins.

NCH in Environmental Space

In environmental space, there is support for an abundance niche center relationship when cumulative data across the Pleistocene through Holocene is used to designate an environmental niche encompassing deep time environmental fluctuations (Figures 6-7). All six of the species analyzed fit a statistically significant linear model R^2 value in the correlation between Mahalanobis distance and frequency of occurrence – i.e., the frequency at which specimens occur within each set of binned distances from the e-space centroid. In this Mahalanobis distance–frequency analysis, the cumulative data for each species presented statistically significant correlation and R^2 values ranging from 0.30–0.56, depending on the species. Spearman rank correlation coefficients for cumulative data were also all statistically significant, with $|\rho| = 0.82 - 0.95$.

When e-space frequency data was transformed logarithmically, R^2 values were similar to models fit using raw data, and again all six species fit linear models with statistical significance ($p < 0.05$) when analyzed using the cumulative dataset. Figure 7 shows linear models fit to the log of distance frequency for each species. The relationship between cumulative data and time bin analysis is very distinct in this analysis – in cumulative datasets, all six species fit a linear model with low R^2 ($0.33 < R^2 < 0.52$), but for several species all available time bins differ from the cumulative data results. In *A. aequalis,* cumulative $R^2 = 0.51$ with a p-value of 1.09e-05, but the middle and late Pleistocene models were non-significant, and the Holocene model shows log-frequency and Mahalanobis distance are once again highly linearly correlated. In *L. ovalis* and *M. campechiensis* all the models were significant, but every time bin model had lower R^2 than cumulative data for that species. Then, *P. tridentata* fit a linear model with an R^2 of 0.37 when measured cumulatively, but no time bin model was significant.

ENMs, often applying the Maxent algorithm, are widely used in interdisciplinary research across the fields of biology, conservation ecology, zoonotic pathology, and macro- and paleoecology for its applications in predicting unsampled occurrences and distributional changes in an evolving landscape. Maxent models run in this study generally reported AUC values > 0.75, with exceptions only in Holocene *P. tridentata,* Middle Pleistocene *M. campechiensis,* and Early Pleistocene *S. floridana* models. As an AUC value of 0.5 would indicate the model was no better than a randomly generated model, a high AUC value indicates these models have significant predictive power (Elith et al., 2011b, 2006; Franklin, 2010). However, response curves generated for the niche models (see Figures B1-B3, Appendix B) do not suggest good model fits. A response curve indicating the dependence of

habitat suitability on a variable would be expected to present with an approximately normal distribution plot, but response curved generated for these species did not show this relationship. Because Maxent models, although potentially significant, were not likely to be very informative (given incomplete response curves), these models are not considered further.

DISCUSSION

Potential Biases

The primary taphonomic biases that arise in fossil occurrence data include sampling bias, prevalent in less well-studied taxa or taxa that occur primarily in remote locations, and preservation bias, which can cause an artifactual pattern of low abundance particularly in taxa that lack hard skeletons or occur in environments where preservation potential is low. Time averaging in the paleo-record of shallow marine environments is most often caused by slow sedimentation rates, erosion and reworking that result in assemblages of bivalves accumulated over periods ranging from years to millennia (Behrensmeyer et al., 2000; Fürsich and Aberhan, 1990). Whereas community-level structure is often disturbed during deposition and diagenesis at the species level averaged over geologic time frames, presence data is not likely to be significantly altered (Kidwell and Bosence, 1991). Using presence frequency as a proxy for abundance, while not a perfect equivalence, will therefore preserve overarching distributional patterns between abundance and distance from a central feature.

Sampling biases can particularly affect studies done on organismal distributions, so in this case study only robustly sampled taxa were selected. The invertebrate fauna of the GCP and ACP have been comprehensively described in previous studies (e.g., Allmon et al., 1993; Anderson et al., 1991; Kolbe et al., 2011; Stanley, 1986). Marine bivalves have the highest preservation potential of any marine organism, providing an ideal dataset for testing ecological hypotheses in the fossil record - up to 85% of bivalve specimens likely get preserved in the fossil record (Valentine, 1989).The high preservation potential of bivalve taxa and net-sedimentation in the passive margin of this region result in a relatively

comprehensive fossil record, lending credibility to potential distributional patterns (Behrensmeyer et al., 2005, 2000; Kidwell, 2005, 1988).

Because it is not feasible to fully eliminate the effect of sampling and preservation bias from fossil data, it is possible the occurrence data used in this study show a "pull of the recent" effect, where specimens dated closer to the present may be more numerous or better preserved. In the case of a pattern originating exclusively from sampling bias, the cumulative patterns observed in the data would be expected to be driven primarily by Holocene data. Datapoints available within each time bin (see Table S1, Appendix A) are not overwhelmingly weighted toward the Holocene, apart from *Abra aequalis*, which does contain twice as many Holocene datapoints as datapoints in any other individual time bin, and represents 44% of the data, despite spanning a very small percentage of the total time. If bias is driving the patterns in the data, it should be apparent as increased noise that might obscure observed patterns.

It is possible that bias could account for some of the differences between time bins, particularly where time bins contained relatively few occurrence points. However, a lack of occurrence points due to lost preservation would not likely increase patterns in the distance between occurrence points and geographic/niche centrality, so any observed strong patterns likely indicate a realized signal within the data. Biases may be responsible for some of the non-significant model fits or the lack of clear pattern within individual time bins (Figures 6-7) but would not likely contribute to a stronger NCH pattern, so where NCH patterns are strong, they are most likely to be a signal of a realized distributional pattern. In g-space it is more likely that bias could be driving ACH patterns, because any bias toward sampling modern coastlines more robustly than further inland or offshore areas might bias the data

toward the artificial centerline of the modern coast. An example centerline determined for *Mercenaria campechiensis* in Figure 3 shows that the centerline does not perfectly line up with the modern coast, in some locations appearing offshore and in others further inland, as might be expected when a species is sampled across the Pleistocene and Holocene and experienced a wide range of sea levels. This does not preclude the possibility of sampling bias in all species but does add credibility to visible patterns that would need to be reevaluated if their centerlines followed the modern coastline well. Given that ACH patterns measured this way are more likely to be altered by sampling bias, these results are further interpreted with caution.

ACH in Geographic Space

Although neither raw nor log-transformed abundance data were strongly linearly related to geographic distance data ($R^2 < 0.5$), the statistically significant correlations between the two suggest there is some presence of an ACH pattern in shallow marine bivalves. Most apparent in the cases of *Abra aequalis, Mercenaria campechiensis and Pleuromeris tridentata,* the g-space log-transformed plots (Figure 5) show a sudden drop from abundant occurrences to zero. It is possible this could be a signal of a threshold relationship to physical environmental change, such as a steep depth increase, where the majority of the population lives within a habitable zone along the coast, and a few outliers explain the extreme distances represented. However, the presence of some ACH patterns in this data could be ascribed to the methods used in geographic range determination. Using a narrow range and a centerline is a more accurate depiction of the geographic ranges of bivalves than the more typical convex hull and centroids but might bias the data toward an

inevitable ACH pattern. This might be particularly observed where data is scarce and the width of the range might be represented by very few points, which inevitably decreases the distance to the centerline for those points.

The lack of strong patterns in individual time bins also suggests that smaller time slices again may not encompass sufficient time for populations to conform to an established geographic mean. In these cases, cumulative analysis of occurrences across the Pleistocene and Holocene reveals an emergent geographic pattern in some species that cannot be observed in any single time bin. The time bins used in this study vary significantly in duration, from 11,000 years of the Holocene to 1.8 million years of the Early Pleistocene, as demarcated in the PBDB database. The fact that even the 1.8 Ma represented by the Early Pleistocene bin is insufficient time and/or data to contain a strong geographic pattern further emphasizes the extent to which distributional patterns in modern data must differ from generalized patterns occurring across the lifetime of a species in geologic time.

As benthic, primarily sessile organisms with a narrow water depth tolerance, bivalves may be a species more likely to conform to an ACH distribution, simply by virtue of inhabiting a geographic range that can be roughly approximated as a linear feature. Studies such as Sagarin and Gaines (2002) have shown that using a 2-dimensionally defined range and measuring centrality along a line cannot be used as a general tool for predicting abundance in several species of marine invertebrates. As an organism-specific tool for determining abundance, geographic distance from a range centerline may hold some predictive power in bivalves, and possibly in other benthic sessile invertebrates, but alternative models may be needed for taxa with more complex life modes. Because of the dearth of evidence for g-space ACH relationships in motile marine and terrestrial species,

these results do not support the validity of a g-space ACH as a generalized trend among fauna over geological time scales (e.g., Dallas et al., 2017; Jin et al., 2020).

NCH in Environmental Space

In models relating Mahalanobis distances measured in 3-D e-space with frequency of occurrence (i.e., abundance), the lack of consistent relationships within each species' individual time bins suggests that limitations on dispersal or other biotic factors may play a significant role in determining distribution structure. Even these shallow marine bivalve species, all with the same general life modes and environmental tolerances, might possess biotic traits controlling distributions in ways not accounted for by environmental variables. It is also possible that the time-specific patterns are a result of a dispersal lag, where the time bins chosen do not encompass sufficient time for each species to reach a distributional equilibrium, accounting for the heterogeneity in e-space abundance-centrality patterns.

There were cases within log-transformed e-space models where cumulative linear fits had higher R^2 values than any of the individual time bins within that species. For example, *M. campechiensis* has much lower R^2 values in the late Pleistocene and Holocene than in cumulative time, but all are statistically significant. This could be a case of a pattern that is weak in any given time slice but is more predictive of general patterns in deep time, potentially driven by a logarithmic (threshold) vs. linear (gradient) relationship between abundance and the center of the fundamental environmental niche rather than the realized niche (supporting NCH). If the relationship is only present within the fundamental niche of the species, it might be weaker when a smaller portion of the niche realized in each time slice is measured. This is expected, as the portion of the available niche shifts with climatic

change, and only ever occupies part of the broader fundamental niche (Soberón and Peterson, 2011). If any given realized niche does not contain the center of the fundamental niche of a species, it would not be surprising if populations extant during that time did not appear to conform to any sort of abundance-centrality pattern.

In another interesting case, the log abundance of the bivalve *P. tridentata* fits a linear model when e-space is measured cumulatively, but all three available time bins showed no significant linear relationships with e-space distance. This could be a more extreme case of the disparity between realized and fundamental niche, where a cumulative analysis is the only way to measure enough of the fundamental niche to observe any relationship between abundance and niche centrality.

In all but one of the species analyzed, cumulative data displays stronger correlation between abundance and niche centrality than any single time bin, but the fact that they do so with varying degrees of support suggests there may be other factors vital in the differences between the species themselves that are influencing patterns in distribution within e-space. The one exception to this pattern, *A. aequalis*, showed better model fit in the Holocene than cumulatively, which might be expected if an abundance and intensive sampling of modern specimens is driving the patterns in the data. This could be an indication of bias in the data, particularly as the same pattern of Holocene data appearing to drive the cumulative data slope and pattern is present for *A. aequalis* in geographic space as well. Alternatively, it could be evidence that for this species, the Holocene climatic regime is a good match for the species' fundamental niche. In this case, Holocene abundances of *A. aequalis* would better match the NCH pattern than earlier time bins that do not sufficiently describe the environments of the fundamental niche.

What Structures Marine Bivalve Distributions?

Geographic range has been described as one of the emergent properties of a species that can be most easily quantified and used to predict species survivorship over time and across changing environmental conditions (Hunt et al., 2005; Jablonski, 1986). Brown's original concept of the Abundant Center Hypothesis envisioned a pattern observable across the geographic range of a species, but because such a pattern has only been variably observed, it is likely not a generalizable biogeographic rule that can be used to predict distributions across time and taxa (Brown, 1984; Dallas et al., 2017; Osorio-Olvera et al., 2020). However, geographic range is a complex variable controlled by several factors including abiotic habitat suitability, dispersal ability, and biotic interactions between individuals (Soberón and Peterson, 2005). In this study we focused on the characterization of abiotic habitat suitability (using two methods: MVEs and ENMs) to test for abundance-niche center relationships because many studies have found that the spatial extent of the fundamental ecological niche determines population density and growth rates (e.g., Lira-Noriega and Manthey, 2014; Martínez-Meyer et al., 2013). However, the results showed inconsistent NCH model fits across time slices within the Pleistocene through Holocene even within the same species, which suggests the dynamics controlling the fundamental niche may be more complex than these abiotic, PCA-based niche estimations can predict. The NCH predicts the highest abundance of each species, as a cumulative set of occurrences in e-space, at the center of the species environmental niche. In these species of shallow marine bivalves, the NCH pattern doesn't appear to be generally present or absent in all species, but rather varies by species, as well as over time. Previous studies have suggested that biotic interactions may play a large role in constraining dispersal patterns, and biotic factors are

purposefully omitted from this ENM-based approach in order to isolate the influence of abiotic environments. However, the variability in results of our NCH tests here may reflect factors, either biotic or sampling/taphonomic, that influence the structure of species distributions in environmental space.

Indeed, there is considerable debate among modern ecologists regarding the generality of NCH patterns. Some studies suggest that life history traits, such as reproductive cycles and larval survival, may have a greater impact on distribution than environmental factors imposing physiological constraints (Dallas and Santini, 2020; Rivadeneira et al., 2010). Factors such as reproductive output variability, local-scale topography and landscape variability, resource availability and population connectivity also have a significant effect on distributional patterns, and prevent clear abundance center patterns from appearing (Dallas and Santini, 2020; Lester et al., 2007). The relationship between population density and spatial distribution has also been characterized in models such as the ideal free distribution model, which predicts a distribution that maximizes individual fitness and prioritizes resource availability rather than abiotic factors (Kacelnik et al., 1992). For example, among the six species chosen for this study, only *Arcinella cornuta* is classified as sessile, and is morphologically very distinct from the other species (see Figure 1), with a large spiny shell. Life mode differences such as these could very well result in different distributional patterns in this species, as is possibly the case in the geographic-space models shown in Figure 4, and which would also influence how this species occupied abiotic environmental space. Geographically, although individual time bins show no strong correlations, the cumulative data for *A. cornuta* contains a very sharp drop in frequency, suggesting almost the entire population is contained very close to the center of its geographic range, with a few extreme

outliers. In e-space, the same species shows a similar steep drop in abundance, suggesting that the narrow distributional range is not controlled by geographic space available, but by a narrow environmental tolerance. These patterns are consistent with a population of organisms that is fully sessile as adults and occupies a small environmental niche and/or is only rarely dispersed to locations that extend the e-space the species might sample.

Beyond functional ecological differences that influence species' occupation of environmental and geographic spaces, biotic interactions may very well play a complex role in altering distribution patterns within species. These interactions could reflect food web dynamics of populations from specific communities or resource limitations (e.g., Roopnarine, 2006). Biotic interactions with humans may also influence distribution dynamics, especially in the Holocene time slice. For example, *Lunarca ovalis* is a commonly harvested commercial clam in coastal Georgia, which could greatly affect its observed distribution, through artificial removal from the ecosystem via overexploitation, or through introduction by humans into areas not naturally inhabited by the species (Župan et al., 2012). Interestingly, *L. ovalis* was the only species for which Spearman correlation tests in cumulative and time series data showed significant correlation in environmental space between frequency of occurrence and distance from the environmental center, for every time bin except the Holocene. This could be a signal of anthropogenic influence, causing modern specimens of *L. ovalis* to appear beyond their natural environmental parameters even as previous time slices showed that the species has sufficient dispersal capability to allow it to quickly track preferred habitat when environments change and equilibrate to an abundant-center distributional pattern in environmental space.

CONCLUSIONS

Here we tested the abundance center and niche center hypotheses of the structure of species in geographic and environmental spaces, respectively (ACH and NCH). Uniquely, we utilize fossil and modern occurrence data for six marine bivalve species occupying the Gulf and Atlantic Coastal Plains to extend our study over geologic timescales (0–2.8 Ma) in order to better characterize the geographic range and environments in which these species survived. Model results indicate there is some evidence for a relationship between species abundance and distance from the center of a species' geographic range (ACH test), but that it is not generalizable beyond each species. Species only show greatest abundance closest to the geographic centerline when the geographic range polygon is edited to avoid deep water where the species are known to be absent, and where the polygon center is estimated as a centerline within that polygon (this pattern is not apparent for the standard convex hull method of reconstructing geographic range). Further, the ACH pattern was primarily observed when species' data was modeled cumulatively across the Pleistocene and Holocene (vs. within individual time slices). This is consistent with previous studies that have found no geographic abundance-centrality pattern in short duration data, such as studies performed using exclusively modern species occurrences (Dallas et al., 2017; Martínez-Meyer et al., 2013; Sagarin and Gaines, 2002).

The niche-centrality hypothesis was tested by plotting occurrences along 3-dimensional environmental principal component axes (i.e., environmental versus geographic range) and Mahalanobis distances to a centroid of a minimum volume ellipsoid were used as a measure of distance from the niche center. A relationship between abundance and niche centrality was variably observed across species, and the relationship was generally stronger

with cumulative data rather than separated by substage. The NCH pattern was even less consistently observed when environmental distance was replaced with a Maxent ecological niche model suitability.

The lack of consistent patterns observed here suggests other factors may influence the abundance structure of species distributions. Biotic factors, including life mode and ecological interactions, may play an important role in the distribution of individuals throughout a geographic or ecological range, and life mode or biotic interactions may be as critical to predicting changes in local abundance as climatic parameters. Ultimately, this study tested abundance-centrality patterns within a highly vetted dataset of shallow marine bivalves, which share many aspects of their functional ecology and have similar environmental tolerances by virtue of inhabiting the same region globally. A more comprehensive study across more and ecologically different taxa would better constrain the generalizability of abundance-centrality patterns within the ecological niches of organisms writ large. Broader studies of modern taxa have revealed taxon-specificity in abundance-centrality patterns, which may suggest there are specific biotic factors or environmental interactions that might influence the strength of an abundance-centrality correlation within any given species.

The study of biogeography is often confronted with issues of scale – and this study attempts to elucidate the scale dependency of the abundance structure of species. Ecological change is an inevitable byproduct of climate change, and predicting the directionality of that ecological change is a central tenet of conservation biology and ecology. Given that bivalve species on average likely last 5-10 million years (Lawton and May, 1995), this study used time-averaged sampling to analyze an ecological principle in its geologic context, with the

goal of better informing a species' biogeographic future by looking to its past. Heading into a regime of extreme climate change, the distributional patterns of marine bivalves may not be fully predictable using purely abiotic approaches, but an expectation that they will shift to maintain a partially abundant center can complement studies of biotic interactions and human exploitation and habitat change that will all experience complex and interdependent changes going forward.

APPENDIX A: ANALYSIS RESULTS TABLES

Table 1: Total unique datapoints in database.

Species	***A. aequalis***	***A. cornuta***	***L. ovalis***	***M. campechiensis***	***P. tridentata***	***S. floridana***
E. Pleistocene	29	78	20	83	56	32
M. Pleistocene	17	34	N/A	29	13	30
L. Pleistocene	29	10	39	59	31	33
Holocene	61	69	29	87	34	N/A
Total	137	191	88	258	135	85

Table 2: Geographic-space linear model R^2 and p-values using convex hull model for geographic range, with distances calculated from polygon centroid.

Species	***A. aequalis***	***A. cornuta***	***L. ovalis***	***M. campechiensis***	***P. tridentata***	***S. floridana***
Cumulative	R^2= 0.004, p=0.88			R^2= 0.12, p= 0.46	R^2=0.13, p=0.28	R^2= 0.02, p= 0.70
E. Pleist	R^2= 0.05, p= 0.66			R^2 = 0.037, p= 0.55	R^2=0.19, p=0.29	R^2= 0.024, p= 0.69
M. Pleist	R^2= 0.01, p= 0.78			R^2= 0.68, p= 0.08	R^2 = 0.12, p=0.57	R^2= 0.005, p= 0.89
L. Pleist	R^2= 0.14, p= 0.46			R^2= 0.07, p= 0.57	R^2=0.17, p=0.42	R^2= 0.07, p=0.54
Holocene	R^2= 0.47, p= 0.09			R^2= 0.17, p= 0.37	R^2=0.31, p=0.16	

Table 3: Geographic-space linear model R^2 and p-values using geographic range polygons edited to avoid deep uninhabitable water, with distances calculated from occurrence point locations to range centerlines.

Species	***A. aequalis***	***A. cornuta***	***L. ovalis***	***M. campechiensis***	***P. tridentata***	***S. floridana***
Cumulative	R^2= 0.73, p= 1.2e-09	R^2=0.25, p=0.02	R^2=0.24, p=0.005	R^2=0.61, p=4.60e-08	R^2=0.41, p=0.0001	R^2=0.34 , p=0.0002
E. Pleist	R^2= 0.01, p= 0.65	R^2=0.30, p=0.01	R^2=0.11, p=0.07	R^2=0.47, p=0.001	R^2=0.33, p=0.003	R^2=0.15 , p=0.10
M. Pleist	R^2= 0.002 , p= 0.83	R^2=0.10, p=0.07		R^2=0.05, p=0.37	R^2=0.13 , p=0.09	R^2= 2.4e-32, p=1
L. Pleist	R^2= 0.30, p= 0.018	R^2=0.03, p=0.46	R^2=0.45, p=0.0004	R^2=0.02, p=0.47	R^2=0.15, p=0.07	R^2=0.06 , p=0.22
Holocene	R^2= 0.49 , p= 9.3e-05	R^2=0.05, p=0.28	R^2=0.29, p=0.002	R^2=0.34, p=0.004	R^2=0.20, p=0.02	R^2=0.27, p=0.02

Table 4: Logarithmic frequency linear model R^2 statistics and p-values of geographic-space distances from centerline.

Species	***A. aequalis***	***A. cornuta***	***L. ovalis***	***M. campechiensis***	***P. tridentata***	***S. floridana***
Cumulative	R^2= 0.31, p= 0.0006	R^2=0.005, p=0.7	R^2=0.10, p=0.18	R^2=0.21, p=0.07	R^2=0.35, p=0.006	R^2=0.52 , p=1.6e-05
E. Pleist	R^2= 0.23, p= 0.01	R^2=0.40, p=0.0002	R^2=0.20, p=0.02	R^2=0.46, p=0.003	R^2=0.21, p=0.02	R^2=0.30 , p=0.004
M. Pleist	R^2= 0.21 , p= 0.01	R^2=0.24, p=0.02		R^2=0.22, p=0.01	R^2=0.16, p=0.04	R^2= 16, p=0.03
L. Pleist	R^2= 0.28, p= 0.002	R^2=0.22, p=0.02	R^2=0.09, p=0.10	R^2=0.38, p=0.0003	R^2=0.14, p=0.04	R^2=0.45, p=0.0001
Holocene	R^2= 0.34 , p= 0.0002	R^2=0.37, p=0.01	R^2=0.16, p=0.11	R^2=0.14, p=0.13	R^2=0.45, p=0.005	R^2=0.16, p=0.07

Table 5: Spearman tests for correlation between geographic space distances from the centerline and frequency of occurrence. Here, ρ indicates the degree of correlation between the variables, regardless of linearity, and the sign of ρ indicates the direction of the correlation (positive or negative relationship) and a p-value of < 0.05 is considered significant.

Species	***A. aequalis***	***A. cornuta***	***L. ovalis***	***M. campechiensis***	***P. tridentata***	***S. floridana***
Cumulative	ρ= -0.86, p= 4.0e-10	ρ= -0.76, p= 0.0001	ρ= -0.91, p=8.1e-8	ρ= -0.80, p=1.7e-8	ρ= -0.70, p=1.1e-5	ρ= -0.62, p= 1.0e-4
E. Pleist	ρ= -0.04, p= 0.8	ρ= -0.54, p= 0.01	ρ= -0.31, p= 0.09	ρ= -0.75, p=0.0003	ρ= -0.68, p=0.0002	ρ= -0.47, p=0.04
M. Pleist	ρ= -0.04, p= 0.8	ρ= -0.3, p= 0.09		ρ= -0.38, p=0.12	ρ= 0.12, p=0.6	ρ= -0.01, p=0.9
L. Pleist	ρ= -0.49, p= 0.04	ρ= 0.25, p= 0.3	ρ= -0.63, p= 0.001	ρ= 0.28, p=0.2	ρ= -0.20, p=0.4	ρ= -0.24, p=0.2
Holocene	ρ= -0.68, p= 0.0002	ρ= -0.41, p= 0.04	ρ= -0.56, p=0.001	ρ= -0.62, p=0.002	ρ= -0.28, p=0.2	ρ= -0.52, p=0.01

Table 6: Linear model R^2 and p-value of environmental space mahalanobis distance frequencies.

Species	***A. aequalis***	***A. cornuta***	***L. ovalis***	***M. campechiensis***	***P. tridentata***	***S. floridana***
Cumulative	R^2= 0.54 p=5.0e-06	R^2=0.56, p=6.4e-5	R^2= 0.56, p= 3.7e-5	R^2=0.45, p=0.0003	R^2=0.30, p=0.0005	R^2=0.37, p=0.003
E. Pleist			R^2= 0.14, p= 0.04			
M. Pleist	R^2= 0.06 p= 0.21					R^2=0.23, p=0.02
L. Pleist	R^2= 7.1e-05, p= 0.96		R^2= 0.28 p= 0.008	R^2= 0.25, p= 0.004	R^2=0.05, p=0.37	R^2=0.05, p=0.34
Holocene	R^2= 0.55, p= 0.0002	R^2=0.32, p=0.003	R^2= 0.32, p= 0.02	R^2= 0.19, p= 0.01		

Table 7: Logarithmic frequency linear model R^2 statistics and p-values of environmental space mahalanobis distances from niche center.

Species	***A. aequalis***	***A. cornuta***	***L. ovalis***	***M. campechiensis***	***P. tridentata***	***S. floridana***
Cumulative	R^2= 0.51 p= 1.9e-05	R^2 = 0.33, p=0.004	R^2= 0.51, p= 0.0001	R^2= 0.52, p=8.9e-05	R^2=0.36, p=0.0001	R^2=0.37, p=0.003
E. Pleist			R^2= 0.14, p=0.04			
M. Pleist	R^2= 0.06, p= 0.21				R^2 = 0.11, p=0.09	R^2=0.23, p=0.02
L. Pleist	R^2=7.1e-05, p= 0.96		R^2= 0.28, p= 0.008	R^2= 0.26, p=0.004	R^2=0.5, p=0.38	R^2=0.06, p=.34
Holocene	R^2=0.55, p=0.0003	R^2 = 0.32, p=0.003	R^2= 0.31, p= 0.02	R^2= 0.19, p= 0.01	R^2 = 0.07, p=0.28	

Table 8: Spearman tests for correlation between environmental space Mahalanobis distances from the niche center vs frequency of occurrence.

Species	***A. aequalis***	***A. cornuta***	***L. ovalis***	***M. campechiensis***	***P. tridentata***	***S. floridana***
Cumulative	ρ= -0.95, p=0.0001	ρ= -0.84, p= 8.8e-07	ρ= -0.82, p=0.0009	ρ= -0.87, p=0.0002	ρ= -0.9, p=0.002	ρ= -0.90, p= 0.0002
E. Pleist			ρ= -0.81, p= 0.02			
M. Pleist	ρ= -0.52, p= 0.28					ρ= -0.37, p=0.41
L. Pleist	ρ= 0, p=1		ρ= -0.81, p= 0.049	ρ= -0.40, p=0.33	ρ= -0.21, p=0.6	ρ= -0.13, p=0.7
Holocene	ρ= -0.63, p= 0.048	ρ= -0.62, p= 0.0008	ρ= 0.8, p=0.13	ρ= -0.61, p=0.11		

Table 9: Linear model R^2 and p-values of environmental-space Mahalanobis distances vs Maxent habitat suitability predictions.

Species	***A. aequalis***	***A. cornuta***	***L. ovalis***	***M. campechiensis***	***P. tridentata***	***S. floridana***
Cumulative	R^2= 0.067 p= 0.007	R^2=0.002, p=0.54	R^2= 0.008, p= 0.40	R^2= 0.04, p=0.02	R^2=0.01, p=0.26	R^2=0.35, p=3.2e-06
E. Pleist			R^2= 0.76, p= 4.6e-07			
M. Pleist	R^2= 0.83, p= 2.7e-07					R^2=0.18, p=0.07
L. Pleist	R^2= 0.56, p= 5.2e-06		R^2= 0.26, p= 0.0008	R^2= 0.13, p=0.01	R^2=0.46, p=2.5e-5	R^2=0.62, p=5.1e-08
Holocene	R^2= 0.05, p= 0.09	R^2=0.12, p=0.003	R^2= 0.23, p= 0.008	R^2= 0.0003, p= 0.88		

Table 10: Spearman tests for correlation between environmental-space Mahalanobis distances vs Maxent habitat suitability predictions.

Species	***A. aequalis***	***A. cornuta***	***L. ovalis***	***M. campechiensis***	***P. tridentata***	***S. floridana***
Cumulative	ρ= 0.14, p=0.14	ρ=0.17, p=0.02	ρ= -0.11, p= 0.30	ρ= 0.27. p= 0.001	ρ=-0.09, p=0.29	ρ=0.51, p=9.8e-05
E. Pleist			ρ= -0.84, p= 2.7e-06			
M. Pleist	ρ= 0.81, p= 6.8e-05					ρ=-0.80, p=2.7e-05
L. Pleist	ρ= 0.50, p= 0.007		ρ= 0.46, p= 0.003	ρ= 0.44, p= 0.0006	ρ=0.73, p=2.9e-06	ρ=0.70, p=6.3e-06
Holocene	ρ= -0.43, p= 0.0005	ρ= -0.43, p=0.0002	ρ= -0.74, p= 4.1e-06	ρ= 0.25, p= 0.02		

Table 11: Variables Used in Environmental Analysis

Variable	**Definition**	**Unit**	**Scaling Factor**
biogeo08	Mean Annual SSS	psu	100×
biogeo09	Minimum Monthly SSS	psu	100×
biogeo10	Maximum Monthly SSS	psu	100×
biogeo11	Annual Range in SSS	psu	100×
biogeo13	Mean Annual SST	degrees C	100×
biogeo14	SST of the coldest ice-free month	degrees C	100×
biogeo15	SST of the warmest ice-free month	degrees C	100×
biogeo16	Annual Range in SST	degrees C	100×

APPENDIX B: MAXENT MODELING

Table 12: Maximum training sensitivity plus specificity (MSS) and Area Under the Curve (AUC) of the Receiver Operating Characteristic (ROC) for each Maxent model. AUC describes the relationship between model sensitivity vs. specificity as a predictor of model performance (e.g., Elith et al., 2006). AUC values over 0.75 are considered reasonably predictive models for fossil data, which are necessarily less robust than modern occurrence data. MSS describes the threshold value above which Maxent models predict that habitat is suitable for the species being modeled.

	A. aequalis	***A. cornuta***	***L. ovalis***	***M. campechiensis***	***P. tridentata***	***S. floridana***
	Holocene					
AUC	0.92	0.87	0.83	0.82	0.60	N/A
MSS	0.27	0.31	0.38	0.58	0.48	N/A
	Late Pleistocene					
AUC	0.95	0.93	0.88	0.88	0.85	0.75
MSS	0.34	0.48	0.27	0.33	0.35	0.49
	Middle Pleistocene					
AUC	0.95	0.92	N/A	0.71	0.80	0.77
MSS	0.39	0.49	N/A	0.47	0.49	0.44
	Early Pleistocene					
AUC	N/A	0.89	0.81	0.80	0.79	0.70
MSS	N/A	0.49	0.55	0.45	0.46	0.51

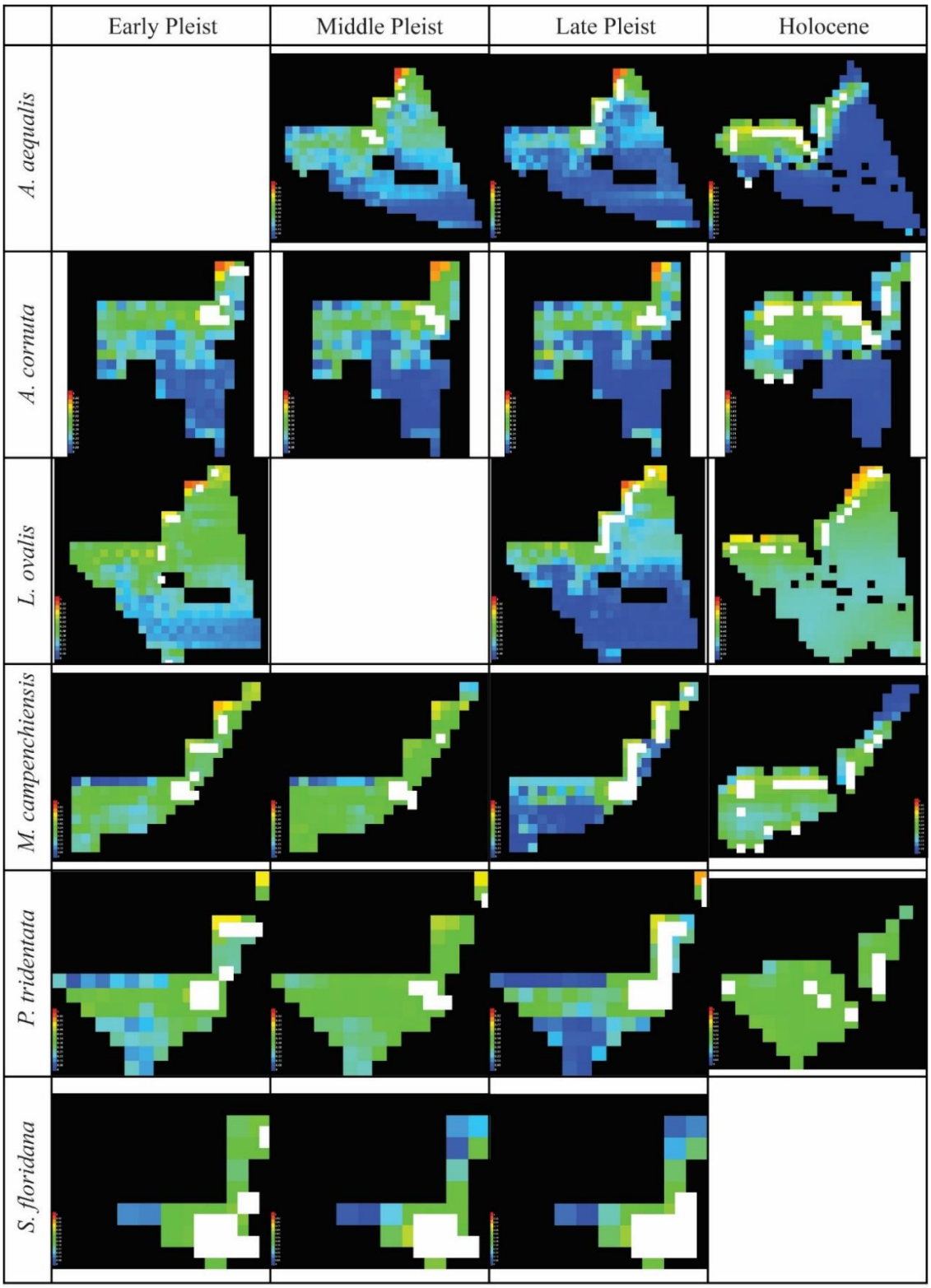

Figure A: Maxent model training maps. Empty boxes indicate insufficient data to create model. Warmer colors indicate regions with more suitable habitat; white squares indicate species occurrence data. MSS and AUC values in Table 11 above.

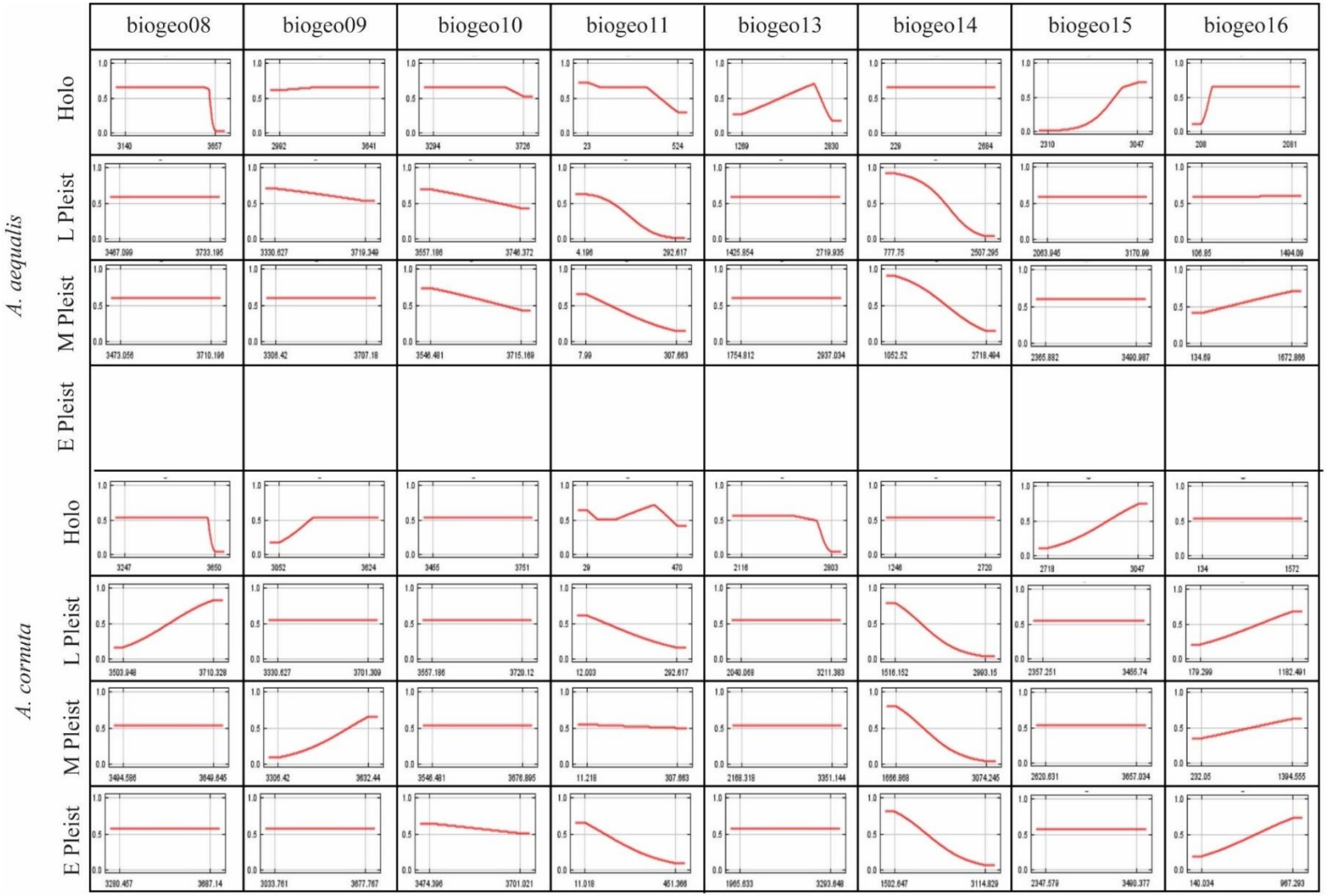

Figure B1: Maxent model response curves. A curve for a well-fit model should appear as a normal distribution curve. Lack of normality indicates that environmental data provided for training the model did not sufficiently sample the range of environments in which the species may survive Saupe et al., 2012; Owens et al., 2013). These models may be interpreted only within the range of environmental values sampled – that is, any model extrapolation in these instances is likely to be artifactual.

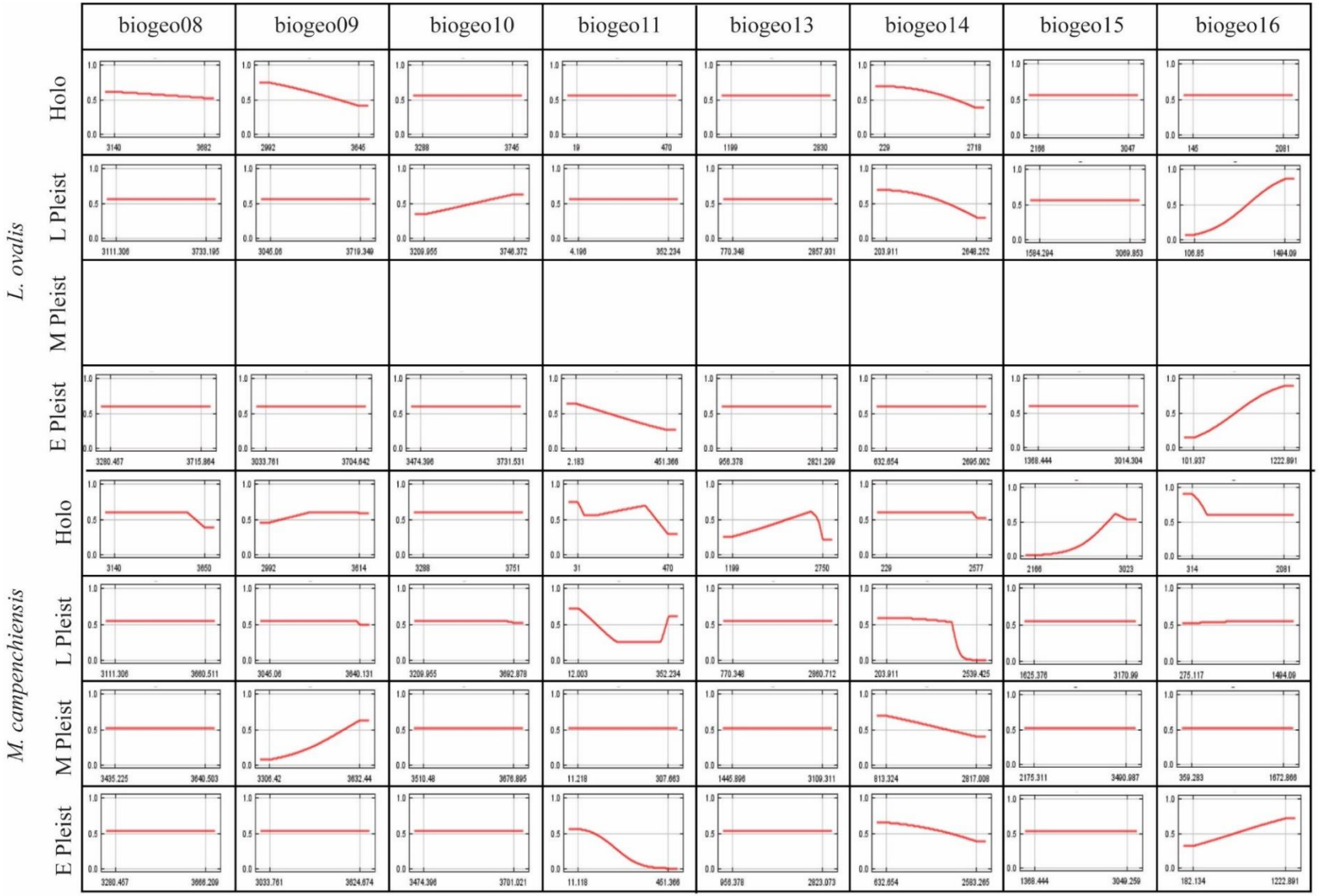

Figure B2: Maxent model response curves. A curve for a well-fit model should appear as a normal distribution curve. See Figure B1 for detailed explanation of figure interpretation.

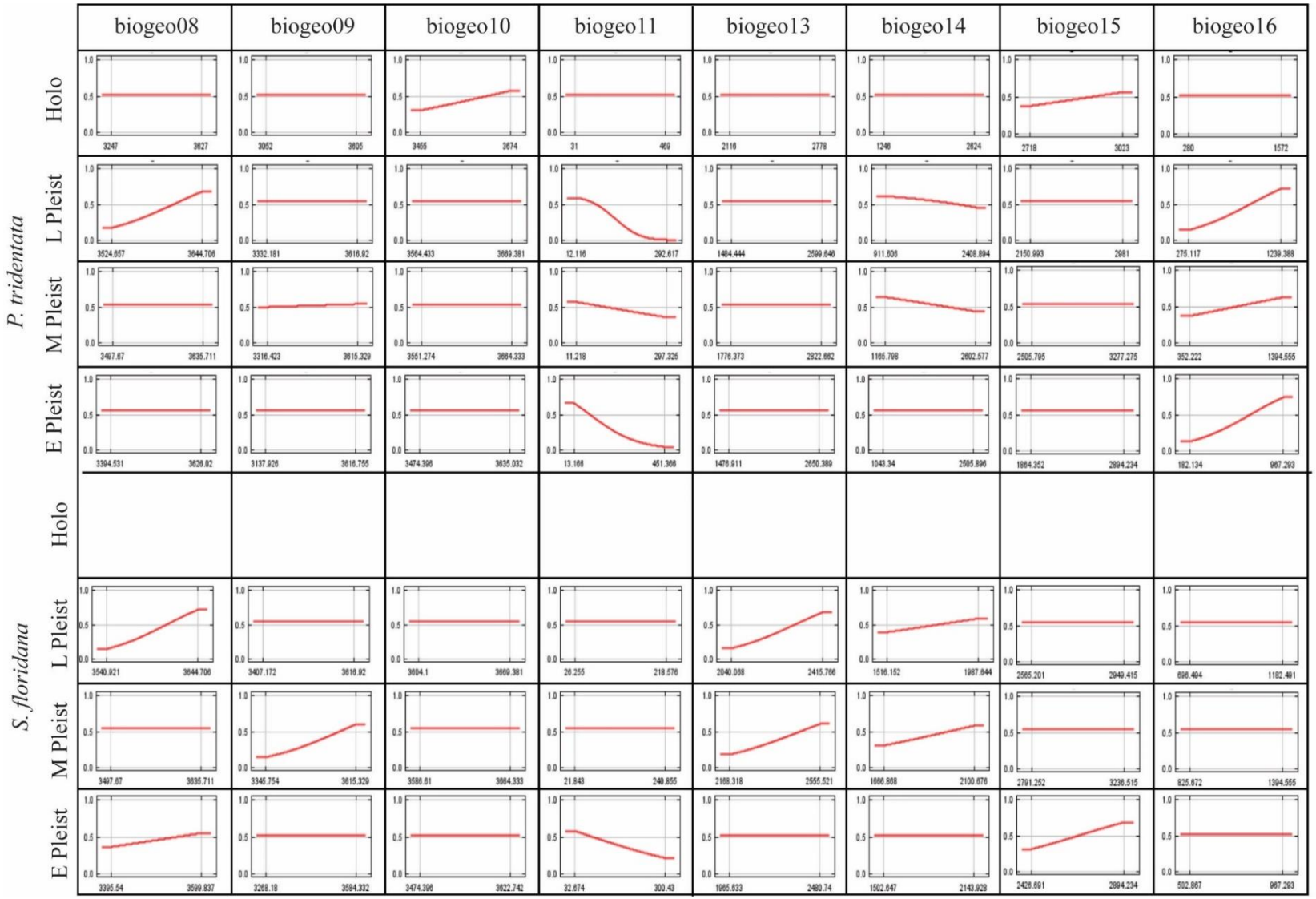

Figure B3: Maxent model response curves. A curve for a well-fit model should appear as a normal distribution curve. See Figure B1 for detailed explanation of figure interpretation.

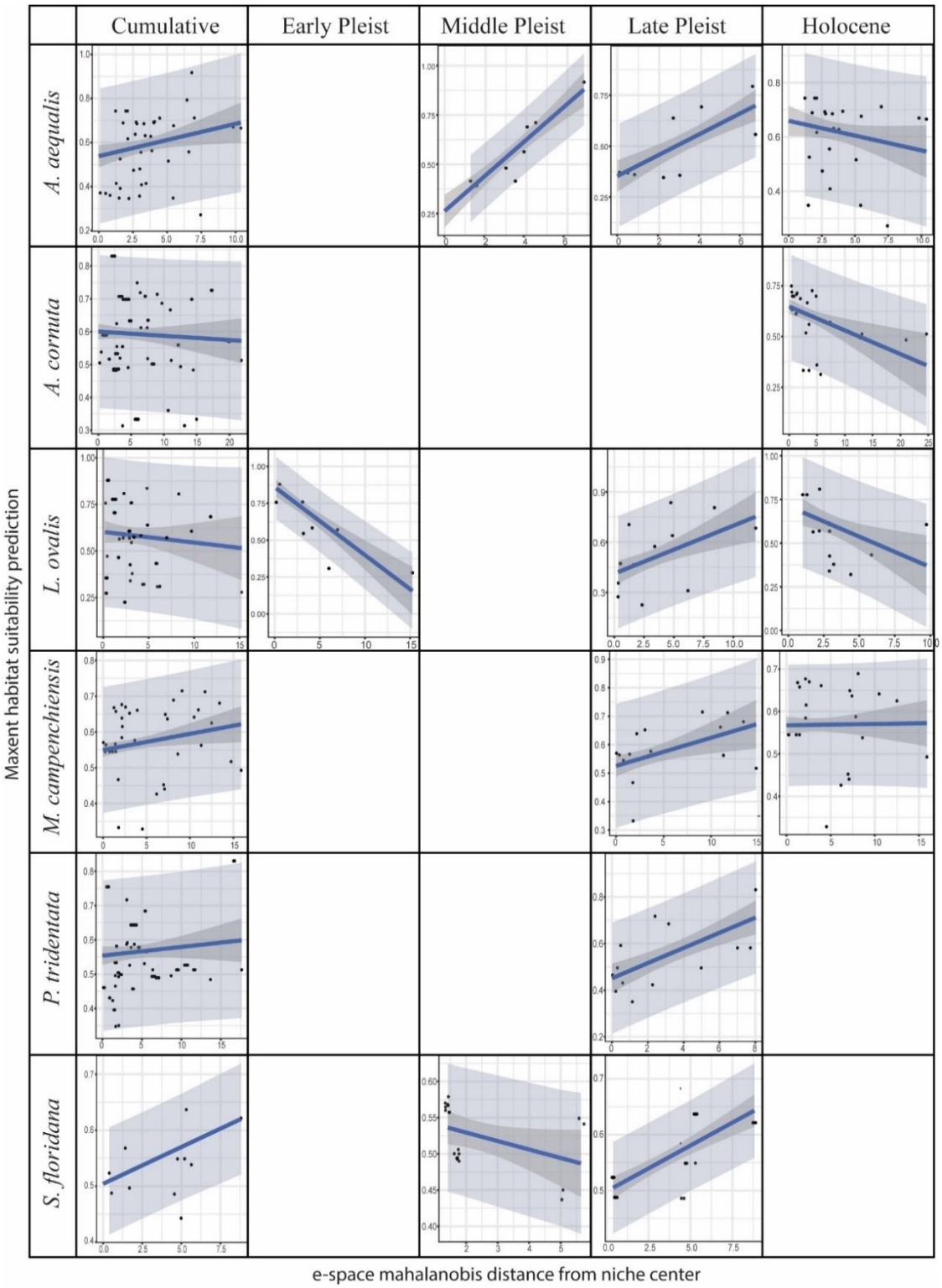

Figure C: Linear model fits to Mahalanobis distances calculated in e-space plotted against logistic probability predictions using Maxent ecological niche models. Blue line shows linear model fit, darker gray band indicates 95% confidence interval for linear model fit, and lighter gray band shows 95% data prediction band.

APPENDIX C: R CODE

````
---
title: "Geographic ACH - Centerline Analysis"
author: "Rhiannon Nolan"
date: "2023-05-10"
output:
  html_document:
  toc: true
---

```{R, echo=FALSE}
#GLOBAL R chunk options.
#   (to hide this message add "echo=FALSE" to the code chunk
options)

knitr::opts_chunk$set(comment = NA, message = FALSE, warning =
FALSE, width = 100)
knitr::opts_chunk$set(fig.align = "center", fig.height = 4,
fig.width = 6)

#knitr::opts_chunk$set(cache = TRUE, autodep=TRUE)
```

```{r}
#run this for the version of the geog range traced in ArcMap, using
distance from a centerline rather than a point in an alpha hull
library(tidyverse)
master <- read_csv("C:/Users/rznol/Dropbox/Rhiannon-Cori/ACH
Datasets/Climate/Geog_Centerlines/master_geogcent.csv")

#choose which species to use this run by deleting leading #
#centln <- subset(master, Taxa == "Abra_aequalis")
#centln <- subset(master, Taxa == "Arcinella_cornuta")
#centln <- subset(master, Taxa == "Lunarca_ovalis")
#centln <- subset(master, Taxa == "Mercenaria_campechiensis")
#centln <- subset(master, Taxa == "Pleuromeris_tridentata")
centln <- subset(master, Taxa == "Stewartia_floridana")

#separate out time bins
centln_all <- subset(centln, Time == "Cumulative")
#get rid of extreme outliers
centln_all <- centln_all %>% filter(Distance < 300)

centln_EPleist <- subset(centln, Time == "Early Pleistocene")
centln_MPleist <- subset(centln, Time == "Middle Pleistocene")
centln_LPleist <- subset(centln, Time == "Late Pleistocene")
centln_Holo <- subset(centln, Time == "Holocene")

#give them the names from the previous iteration
````
````

```
euc_geog <- centln_all
euc_epleist <- centln_EPleist
euc_mpleist <- centln_MPleist
euc_lpleist <- centln_LPleist
euc_holo <- centln_Holo

#create data frames for histograms
euc_geog_df <- data.frame(centln_all)
euc_epleist_df <- data.frame(centln_EPleist)
euc_mpleist_df <- data.frame(centln_MPleist)
euc_lpleist_df <- data.frame(centln_LPleist)
euc_holo_df <- data.frame(centln_Holo)
```

```{r}
#create objects that are the histogram frequencies as values
hist_euc_geog <- hist(euc_geog$Distance, freq=TRUE, breaks = 25)
hist_euc_epleist <- hist(euc_epleist$Distance, freq=TRUE, breaks =
25)
hist_euc_mpleist <- hist(euc_mpleist$Distance, freq=TRUE, breaks =
25)
hist_euc_lpleist <- hist(euc_lpleist$Distance, freq=TRUE, breaks =
25)
hist_euc_holo <- hist(euc_holo$Distance, freq=TRUE, breaks = 25)
```

```{r}
#linear models of histograms
#cumulative data
library(dbplyr)
euc_freq_geog <- data.frame(hist_euc_geog$counts)
euc_freq_geog <-euc_freq_geog %>%  mutate(id = 1:n())
str(euc_freq_geog)
lm_geog <- lm(hist_euc_geog.counts ~ id, data = euc_freq_geog)

#epleist
library(dbplyr)
euc_freq_epleist <- data.frame(hist_euc_epleist$counts)
euc_freq_epleist <-euc_freq_epleist %>%  mutate(id = 1:n())
str(euc_freq_epleist)
lm_epleist_geog <- lm(hist_euc_epleist.counts ~ id, data =
euc_freq_epleist)

#mpleist
library(dbplyr)
euc_freq_mpleist <- data.frame(hist_euc_mpleist$counts)
euc_freq_mpleist <-euc_freq_mpleist %>%  mutate(id = 1:n())
str(euc_freq_mpleist)
lm_mpleist_geog <- lm(hist_euc_mpleist.counts ~ id, data =
euc_freq_mpleist)
```

```
#lpleist
library(dbplyr)
euc_freq_lpleist <- data.frame(hist_euc_lpleist$counts)
euc_freq_lpleist <-euc_freq_lpleist %>%  mutate(id = 1:n())
str(euc_freq_lpleist)
lm_lpleist_geog <- lm(hist_euc_lpleist.counts ~ id, data =
euc_freq_lpleist)

#holo
library(dbplyr)
euc_freq_holo <- data.frame(hist_euc_holo$counts)
euc_freq_holo <-euc_freq_holo %>%  mutate(id = 1:n())
str(euc_freq_holo)
lm_holo_geog <- lm(hist_euc_holo.counts ~ id, data = euc_freq_holo)
```

```{r}
summary(lm_geog)
summary(lm_epleist_geog)
summary(lm_mpleist_geog)
summary(lm_lpleist_geog)
summary(lm_holo_geog)
```

Spearman Tests

```{r}
#spearman tests for environmental data - freq vs mahal
#this tests for correlation between variables, with result values
between {-1,1}, where a 0 indicates no correlation
cumul_mahal <- as.numeric(euc_freq_geog$hist_euc_geog.counts)
cumul_freq <- as.numeric(euc_freq_geog$id)
corr_cumulat_freq <- cor.test(cumul_freq, cumul_mahal, data =
euc_freq_geog , method = 'spearman')
corr_cumulat_freq

#time bins enviro spearman tests

#epleist
epleist_mahal2 <-
as.numeric(euc_freq_epleist$hist_euc_epleist.counts)
epleist_freq <- as.numeric(euc_freq_epleist$id)
corr_epleist_freq <- cor.test(epleist_freq, epleist_mahal2, data =
euc_freq_epleist , method = 'spearman')
corr_epleist_freq

#mpleist
mpleist_mahal2 <-
as.numeric(euc_freq_mpleist$hist_euc_mpleist.counts)
```

```
mpleist_freq <- as.numeric(euc_freq_mpleist$id)
corr_mpleist_freq <- cor.test(mpleist_freq, mpleist_mahal2, data =
euc_freq_mpleist , method = 'spearman')
corr_mpleist_freq

#lpleist
lpleist_mahal2 <-
as.numeric(euc_freq_lpleist$hist_euc_lpleist.counts)
lpleist_freq <- as.numeric(euc_freq_lpleist$id)
corr_lpleist_freq <- cor.test(lpleist_freq, lpleist_mahal2, data =
euc_freq_lpleist , method = 'spearman')
corr_lpleist_freq

#holo
holo_mahal2 <- as.numeric(euc_freq_holo$hist_euc_holo.counts)
holo_freq <- as.numeric(euc_freq_holo$id)
corr_holo_freq <- cor.test(holo_freq, holo_mahal2, data =
euc_freq_holo , method = 'spearman')
corr_holo_freq
```

```
# Log Transform

```{r, eval=FALSE}
#run this for the version of the geog range traced in ArcMap, using
distance from a centerline rather than a point in an alpha hull
library(tidyverse)
master <- read_csv("C:/Users/rznol/Dropbox/Rhiannon-Cori/ACH
Datasets/Climate/Geog_Centerlines/master_geogcent.csv")

#choose which species to use this run
#centln <- subset(master, Taxa == "Abra_aequalis")
#centln <- subset(master, Taxa == "Arcinella_cornuta")
centln <- subset(master, Taxa == "Lunarca_ovalis")
#centln <- subset(master, Taxa == "Mercenaria_campechiensis")
#centln <- subset(master, Taxa == "Pleuromeris_tridentata")
#centln <- subset(master, Taxa == "Stewartia_floridana")

#separate out time bins
centln_all <- subset(centln, Time == "Cumulative")
centln_EPleist <- subset(centln, Time == "Early Pleistocene")
centln_MPleist <- subset(centln, Time == "Middle Pleistocene")
centln_LPleist <- subset(centln, Time == "Late Pleistocene")
centln_Holo <- subset(centln, Time == "Holocene")

#give them the names from the previous iteration
euc_geog <- centln_all
euc_epleist <- centln_EPleist
euc_mpleist <- centln_MPleist
euc_lpleist <- centln_LPleist
```
```

```
euc_holo <- centln_Holo

#create data frames for histograms
euc_geog_df <- data.frame(centln_all)
euc_epleist_df <- data.frame(centln_EPleist)
euc_mpleist_df <- data.frame(centln_MPleist)
euc_lpleist_df <- data.frame(centln_LPleist)
euc_holo_df <- data.frame(centln_Holo)
```

```{r}
#create objects that are the histogram frequencies as values - log
value
hist_log_geog <- hist(euc_geog$Log, freq=TRUE, breaks = 25)
hist_log_epleist <- hist(euc_epleist$Log, freq=TRUE, breaks = 25)
hist_log_mpleist <- hist(euc_mpleist$Log, freq=TRUE, breaks = 25)
hist_log_lpleist <- hist(euc_lpleist$Log, freq=TRUE, breaks = 25)
hist_log_holo <- hist(euc_holo$Log, freq=TRUE, breaks = 25)
```

```{r}
#linear models of histograms
#cumulative data
library(dbplyr)
log_freq_geog <- data.frame(hist_log_geog$counts)
log_freq_geog <-log_freq_geog %>%  mutate(id = 1:n())
str(log_freq_geog)
lm_geog_log <- lm(hist_log_geog.counts ~ id, data = log_freq_geog)

#epleist
library(dbplyr)
log_freq_epleist <- data.frame(hist_log_epleist$counts)
log_freq_epleist <-log_freq_epleist %>%  mutate(id = 1:n())
str(log_freq_epleist)
lm_epleist_geog_log <- lm(hist_log_epleist.counts ~ id, data =
log_freq_epleist)

#mpleist
library(dbplyr)
log_freq_mpleist <- data.frame(hist_log_mpleist$counts)
log_freq_mpleist <-log_freq_mpleist %>%  mutate(id = 1:n())
str(log_freq_mpleist)
lm_mpleist_geog_log <- lm(hist_log_mpleist.counts ~ id, data =
log_freq_mpleist)

#lpleist
library(dbplyr)
log_freq_lpleist <- data.frame(hist_log_lpleist$counts)
log_freq_lpleist <-log_freq_lpleist %>%  mutate(id = 1:n())
str(log_freq_lpleist)
```

```
lm_lpleist_geog_log <- lm(hist_log_lpleist.counts ~ id, data =
log_freq_lpleist)

#holo
library(dbplyr)
log_freq_holo <- data.frame(hist_log_holo$counts)
log_freq_holo <-log_freq_holo %>%  mutate(id = 1:n())
str(log_freq_holo)
lm_holo_geog_log <- lm(hist_log_holo.counts ~ id, data =
log_freq_holo)
```

```{r}
summary(lm_geog_log)
summary(lm_epleist_geog_log)
summary(lm_mpleist_geog_log)
summary(lm_lpleist_geog_log)
summary(lm_holo_geog_log)
```

````
---
title: "Environmental NCH Analysis"
author: "Rhiannon Nolan"
date: "April 02, 2023"
output:
  html_document:
  toc: true
---

```{R, echo=FALSE}
#GLOBAL R chunk options.
#   (to hide this message add "echo=FALSE" to the code chunk
options)

knitr::opts_chunk$set(comment = NA, message = FALSE, warning =
FALSE, width = 100)
knitr::opts_chunk$set(fig.align = "center", fig.height = 4,
fig.width = 6)

#knitr::opts_chunk$set(cache = TRUE, autodep=TRUE)
```

### Data Conversion

```{R}
library(tidyverse)

#load in env csv file with lat, long, time bins,environmental data,
and logistic predictions from MAXENT
env <- read_csv("C:/Users/rznol/Dropbox/Rhiannon-Cori/ACH Datasets/R
Files/occ_near_csv/full_env/env_all.csv")

#change names of Time bins for easier reading of figures
env <-  env %>%  mutate(    Time =
      fct_collapse(
        Time
      , "E Pleist" = c("Early Pleistocene")
      , "M Pleist" = c("Middle Pleistocene")
      , "L Pleist" = c("Late Pleistocene")
      , "Holo" = c("Holocene")
      )
  )

#set order of time bins to chronological
env$Time <-  factor(    env$Time
  , levels = c( "E Pleist"
                , "M Pleist"
                , "L Pleist"
                , "Holo"
     )
````
````

```
  )

```

```{r}
#separate out which species you want to run
library(tidyverse)

#choose which species to use this run by deleting the # in front of
it
env_sp <- subset(env, species == "Abra_aequalis")
#env_sp <- subset(env, species == "Arcinella_cornuta")
#env_sp <- subset(env, species == "Lunarca_ovalis")
#env_sp <- subset(env, species == "Mercenaria_campechiensis")
#env_sp <- subset(env, species == "Pleuromeris_tridentata")
#env_sp <- subset(env, species == "Stewartia_floridana")

#separate out time bins
env_sp_all <- env_sp
env_sp_EPleist <- subset(env_sp, Time == "E Pleist")
env_sp_MPleist <- subset(env_sp, Time == "M Pleist")
env_sp_LPleist <- subset(env_sp, Time == "L Pleist")
env_sp_Holo <- subset(env_sp, Time == "Holo")
```

Environmental Space

```{R, fig.height = 8, fig.width = 8}
# Scatterplot matrix of variables to check for colinearity
library(ggplot2)
library(GGally)

#columns to use for analysis - check using names(env) to find which
columns are environmental variables
use_col_ind <- c(7:14)
use_col_names <- names(env_sp)[use_col_ind]
use_col_names

#scatterplot - takes a while to plot so just run when needed

#p <-
#  ggpairs(
#    data = env,
#    columns = use_col_names
#  )
#print(p)
```

```{R, fig.height = 5, fig.width = 8}
#create PCA
```

```
pca <-
  princomp(
    env_sp[, use_col_ind]
  , cor = TRUE
  )
summary(pca)

#show loadings to find what each PC axis formula is being used
pca %>% loadings() %>% print(cutoff = 0.2) # cutoff = 0 to show all
values

```

```{R}
par(mfrow=c(1,2))
#histogram showing percent of variance explained by each component
screeplot(pca)
#vector plot of varianbles and data, plotted in the first two
components PC-space
biplot(pca)
par(mfrow=c(1,1))
```

```{r}
# create data frame of PC values
pca_coord <- data.frame(pca$scores)
df_env <- data.frame(pca_coord$Comp.1, pca_coord$Comp.2,
pca_coord$Comp.3)
df_env <-df_env %>%  mutate(id = 1:n())
```

Create Separate PCs for each Time Bin, using only environmental data from that time bin

```{r}
library(MASS)

#time bins using enviro data

pca_EPleist <- princomp(env_ep_df[, use_col_ind], cor = TRUE)
pca_MPleist <- princomp(env_mp_df[, use_col_ind], cor = TRUE)
pca_LPleist <- princomp(env_lp_df[, use_col_ind], cor = TRUE)
pca_Holo <- princomp(env_h_df[, use_col_ind], cor = TRUE)

#create data frames of first 3 components, convert to matrix

pca12 <- data.frame(pca_EPleist$scores)
df12 <- data.frame(pca12$Comp.1, pca12$Comp.2, pca12$Comp.3)

pca13 <- data.frame(pca_MPleist$scores)
```

```
df13 <- data.frame(pca13$Comp.1, pca13$Comp.2, pca13$Comp.3)

pca14 <- data.frame(pca_LPleist$scores)
df14 <- data.frame(pca14$Comp.1, pca14$Comp.2, pca14$Comp.3)

pca15 <- data.frame(pca_Holo$scores)
df15 <- data.frame(pca15$Comp.1, pca15$Comp.2, pca15$Comp.3)
```

Centroids and Mahalanobis Distances in E-Space

```{r}
# cumulative range centroid
library(MASS)
mve_env <- cov.rob(df_env)

# Time bin range centroids
library(MASS)
env_mve_epleist <- cov.mve(df12)
env_mve_mpleist <- cov.mve(df13)
env_mve_lpleist <- cov.mve(df14)
env_mve_holo <- cov.mve(df15)

# mahalanobis distances
#cumulative
mahal_env <- mahalanobis(df_env, mve_env$center, cov(df_env))

#EPleist
env_mahal_epleist <- mahalanobis(df12, env_mve_epleist$center,
cov(df12))
#MPleist
env_mahal_mpleist <- mahalanobis(df13, env_mve_mpleist$center,
cov(df13))
#LPleist
env_mahal_lpleist <- mahalanobis(df14, env_mve_lpleist$center,
cov(df14))
#Holo
env_mahal_holo <- mahalanobis(df15, env_mve_holo$center, cov(df15))
```

```{r}
#create objects that are the histogram frequencies as values
env_hist_mahal <- hist(mahal_env, breaks = 25)
env_hist_mahal_epleist <- hist(env_mahal_epleist, breaks = 25)
env_hist_mahal_mpleist <- hist(env_mahal_mpleist, breaks = 25)
env_hist_mahal_lpleist <- hist(env_mahal_lpleist, breaks = 25)
env_hist_mahal_holo <- hist(env_mahal_holo, breaks = 25)
```

Linear Modeling

```{r}
#linear models of histograms
#cumulative
library(dbplyr)
env_mahal_freq <- data.frame(env_hist_mahal$counts)
env_mahal_freq <-env_mahal_freq %>%  mutate(id = 1:n())
str(env_mahal_freq)
lm_env <- lm(env_hist_mahal.counts ~ id, data = env_mahal_freq)
```

```{r}
#linear models of time bin environmental data - mahalanobis
distances vs frequencies

#epleist
library(dbplyr)
env_mahal_freq_epleist <- data.frame(env_hist_mahal_epleist$counts)
env_mahal_freq_epleist <-env_mahal_freq_epleist %>%  mutate(id =
1:n())
str(env_mahal_freq_epleist)
lm_epleist_env <- lm(env_hist_mahal_epleist.counts ~ id, data =
env_mahal_freq_epleist)

#mpleist
library(dbplyr)
env_mahal_freq_mpleist <- data.frame(env_hist_mahal_mpleist$counts)
env_mahal_freq_mpleist <-env_mahal_freq_mpleist %>%  mutate(id =
1:n())
str(env_mahal_freq_mpleist)
lm_mpleist_env <- lm(env_hist_mahal_mpleist.counts ~ id, data =
env_mahal_freq_mpleist)

#lpleist
library(dbplyr)
env_mahal_freq_lpleist <- data.frame(env_hist_mahal_lpleist$counts)
env_mahal_freq_lpleist <-env_mahal_freq_lpleist %>%  mutate(id =
1:n())
str(env_mahal_freq_lpleist)
lm_lpleist_env <- lm(env_hist_mahal_lpleist.counts ~ id, data =
env_mahal_freq_lpleist)

#holo
library(dbplyr)
env_mahal_freq_holo <- data.frame(env_hist_mahal_holo$counts)
env_mahal_freq_holo <-env_mahal_freq_holo %>%  mutate(id = 1:n())
str(env_mahal_freq_holo)
lm_holo_env <- lm(env_hist_mahal_holo.counts ~ id, data =
env_mahal_freq_holo)
```

```{R}
#summaries of linear models using mahal dist and frequency
summary(lm_env)
summary(lm_epleist_env)
summary(lm_mpleist_env)
summary(lm_lpleist_env)
summary(lm_holo_env)
```

```{r}
#add col that contains mahal dists
env_sp$mahal <- mahal_env
env_sp_EPleist$mahal <- env_mahal_epleist
env_sp_MPleist$mahal <- env_mahal_mpleist
env_sp_LPleist$mahal <- env_mahal_lpleist
env_sp_Holo$mahal <- env_mahal_holo

#add all the working time bins back into one dataset
env_sp_2 <- rbind(
  env_sp_EPleist,
  env_sp_MPleist,
  env_sp_LPleist,
  env_sp_Holo
  )
```

Spearman Tests

```{r}
#spearman tests for environmental data - freq vs mahal
#this tests for correlation between variables, with result values
between {-1,1}, where a 0 indicates no correlation
cumul_mahal <- as.numeric(env_mahal_freq$env_hist_mahal.counts)
cumul_freq <- as.numeric(env_mahal_freq$id)
corr_cumulat_freq <- cor.test(cumul_freq, cumul_mahal2, data =
env_mahal_freq , method = 'spearman')
corr_cumulat_freq

#time bins enviro spearman tests

#epleist
epleist_mahal2 <-
as.numeric(env_mahal_freq_epleist$env_hist_mahal_epleist.counts)
epleist_freq <- as.numeric(env_mahal_freq_epleist$id)
corr_epleist_freq <- cor.test(epleist_freq, epleist_mahal2, data =
env_EPleist , method = 'spearman')
corr_epleist_freq

#mpleist
```

```
mpleist_mahal2 <-
as.numeric(env_mahal_freq_mpleist$env_hist_mahal_mpleist.counts)
mpleist_freq <- as.numeric(env_mahal_freq_mpleist$id)
corr_mpleist_freq <- cor.test(mpleist_freq, mpleist_mahal2, data =
env_MPleist , method = 'spearman')
corr_mpleist_freq

#lpleist
lpleist_mahal2 <-
as.numeric(env_mahal_freq_lpleist$env_hist_mahal_lpleist.counts)
lpleist_freq <- as.numeric(env_mahal_freq_lpleist$id)
corr_lpleist_freq <- cor.test(lpleist_freq, lpleist_mahal2, data =
env_Lpleist , method = 'spearman')
corr_lpleist_freq

#holo
holo_mahal2 <-
as.numeric(env_mahal_freq_holo$env_hist_mahal_holo.counts)
holo_freq <- as.numeric(env_mahal_freq_holo$id)
corr_holo_freq <- cor.test(holo_freq, holo_mahal2, data = env_holo ,
method = 'spearman')
corr_holo_freq
```

Log Scale Linear Models

```{r}
#create objects that are the histogram frequencies as values
env_hist_mahal <- hist(mahal_env, breaks = 25)
env_hist_mahal_epleist <- hist(env_mahal_epleist, breaks = 25)
env_hist_mahal_mpleist <- hist(env_mahal_mpleist, breaks = 25)
env_hist_mahal_lpleist <- hist(env_mahal_lpleist, breaks = 25)
env_hist_mahal_holo <- hist(env_mahal_holo, breaks = 25)
```

```{r}
library(dbplyr)
log_env <- data.frame(env_hist_mahal$counts)
log_env <- log(abs(log_env))

log_env_ep <- data.frame(env_hist_mahal_epleist$counts)
log_env_ep <- log(abs(log_env_ep))

log_env_mp <- data.frame(env_hist_mahal_mpleist$counts)
log_env_mp <- log(abs(log_env_mp))

log_env_lp <- data.frame(env_hist_mahal_lpleist$counts)
log_env_lp <- log(abs(log_env_lp))

log_env_h <- data.frame(env_hist_mahal_holo$counts)
```

```
log_env_h <- log(abs(log_env_h))
```

```{r}
#linear models of histograms
#cumulative data
library(dbplyr)

#log_env <- log_env %>% filter(log_env >= 0)
log_env <-log_env %>%  mutate(id = 1:n())
lm_log <- lm(env_hist_mahal$counts ~ id, data = log_env)

log_env_ep <-log_env_ep %>%  mutate(id = 1:n())
log_env_ep <-log_env_ep %>%  mutate(id = 1:n())
lm_log_ep <- lm(env_hist_mahal_epleist$counts ~ id, data =
log_env_ep)

log_env_mp <-log_env_mp %>%  mutate(id = 1:n())
lm_log_mp <- lm(env_hist_mahal_mpleist$counts ~ id, data =
log_env_mp)

log_env_lp <-log_env_lp %>%  mutate(id = 1:n())
lm_log_lp <- lm(env_hist_mahal_lpleist$counts ~ id, data =
log_env_lp)

log_env_h <-log_env_h %>%  mutate(id = 1:n())
lm_log_h <- lm(env_hist_mahal_holo$counts ~ id, data = log_env_h)
```

```{r}
summary(lm_log)
summary(lm_log_ep)
summary(lm_log_mp)
summary(lm_log_lp)
summary(lm_log_h)
```

REFERENCES

Adhikari, D., Mir, A.H., Upadhaya, K., Iralu, V., Roy, D.K., 2018. Abundance and habitat-suitability relationship deteriorate in fragmented forest landscapes: a case of Adinandra griffithii Dyer, a threatened endemic tree from Meghalaya in northeast India. Ecol Process 7, 3. https://doi.org/10.1186/s13717-018-0114-z

Allmon, W.D., Rosenberg, G., Portell, R.W., Schindler, K.S., 1993. Diversity of Atlantic Coastal Plain Mollusks Since the Pliocene. Science 260, 1626–1629. https://doi.org/10.1126/science.260.5114.1626

Allouche, O., Tsoar, A., Kadmon, R., 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). Journal of Applied Ecology 43, 1223–1232. https://doi.org/10.1111/j.1365-2664.2006.01214.x

Anderson, L.C., Geary, D.H., Nehm, R.H., Allmon, W.D., 1991. A comparative study of naticid gastropod predation on Varicorbula caloosae and Chione cancellata, Plio-Pleistocene of Florida, U.S.A. Palaeogeography, Palaeoclimatology, Palaeoecology 85, 29–46. https://doi.org/10.1016/0031-0182(91)90024-L

Attorre, F., De Sanctis, M., Farcomeni, A., Guillet, A., Scepi, E., Vitale, M., Pella, F., Fasola, M., 2013. The use of spatial ecological modelling as a tool for improving the assessment of geographic range size of threatened species. Journal for Nature Conservation 21, 48–55. https://doi.org/10.1016/j.jnc.2012.10.001

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

Behrensmeyer, A.K., Fürsich, F.T., Gastaldo, R.A., Kidwell, S.M., Kosnik, M.A., Kowalewski, M., Plotnick, R.E., Rogers, R.R., Alroy, J., 2005. Are the most durable shelly taxa also the most common in the marine fossil record? Paleobiology 31, 607–623. https://doi.org/10.1666/0094-8373(2005)031[0607:ATMDST]2.0.CO;2

Behrensmeyer, A.K., Kidwell, S.M., Gastaldo, R.A., 2000. Taphonomy and paleobiology. The Paleontological Society Papers.

Bird, D.E., Burke, K., Hall, S.A., Casey, J.F., 2005. Gulf of Mexico tectonic history: Hotspot tracks, crustal boundaries, and early salt distribution. AAPG Bulletin 89, 311–328. https://doi.org/10.1306/10280404026

Blackwelder, B.W., 1981. Late Cenozoic marine deposition in the United States Atlantic coastal plain related to tectonism and global climate, Palaeogeography, Palaeoclimatology, Palaeoecology.

Blois, J.L., Zarnetske, P.L., Fitzpatrick, M.C., Finnegan, S., 2013. Climate Change and the Past, Present, and Future of Biotic Interactions. Science 341, 499–504. https://doi.org/10.1126/science.1237184

Brame, H.-M.R., Stigall, A.L., 2013. Controls on niche stability in geologic time: congruent responses to biotic and abiotic environmental changes among Cincinnatian (Late Ordovician) marine invertebrates. Paleobiology 40, 70–90. https://doi.org/10.1666/13035

Brokaw, R.J., Subrahmanyam, B., Morey, S.L., 2019. Loop Current and Eddy-Driven Salinity Variability in the Gulf of Mexico. Geophysical Research Letters 46, 5978–5986. https://doi.org/10.1029/2019GL082931

Brown, J.H., 1995. Macroecology. University of Chicago Press, Chicago.

Brown, J.H., 1984. On the Relationship between Abundance and Distribution of Species. The American Naturalist 124, 255–279.

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

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

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

Davis, J.L., Mitrovica, J.X., 1996. Glacial isostatic adjustment and the anomalous tide gauge record of eastern North America. Nature 379, 331–333. https://doi.org/10.1038/379331a0

Di Marco, M., Pacifici, M., Maiorano, L., Rondinini, C., 2021. Drivers of change in the realised climatic niche of terrestrial mammals. Ecography 44, 1180–1190. https://doi.org/10.1111/ecog.05414

Elith, J., H. Graham*, C., P. Anderson, R., Dudík, M., Ferrier, S., Guisan, A., J. Hijmans, R., Huettmann, F., R. Leathwick, J., Lehmann, A., Li, J., G. Lohmann, L., A. Loiselle, B., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., McC. M. Overton, J., Townsend Peterson, A., J. Phillips, S., Richardson, K., Scachetti-Pereira, R., E. Schapire, R., Soberón, J., Williams, S., S. Wisz, M., E. Zimmermann, N., 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29, 129–151. https://doi.org/10.1111/j.2006.0906-7590.04596.x

Elith, J., Phillips, S.J., Hastie, T., Dudík, M., Chee, Y.E., Yates, C.J., 2011a. A statistical explanation of MaxEnt for ecologists: Statistical explanation of MaxEnt. Diversity and Distributions 17, 43–57. https://doi.org/10.1111/j.1472-4642.2010.00725.x

Elith, J., Phillips, S.J., Hastie, T., Dudík, M., Chee, Y.E., Yates, C.J., 2011b. A statistical explanation of MaxEnt for ecologists. Diversity and Distributions 17, 43–57. https://doi.org/10.1111/j.1472-4642.2010.00725.x

Fairbanks, R.G., 1989. A 17,OOO-year glacio-eustatic sea level record: influence of glacial melting rates on the Younger Dryas event and deep-ocean circulation.

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

Fürsich, F.T., Aberhan, M., 1990. Significance of time-averaging for palaeocommunity analysis. Lethaia 23, 143–152. https://doi.org/10.1111/j.1502-3931.1990.tb01355.x

Gibbard, P.L., Head, M.J., Walker, M.J.C., the Subcommission on Quaternary Stratigraphy, 2010. Formal ratification of the Quaternary System/Period and the Pleistocene Series/Epoch with a base at 2.58 Ma. J. Quaternary Sci. 25, 96–102. https://doi.org/10.1002/jqs.1338

Grams, T.E.E., Andersen, C.P., 2007. Competition for Resources in Trees: Physiological Versus Morphological Plasticity. Progress in Botany 68, 356–381.

Guisan, A., Zimmermann, N.E., 2000. Predictive habitat distribution models in ecology. Ecological Modelling 135, 147–186. https://doi.org/10.1016/S0304-3800(00)00354-9

Hübscher, C., Nürnberg, D., 2023. Loop Current attenuation after the Mid-Pleistocene Transition contributes to Northern hemisphere cooling. Marine Geology 456, 106976. https://doi.org/10.1016/j.margeo.2022.106976

Hunt, G., Roy, K., Jablonski, D., 2005. Species-Level Heritability Reaffirmed: A Comment on “On the Heritability of Geographic Range Sizes.” The American Naturalist 166, 129–135. https://doi.org/10.1086/430722

Huntley, B., Allen, J.R.M., Forrest, M., Hickler, T., Ohlemüller, R., Singarayer, J.S., Valdes, P.J., 2023. Global biome patterns of the Middle and Late Pleistocene. Journal of Biogeography.

Hutchinson, G.E., 1957. Concluding Remarks. Cold Spring Harbor Symposia on Quantitative Biology 22, 415–427. https://doi.org/10.1101/sqb.1957.022.01.039

Jablonski, D., 1986. Background and Mass Extinctions: The Alternation of Macroevolutionary Regimes.

Jin, P.Y., Sun, J.T., Chen, L., Xue, X.F., Hong, X.Y., 2020. Geography alone cannot explain Tetranychus truncatus (Acari: Tetranychidae) population abundance and genetic diversity in the context of the center–periphery hypothesis. Heredity 124, 383–396. https://doi.org/10.1038/s41437-019-0280-5

Johnston, T.H., 1924. The relation of climate to the spread of prickly pear. Transactions of the Royal Society of South Australia 48, 269–295.

Kacelnik, A., Krebs, J.R., Bernstein, C., 1992. The ideal free distribution and predator-prey populations. Trends in Ecology & Evolution 7, 50–55. https://doi.org/10.1016/0169-5347(92)90106-L

Kidwell, S.M., 2005. Shell Composition Has No Net Impact on Large-Scale Evolutionary Patterns in Mollusks. Science 307, 914–917. https://doi.org/10.1126/science.1106654

Kidwell, S.M., 1988. TAPHONOMIC COMPARISON OF PASSIVE AND ACTIVE CONTINENTAL MARGINS: NEOGENE SHELL BEDS OF THE ATLANTIC COASTAL PLAIN AND NORTHERN GULF OF CALIFORNIA, Palaeogeography, Palaeoclimatology, Palaeoecology.

Kidwell, S.M., Bosence, D.W.J., 1991. Taphonomy and Time-Averaging of Marine Shelly Faunas, in: Taphonomy: Releasing the Data Locked in the Fossil Record.

Kolbe, S.E., Lockwood, R., Hunt, G., 2011. Does morphological variation buffer against extinction? A test using veneroid bivalves from the Plio-Pleistocene of Florida. Paleobiology 37, 355–368. https://doi.org/10.1666/09073.1

Lawton, E. by J.H., May, R.M. (Eds.), 1995. Extinction Rates. Oxford University Press, Oxford, New York.

Lester, S.E., Gaines, S.D., Kinlan, B.P., 2007. Reproduction on the Edge: Large-Scale Patterns of Individual Performance in a Marine Invertebrate. Ecology 88, 2229–2239.

Lira-Noriega, A., Manthey, J.D., 2014. Relationship of genetic diversity and niche centrality: A survey and analysis. Evolution 68, 1082–1093. https://doi.org/10.1111/evo.12343

Ludt, W.B., Rocha, L.A., 2015. Shifting seas: the impacts of Pleistocene sea-level fluctuations on the evolution of tropical marine taxa. J. Biogeogr. 42, 25–38. https://doi.org/10.1111/jbi.12416

Malakhova, V.V., Eliseev, A.V., 2020. Uncertainty in temperature and sea level datasets for the Pleistocene glacial cycles: Implications for thermal state of the subsea sediments. Global and Planetary Change 192, 103249. https://doi.org/10.1016/j.gloplacha.2020.103249

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

Mayr, E., 1963. Animal Species and Evolution, Animal Species and Evolution. Harvard University Press. https://doi.org/10.4159/harvard.9780674865327

Myers, C.E., Saupe, E.E., 2013. A macroevolutionary expansion of the modern synthesis and the importance of extrinsic abiotic factors. Palaeontology 56, 1179–1198. https://doi.org/10.1111/pala.12053

Osorio-Olvera, L., Yañez-Arenas, C., Martínez-Meyer, E., Peterson, A.T., 2020. Relationships between population densities and niche-centroid distances in North American birds. Ecology Letters 23, 555–564. https://doi.org/10.1111/ele.13453

Parmesan, C., 2006. Ecological and Evolutionary Responses to Recent Climate Change. Annual Review of Ecology, Evolution, and Systematics 37, 637–669. https://doi.org/10.1146/annurev.ecolsys.37.091305.110100

Peterson, A.T., 2001. Predicting Species' Geographic Distributions Based on Ecological Niche Modeling. The Condor 103, 599–605.

Peterson, A.T., Soberón, J., 2012. Species Distribution Modeling and Ecological Niche Modeling: Getting the Concepts Right, Brazilian Journal of Nature Conservation Essays & Perspectives Natureza & Conservação.

Peterson, A.T., Soberón, J., Sánchez-Cordero, V., 1999. Conservatism of Ecological Niches in Evolutionary Time. Science 285, 1265–1267. https://doi.org/10.1126/science.285.5431.1265

Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modeling of species geographic distributions. Ecological Modelling 190, 231–259. https://doi.org/10.1016/j.ecolmodel.2005.03.026

Phillips, S.J., Dudík, M., Schapire, R.E., 2017. Maxent software for modeling species niches and distributions.

Phillips, S.J., Dudík, M., Schapire, R.E., 2004. A maximum entropy approach to species distribution modeling, in: Twenty-First International Conference on Machine Learning - ICML '04. Presented at the Twenty-first international conference, ACM Press, Banff, Alberta, Canada, p. 83. https://doi.org/10.1145/1015330.1015412

Pillans, B., Chappell, J., Naish, T.R., 1998. A review of the Milankovitch climatic beat: template for Plio–Pleistocene sea-level changes and sequence stratigraphy. Sedimentary Geology 122, 5–21. https://doi.org/10.1016/S0037-0738(98)00095-5

Pisias, N.G., Moore, T.C., 1981. The evolution of Pleistocene climate: A time series approach. Earth and Planetary Science Letters 52, 450–458. https://doi.org/10.1016/0012-821X(81)90197-7

Qiao, H., Soberón, J., Peterson, A.T., 2015. No silver bullets in correlative ecological niche modelling: insights from testing among many potential algorithms for niche estimation. Methods Ecol Evol 6, 1126–1136. https://doi.org/10.1111/2041-210X.12397

Rivadeneira, M.M., Hernáez, P., Antonio Baeza, J., Boltaña, S., Cifuentes, M., Correa, C., Cuevas, A., del Valle, E., Hinojosa, I., Ulrich, N., Valdivia, N., Vásquez, N., Zander, A., Thiel, M., 2010. Testing the abundant-centre hypothesis using intertidal porcelain crabs along the Chilean coast: linking abundance and life-history variation. Journal of Biogeography 37, 486–498. https://doi.org/10.1111/j.1365-2699.2009.02224.x

Roopnarine, P.D., 2006. Extinction cascades and catastrophe in ancient food webs. Paleobiology 32, 1–19. https://doi.org/10.1666/0094-8373(2006)032[0001:ECACIA]2.0.CO;2

Sagarin, R.D., Gaines, S.D., 2002. The "abundant centre" distribution: to what extent is it a biogeographical rule? Ecology Letters 5, 137–147.

Sbrocco, E.J., Barber, P.H., 2013. MARSPEC: ocean climate layers for marine spatial ecology. Ecology 94, 979–979. https://doi.org/10.1890/12-1358.1

Shelley, J.J., Swearer, S.E., Dempster, T., Adams, M., Le Feuvre, M.C., Hammer, M.P., Unmack, P.J., 2020. Plio-Pleistocene sea-level changes drive speciation of freshwater fishes in north-western Australia. Journal of Biogeography 47, 1727–1738. https://doi.org/10.1111/jbi.13856

Soberón, J., Peterson, A.T., 2005. Interpretation of Models of Fundamental Ecological Niches and Species' Distributional Areas. Biodiv. Inf. 2. https://doi.org/10.17161/bi.v2i0.4

Soberón, J., Peterson, T., 2011. Ecological niche shifts and environmental space anisotropy: a cautionary note. Revista mexicana de biodiversidad 82, 1348–1355.

Stanley, S.M., 1986. Anatomy of a Regional Mass Extinction: Plio-Pleistocene Decimation of the Western Atlantic Bivalve Fauna. PALAIOS 1, 17–36. https://doi.org/10.2307/3514456

Stigall, A.L., 2014. When and how do species achieve niche stability over long time scales? Ecography 37, 1123–1132. https://doi.org/10.1111/ecog.00719

Taylor, J.D., Glover, E.A., 2010. Chemosymbiotic Bivalves, in: Kiel, S. (Ed.), The Vent and Seep Biota: Aspects from Microbes to Ecosystems, Topics in Geobiology. Springer Netherlands, Dordrecht, pp. 107–135. https://doi.org/10.1007/978-90-481-9572-5_5

Thomas, Y., Bacher, C., 2018. Assessing the sensitivity of bivalve populations to global warming using an individual-based modelling approach. Global Change Biology 24, 4581–4597. https://doi.org/10.1111/gcb.14402

Valentine, J.W., 1989. How good was the fossil record? Clues from the Californian Pleistocene, Paleobiology.

Visher, S.S., 1915. Notes on the Significance of the Biota and of Biogeography. Bulletin of the American Geographical Society 47, 509. https://doi.org/10.2307/201434

Wiens, D., Slaton, M.R., 2012. The mechanism of background extinction. Biological Journal of the Linnean Society 105, 255–268. https://doi.org/10.1111/j.1095-8312.2011.01819.x

Župan, I., Peharda, M., Ezgeta-Balić, D., Šarić, T., 2012. NOAH'S ARK SHELL (ARCA NOAE LINNAEUS, 1758) – WHAT DO WE NEED TO KNOW FOR STARTING UP ITS AQUACULTURE? Croatian Journal of Fisheries 70, 71–81.