PaleoENM: applying ecological niche modeling to the fossil record

Corinne E. Myers, Alycia L. Stigall, and Bruce S. Lieberman

Abstract.—Ecological niche modeling (ENM) is a quantitative approach to predict species’ abiotic requirements. It is a correlative technique, requiring geographically explicit information on species occurrences and the suites of environmental conditions experienced at each occurrence point. The output of these models is a set of environmental suitability rules that can be projected geographically and through time to test biogeographic, ecologic, and evolutionary hypotheses. Although developed by biologists and used extensively in the modern, ENM is in its early stages of application to the deep-time fossil record (hence PaleoENM). In part its limited use in the fossil record thus far reflects the methodological challenge of constructing paleoenvironmental layers needed for PaleoENM analysis, whereas in the modern these layers are available from large public databases (e.g., WorldClim). This paper provides a contextual and methodological framework for appropriately applying PaleoENM, including best practices for developing species occurrence and paleoenvironmental data sets for PaleoENM analyses.

Introduction

Ecological niche modeling (ENM) is a widely used technique developed by biologists for estimating species’ environmental requirements (i.e., abiotic niche attributes) by correlating known species occurrences with spatially explicit environmental characteristics (Guisan and Zimmerman 2000; Guisan and Thuiller 2005; Elith and Leathwick 2009; Peterson et al. 2011). This method allows biologists to test quantitative hypotheses of species’ interactions with their environment across space or into the near future or recent past. Over the last 25 years, ENM has been used to investigate species biogeography (e.g., Svenning and Svok 2004; Guisan and Hérling 2006; Guisan and Thuiller 2005; Elith and Leathwick 2009; Peterson et al. 2011). This method allows biologists to test quantitative hypotheses of species’ interactions with their environment across space or into the near future or recent past. Over the last 25 years, ENM has been used to investigate species biogeography (e.g., Svenning and Svok 2004; Guisan and Hérling 2006; Guisan and Thuiller 2005; Elith and Leathwick 2009; Peterson et al. 2011). This method allows biologists to test quantitative hypotheses of species’ interactions with their environment across space or into the near future or recent past.
both sampling regimes. PaleoENM does differ, however, from more modern applications in the process of acquiring environmental layers across the study area.

For modern analyses, spatially explicit global environmental layers (e.g., temperature, precipitation) can be downloaded at a variety of spatial scales from publicly available online databases (e.g., WorldClim). Increasingly, climate models that reach into the past have also made environmental layers available for parts of the Quaternary period (e.g., Petit et al. 1999; Braconnot et al. 2007; Haywood et al. 2011). However, these layers are not available for intervals in deep time. Thus, in order to use ENM techniques to quantify niche characteristics, deep-time paleobiologists must construct their own environmental layers by using information from sedimentological and geochemical analyses. Although this currently may be more time consuming than a simple download from an online database, it mainly reflects the early stage of PaleoENM development and application. Available environmental information from online databases spanning the Quaternary to recent demonstrates the time already spent compiling large global data sets and climate models for these time periods. Such compilations are important for analyses in deep time as well, and are expected to increase in availability as data digitization efforts (e.g., EarthCube) and PaleoENM studies become more widespread.

Notably, the fossil record has the distinct advantage of preserving species in the environments in which they lived throughout their evolutionary history. In other words, the fossil record preserves a time series of shifting environments and their concomitant effects on species distributions. This provides unique insight compared to the modern record, where species are limited to the single temporal snapshot of the environments they occupy at present. Differences in the temporal scale and geographic resolution achievable in modern and paleoenvironmental datasets dictate that PaleoENM and ENM analyses will have methodological and theoretical differences in their development and interpretation. Both approaches have strengths and weaknesses; thus conceptual context is important for PaleoENM application, as is a standardized and quantitative framework for reconstructing past environments. Our goal in this contribution is to encourage and expand PaleoENM use in a broader range of paleobiological studies. Accordingly, we describe best practices for the application of PaleoENM with emphasis on methods for constructing paleoenvironmental layers for this type of research. We present an illustrative example for creating paleoenvironmental layers based on Late Cretaceous data. Additional examples discuss PaleoENM analyses in the Ordovician of the Cincinnati Basin, the Devonian of the Appalachian Basin, and the Miocene of the Great Plains.

**ENM: Basic Theory and Methods**

**Conceptual Framework**

Over the last decade, ENM has enjoyed increasing popularity in a number of disciplines and research groups. In part, this reflects the user-friendly nature of many modeling algorithms (e.g., Maxent). However, it is vital that ENMs be applied using an explicit conceptual framework and consideration of species-specific characteristics (Austin 2002, 2007; Guisan and Thuiller 2005; Jiménez-Valverde et al. 2008; Peterson et al. 2011; Araújo and Peterson 2012).

Species’ geographic distributions are controlled by three main factors: abiotic conditions necessary for the species’ survival and reproduction, necessary and non-exclusive biotic interactions, and the ability to access suitable areas (Soberón and Peterson 2005; Peterson et al. 2011). Together, these factors make up the Biotic-Abiotic-Movement, or **BAM**, framework of Peterson et al. (2011; see also Soberón and Peterson 2005; Barve et al. 2011; Saupe et al. 2012) (Fig. 1). Within this framework, ENM is a multivariate correlative approach for estimating A, species’ abiotic requirements. That is, by comparing species occurrences with the combinations of environmental factors experienced at each location in environmental space (vs. geographic space), these models move beyond a simple mapping of distributions on environment to provide a set of rules predicting the environmental combinations that are suitable vs. unsuitable for a given species. To the degree that species are able to occupy all suitable abiotic
habitat (i.e., are not B- or M-limited), and that the breadth of sampled habitats reflects the full range of a species' environmental tolerances, ENM provides a prediction of the fundamental niche (sensu Grinnell 1917). In reality, the effect of biotic interactions, historical accessibility, and historical environmental availability are often unknown. Thus, ENM predictions can most fruitfully be interpreted as providing a prediction somewhat broader than the realized niche, but not necessarily comprising the entire fundamental niche.

Statistical Approaches to ENM

ENM may be implemented using a variety of modeling algorithms (see Guisan and Zimmerman 2000; Guisan and Thuiller 2005; Elith and Leathwick 2009; Peterson et al. 2011). The result of most algorithms is a geographically explicit suitability surface that predicts where abiotic conditions are suitable vs. unsuitable for a given species. This is achieved by fitting mathematical functions to the multivariate relationship between occurrence data and environmental factors (Elith and Leathwick 2009; Peterson et al. 2011). Models are trained in a region containing all known species occurrences plus some additional area that is inferred to be accessible, but probably unsuitable, to the species (the M region in Fig. 1) (Soberón and Peterson 2005; Barve et al. 2011; Peterson et al. 2011). Once the algorithm establishes a suitability rule-set for this training region, the model can be projected to a new geographic area and/or to another period of time (e.g., Fig. 2). The product of this projection is a new geographic map composed of suitability scores for a given species in the new region or time. These results can then be used to test hypotheses of observed distribution changes, extinction, speciation, or environmental adaptation.

Many research groups have tested the performance of the different ENM algorithms under different environmental conditions with mixed recommendations (e.g., Hirzel et al. 2001; Elith et al. 2006; Austin 2007; Elith and Graham 2009; Saupe et al. 2012). What is pertinent to PaleoENM is choosing a modeling algorithm that works well with "presence-only" data—that is, an algorithm that requires information about known presences, but not about known absences, of the species. Identifying true species absences is challenging in the modern biological record and next to impossible in the fossil record because of issues of sampling bias, fossil
preservation, or availability of geologic outcrop (Hortal et al. 2008; Jiménez-Valverde et al. 2008, 2011a; Maguire and Stigall 2008; Myers and Lieberman 2011; Peterson et al. 2011; Martin et al. 2012). Two presence-only algorithms recommended for PaleoENM use are Maxent (a maximum entropy algorithm [Phillips et al. 2004, 2006]) and GARP (genetic algorithm for rule-set prediction [Stockwell and Peters 1999]). Both Maxent and GARP appear to function well under many modern scenarios, and are ideally formulated to work with fossil data because they deal well with non-uniform and small sample sizes (Peterson 2001; Stigall Rode and Lieberman 2005a; Hernandez et al. 2006; Guisan et al. 2007; Pearson et al. 2007; Jiménez-Valverde et al. 2008; Peterson et al. 2011). Moreover, ground-truth studies and comparative analyses of deep-time datasets show that both algorithms achieve high predictive accuracy in PaleoENM studies (Malizia and Stigall 2011; Walls and Stigall 2012). A more detailed discussion of available ENM algorithms, including Maxent and GARP, is provided in the supplemental text.

Model Calibration and Evaluation
Model calibration involves selecting appropriate environmental layers and adjusting data and algorithm parameters such that model predictions best match observed species occurrences (Guisan and Zimmerman 2000; Peterson et al. 2011). Once an ENM is calibrated and run, the resulting output is a spatially explicit set of predictions. For Maxent, each pixel value in this surface indicates probability of environmental suitability; for GARP, each pixel registers the sum of “best” models predicting species presence, when employing the “best subsets” procedure (Anderson et al. 2003). Evaluation of
model predictions involves decisions about thresholds, and assessment of both model performance and model significance. Model performance measures include omission error rates; model significance measures whether evaluation data are predicted to be present more often than by random chance (Peterson et al. 2011). An in-depth discussion of model calibration and evaluation for general ENM use is provided in the supplemental text. The reconstruction of environmental layers for PaleoENM is discussed in detail in the PaleoENM Methods in Deep Time section below. Here, we provide a brief discussion of model extent and extrapolation to highlight important conceptual and methodological considerations specific to PaleoENM users.

Model Extent.—Several recent studies have noted the importance of delineating an appropriately sized region within which to train niche models (e.g., Guisan and Thuiller 2005; Barve et al. 2011; Peterson et al. 2011; Saupe et al. 2012; Owens et al. 2013). ENM algorithms use the training region \( M \) in Fig. 1) (Soberón and Peterson 2005; Barve et al. 2011; Peterson et al. 2011) to sample environments both with and without species occurrences in order to determine suitable vs. unsuitable environmental combinations. Thus, \( M \) is the region that could feasibly be sampled (or reached) by a given species and is delimited using information about species’ dispersal capabilities and the distribution of geographic barriers (Fig. 1B). That is, a species that can disperse widely (and sample a large number of habitats) should have a larger hypothesized \( M \) (and model training region) than a species with more limited movement capacity. The size of \( M \) is important because overestimation leads models to speciously predict potentially habitable, but inaccessible, areas as unsuitable. Likewise, underestimation of \( M \) prevents models from having enough information to estimate suitability and may lead to model extrapolation (discussed further below and in Barve et al. 2011; Saupe et al. 2012). An important consideration for PaleoENM application is the geographic availability of sedimentary record. Model extent must be limited to available outcrop area as this constrains where environments and occurrence data are sampled (Fig. 2).

Extending models to areas beyond available outcrop is dangerous because environmental interpolations will be heavily extrapolated, and the lack of species occurrences from these areas will cause algorithms to treat those environmental combinations as unsuitable when this is unknown.

Model Extrapolation.—Extrapolation occurs when the ENM algorithm encounters novel environmental conditions not present in the training data as a result of transferring model predictions beyond the training region in space or time (e.g., black dots in Fig. 1B). The Maxent algorithm allows users to modify the process of model extrapolation in three basic ways: Maxent may “clamp” the model, whereby environments beyond those in the training region are given the same suitability value as the closest training-region pixel in environmental space. Alternatively, Maxent may be allowed to extrapolate under these conditions, whereby suitability scores are assigned based on a continuation of the fitted species response curve (Elith et al. 2011; Peterson et al. 2011; Owens et al. 2013). Finally, both clamping and extrapolation can be turned off, resulting in Maxent giving a low suitability score to all conditions outside of the training region. Allowing models to clamp or extrapolate has a significant effect on model output, as does artificially truncating model predictions outside of the environments experienced in the training region by turning clamping and extrapolation off. The GARP algorithm is also subject to model extrapolation (Owens et al. 2013). Methods exist to explore these effects, e.g., MOP (Owens et al. 2013), MESS (Elith et al. 2010), and environmental overlap maps (Zurell et al. 2012), which should be part of any ENM/PaleoENM interpretation. Importantly, extrapolation is reduced when species occurrence points are centrally (vs. peripherally) located within the environments defining \( M \) (e.g., red dots in Fig. 1B). This may be explored using the program Niche Analyst (Qiao et al. 2012) or statistical evaluation, following which \( M \) hypotheses can be modified to reduce the potential for model extrapolation (Owens et al. 2013). Extrapolation may also be reduced by filtering occurrence points to increase environmental “evenness” as discussed by Varela et al. (2014).
PaleoENM: Methods in Deep Time

Species Occurrence Data

Before conducting PaleoENM or ENM, species distribution data need to be assembled. Species occurrences are collected in a similar fashion (and with the same set of potential biases) in both the modern and fossil records. For example, occurrence data are increasingly collected from large online databases (e.g., GBIF [www.gbif.org] in the modern or the PBDB [www.pbdb.org] in the fossil record). However, direct observations—from fieldwork or museum study—and literature surveys are more effective because they allow for hands-on vetting of the data. The greater the time spent validating species assignments and distributions, the more confidence one can have that modeled results are accurate. Where possible, occurrences should be identified to the species level and localities identified to the most precise geographic and stratigraphic context from the specimen label or database. (Guisan and Thuiller 2005; Peterson et al. 2011). It will then be possible to compare the level of resolution with that of paleoenvironmental data and to remove specimen occurrences that are not of similar resolution. The remaining species occurrence data must then be georeferenced (i.e., locality information translated into latitude and longitude) at a geographic resolution that approximately matches the resolution of paleoenvironmental data, and formatted for ArcGIS (ESRI 2006) and ENM software.

In order to minimize model bias, species occurrence data should include samples distributed across the entire known species range (Araújo et al. 2009; Menke et al. 2009; Jiménez-Valverde et al. 2011b; Peterson et al. 2011). This is particularly important when the aim of modeling is to project environmental suitability rules to other geographic regions or times to test species-level hypotheses of niche characteristics. Application of ENM to specific populations is interesting; however, population-level data cannot accurately be used to build ENMs and then extrapolated to the species level (either by directly projecting population models more globally or by interpreting population-level results at the species level) (see Araújo et al. 2009 and Jiménez-Valverde et al. 2011b for commentary). This is because excluded occurrences truncate species occupation of environmental space, which introduces errors in model results that are unpredictable and algorithm specific (Guisan and Thuiller 2005; Hortal et al. 2008; Barve et al. 2011; Jiménez-Valverde et al. 2011b; Araújo and Peterson 2012; Raes 2012; Saupe et al. 2012; Owens et al. 2013).

The step unique to PaleoENM analysis is determining the stratigraphic interval in which the specimens occur. To generate accurate models, species and environmental data should be contemporaneous at the temporal resolution with which the study is conducted. Temporal resolution will be strongly influenced by the resolution of available taxonomic data and by the resolution best fitting the hypotheses being tested. For example, PaleoENMs of a species that persisted for five million years may perform well with temporal bins in the 1–2 Myr range, whereas investigation of population dynamics requires a smaller bin size. When possible, models should be run at multiple temporal bin sizes, which will help establish the most appropriate resolution for a given study system. One can check model sensitivity to temporal bin size by running multiple models that bin species occurrences (and average environmental conditions) at successively larger stratigraphic intervals. As temporal resolution decreases, the ability of the PaleoENM to discern differences in habitat preferences will also decrease, eventually to the point of being uninformative. Similarly, as temporal resolution is increased, the amount and quality of data required to develop informative models will decrease. Thus, eventually an upper resolution limit will be reached where PaleoENMs cannot be run due to an insufficient number of occurrence points or lack of confidence in down-sampled environmental information.

Paleoenvironmental Data

In the most basic sense, environmental layers are universally constructed by assigning environmental characteristics to unique geographic points from field observations and literature survey. Generally, interpolation is
required to construct continuous environmental layers from these point-source measurements. For example, in the fossil record GIS algorithms such as ordinary kriging or inverse distance weighting are used to interpolate between points and create a continuous coverage of values for each environmental factor across the area of interest (e.g., Stigall Rode and Lieberman 2005a; Maguire and Stigall 2009; Dudei and Stigall 2010). Spatially explicit environmental layers are more readily available in the modern and recent past, to the extent that climate models exist. By contrast, such layers must be reconstructed from the sedimentary record in deep time. Certain environmental information (e.g., temperature, precipitation) is not directly measurable in the geologic record; thus, environmental layers are constructed by using sedimentological and geochemical proxies for environmental factors that are considered important for delimiting habitable areas of the focal taxa. These data sources are well archived and available in the literature and online databases (e.g., http://macrostrat.org; www.earthbase.org). When developing environmental layers for PaleoENM (or ENM), the type and number of layers, the relationships between environmental variables, and the spatial resolution of environmental interpolation are important considerations; these concepts are considered in detail in the supplemental text. Here we briefly discuss environmental layer selection as it affects PaleoENM directly, then provide an example of paleoenvironmental layer construction from the Late Cretaceous Western Interior Seaway (WIS) to illustrate this procedure.

**Selecting Paleoenvironmental Layers.**—Before beginning to collect paleoenvironmental data, it is important to establish what environmental layers will be used in the analysis. The types of layers used will depend on the scale of the study, the specific ecology of the species under investigation, and the types of data available (Austin 2002, 2007; Guisan and Thuiller 2005; Elith and Leathwick 2009; Franklin 2009; Jiménez-Valverde et al. 2011a; Peterson et al. 2011). Examples of indirect variables in the marine realm include bathymetry or latitude. Unfortunately, in the fossil record measurement of direct variables is typically impossible (and is often challenging in the modern as well). However, what can be measured are proxies for direct variables (e.g., geochemical proxies for temperature, oxygenation, pH; sedimentary proxies for wave energy, substrate consistency, grain size), which should be preferred over proxies for indirect variables. Further, when the intended purpose of the analysis is to test for the effects of abiotic change on species evolution, environmental layers should be purely abiotic. The use of biotic variables is inappropriate in these cases because they confound the ability to isolate the effect of abiotic factors and, to the degree that biotic variables may reflect abiotic conditions, they are primarily indirect proxies and so less desirable. The use of biotic variables may be informative for understanding biotic limitations to species occupation of suitable habitat; however, this should be tested independently of abiotic hypotheses. Additional detail on
general ENM considerations regarding number of environmental layers used, spatial resolution, and variable correlation is provided in the supplemental text.

Example from the Late Cretaceous.—To clarify the methods and theoretical constructs developed above (and in the supplemental text), we present a detailed example environmental dataset developed for Late Cretaceous strata from the WIS of North America. The first step is the establishment of temporal bins based on a detailed stratigraphic framework, which constrains both species occurrence and environmental information. In keeping with the resolution of WIS taxonomic data, temporal bin size was limited to the geologic stage level. The last large-scale stratigraphic correlation of WIS formations across North America was completed by Cobban and Reeside (1952). However, understanding of the WIS stratigraphic setting has advanced over the past decades, and thus we updated the stratigraphic correlation for this region at the geologic stage level (Supplementary Table 1 and associated references). Stratigraphic correlations were determined by extensive literature survey and the use of various geologic databases (e.g., USGS National Geologic Map Database [http://ngmdb.usgs.gov], Macrostrat [http://macrostrat.org], and COSUNA, Correlation of Stratigraphic Units of North America Project). Biostratigraphic indices were also used when available, following the Late Cretaceous zonation of Cobban et al. (2006).

Paleoenvironmental data were collected for 14 layers within the Late Cretaceous WIS (Table 1). These include percent clay, silt, sand, and chalk; percent siliciclastic vs. carbonate sediments; bedding style; degree of bioturbation; inferred water depth; depositional environment; oxygenation; and total organic carbon (TOC), $\delta^{13}$C, and $\delta^{18}$O. The 14 environmental layers described here have been modified from those used in previous work (e.g., Stigall Rode and Lieberman 2005a) to reflect the taxa, conditions, and specific hypotheses currently being investigated in the WIS. Paleoenvironmental data were collected primarily through literature survey including peer-reviewed articles, master’s theses, doctoral dissertations, and published fieldtrip guidebooks. Data gathered from new fieldwork were incorporated where possible. Because we used many independent references (e.g., Supplementary Table 1), which employed a variety of terminologies, it was important to develop a standardized scheme for paleoenvironmental variables that ensures consistency in coding. Table 2 provides an example of the coding rule-set used for the Late Cretaceous WIS and associated reference material.

Substrate conditions were characterized using environmental layers describing substrate grain size (i.e., percent clay, silt, sand), percent chalk, proportion of siliciclastic vs. carbonate sediment, degree of bioturbation, and bedding style. Grain size percentages and percent siliciclastic vs. carbonate sediments were calculated from stratigraphic columns as the approximate fraction of each grain size/lithology in a given column. Fossil specimens may be present in particular lithologies within a given section; however, averaging conditions across the full sedimentary package places these in the broader environmental context of the temporal resolution of environmental and species data (Abbott and Sweet 2000). Figure 3 provides a sample stratigraphic column and measurement of these properties. In this example, the total vertical extent of each rock type is first calculated by direct measurement: the column is composed of 2.34 m of sandstone, 5.13 m of shale, 1.04 m of siltstone, and 1.64 m of calcareous shale. Using coding rules provided in Table 2, sandstone is coded as 100% sand, shale is composed of 50% silt and 50% clay, and siltstone is composed of 83% silt and 17% clay. Likewise, the “calcareous” modifier of the shale is coded as 10% carbonate. Following the calculations shown in Figure 3, the grain sizes of this section are coded as 23.1% sand, 41.0% silt, 34.3% clay; and the lithology as 98.4% siliciclastic and 1.6% carbonate.

Degree of bioturbation is a measure of the percentage of beds showing signs of burrowing or other trace fossil activity in a sediment package. This layer is a proxy for general environmental habitability of benthic environments, including such factors as oxygenation, current intensity, depth, seafloor hardness, and rate of sedimentation (Droser and Bottjer 1993; Brenchley and Harper 1998). Bedding style was calculated as the abundance-weighted
average thickness of the beds in a sediment package. Beds may range from laminated (thickness <1 cm) to meter-scale, which describes the amount of sediment input into the marine habitat. Thus, bedding style provides information about turbidity and energy level of the environment (Reading 1996; Prothero and Schwab 2004). Information about bioturbation and bedding style may be estimated directly from stratigraphic columns or found in the lithostratigraphic discussion provided in the accompanying text. In the example provided in

<table>
<thead>
<tr>
<th>Environmental factor</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent clay, silt, sand</td>
<td>Approximate fraction of each grain size within a marine sedimentary package.</td>
</tr>
<tr>
<td>Percent siliciclastics, carbonates</td>
<td>Approximate fraction of siliclastic and carbonate (mainly limestone) sediments within a marine sedimentary package.</td>
</tr>
<tr>
<td>Percent chalk</td>
<td>Approximate fraction of chalk within a marine sedimentary package.</td>
</tr>
<tr>
<td>Degree of bioturbation</td>
<td>Approximate degree of burrowing and other within sediment trace fossil activity within a marine sedimentary package. Decimals delineate relative abundance of trace fossil activity within a marine sedimentary package.</td>
</tr>
<tr>
<td>1. Minimal</td>
<td>&lt;25% sediments show bioturbation.</td>
</tr>
<tr>
<td>2. Moderate</td>
<td>25–50% sediments show bioturbation.</td>
</tr>
<tr>
<td>3. Moderate-high</td>
<td>50–75% of sediments show bioturbation.</td>
</tr>
<tr>
<td>4. High</td>
<td>75–100% of sediments show bioturbation.</td>
</tr>
<tr>
<td>Bedding Style</td>
<td>Approximate thickness of sedimentary beds. Decimals delineate relative abundance of bedding thickness within a marine sedimentary package.</td>
</tr>
<tr>
<td>1. Laminated</td>
<td>&lt;1 cm-scale bedding</td>
</tr>
<tr>
<td>2. Thin</td>
<td>cm-scale bedding.</td>
</tr>
<tr>
<td>3. Moderate</td>
<td>dm-scale bedding.</td>
</tr>
<tr>
<td>4. Thick</td>
<td>m-scale bedding.</td>
</tr>
<tr>
<td>Inferred water depth</td>
<td>Relative water depth with respect to storm and fair-weather wave bases. Decimals delineate relative placement within an energy zone.</td>
</tr>
<tr>
<td>0. Subaerial</td>
<td>Above mean tide line; including delta plain and marsh settings.</td>
</tr>
<tr>
<td>1. Upper intertidal</td>
<td>Between mean low tide and mean high tide; including delta plain and marsh settings.</td>
</tr>
<tr>
<td>2. Lower intertidal</td>
<td>Between mean low tide and fair weather wave base; including upper and middle shoreface settings, delta plain and marsh settings.</td>
</tr>
<tr>
<td>3. Shallow subtidal</td>
<td>Between fair weather and storm wave base; including delta front and prodelta slope settings and lower shoreface settings.</td>
</tr>
<tr>
<td>4. Offshore/basin</td>
<td>Below storm wave base; including delta front and prodelta slope settings.</td>
</tr>
<tr>
<td>1. Estuarine/delta plain</td>
<td>Peritidal; beach and channel deposits, high sediment deltaic environments, shallow estuarine.</td>
</tr>
<tr>
<td>2. Lagoonal/delta front</td>
<td>Nearshore, protected subtidal including shelf lagoons, delta platform, and delta front; frequently heterolithic fine-grained lithofacies with storm deposits; wave-agitated environments including bars, oolite shoals, biohermic areas; above wave base, may or may not be steep.</td>
</tr>
<tr>
<td>3. Inner shelf/prodelta</td>
<td>Dominated by sand and silt deposits; shallow open shelf and prodelta environments, below fair-weather wave base, but evidence of storm deposits.</td>
</tr>
<tr>
<td>4. Midshelf</td>
<td>Dominated by dark clay-muds; deeper open shelf and fore-delta environments; fine-grained sediments; low frequency of storm reworking.</td>
</tr>
<tr>
<td>5. Outer Shelf</td>
<td>Dominated by impure clayey carbonate muds; below storm wave base.</td>
</tr>
<tr>
<td>6. Basin</td>
<td>Dominated by carbonate muds; deep water; black shales; lower oxygen concentration.</td>
</tr>
<tr>
<td>1. Subaerial</td>
<td>Above mean tide line; non-marine.</td>
</tr>
<tr>
<td>2. Normal marine/aerobic</td>
<td>Diverse shelly taxa including epifauna and infauna; bioturbated.</td>
</tr>
<tr>
<td>3. Dysaerobic</td>
<td>Shell epifauna and burrowers dominant; laminated to burrowed sediments.</td>
</tr>
<tr>
<td>4. Anaerobic</td>
<td>Macrofauna; anaerobic S-bacteria; laminated sediments; iron speciation geochemistry.</td>
</tr>
<tr>
<td>δ¹³C, δ¹⁸O, TOC</td>
<td>Average δ¹³C, δ¹⁸O, and total organic carbon per marine sediment package.</td>
</tr>
</tbody>
</table>
Table 2. Coding rule-set used for evaluating paleoenvironmental information from literature survey based on PaleoENM development in the Late Cretaceous Western Interior Seaway of North America. See Table 1 for paleoenvironmental layer descriptions and numerical codes used in PaleoENM analyses.

<table>
<thead>
<tr>
<th>Sediment type</th>
<th>Coding</th>
<th>Comments</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Siltstone/Silt-shale</td>
<td>17% clay, 83% silt</td>
<td>&gt;66% silt, rest clay.</td>
<td>Prothero and Schwab 2004; Potter et al. 2005</td>
</tr>
<tr>
<td>Claystone/Clay-shale</td>
<td>83% clay, 17% silt</td>
<td>&gt;66% clay, rest silt.</td>
<td>Prothero and Schwab 2004; Potter et al. 2005</td>
</tr>
<tr>
<td>Mudstone/Mudrock/Shale</td>
<td>50% clay, 50% silt</td>
<td>Potter 2005: mudrock/shale = 50/50 silt/clay.</td>
<td>Prothero and Schwab 2004; Potter et al. 2005</td>
</tr>
<tr>
<td>Bioclastic mudstone</td>
<td>10% carbonate, 45% clay, 45% silt</td>
<td>Prothero and Schwab 2004: &lt;10% limestone, rest mudstone.</td>
<td>Prothero and Schwab 2004; Neuendorf et al. 2005</td>
</tr>
<tr>
<td>Wackestone</td>
<td>10% carbonate, 45% clay, 45% silt</td>
<td>&gt;10% limestone, rest mudstone; typical of a lagoonal environment.</td>
<td>Prothero and Schwab 2004; Neuendorf et al. 2005; Reading 1996</td>
</tr>
<tr>
<td>Calcareous ooze</td>
<td>15% clay, 15% silt, 60% SiO₂</td>
<td>Based on the “3-component system” of pelagic-hemipelagic sediments: 30% non-biogenic mud (clay and silt), 0–20% calcareous ooze (CaCO₃ microfossils), and 50–70% SiO₂ ooze (SiO₂ microfossils).</td>
<td>Reading 1996</td>
</tr>
<tr>
<td>Glauconite</td>
<td>100% clay</td>
<td>Iron-rich clay; most often found as pellets in sandstones; in agitated, oxidized, normal shallow marine water (max 50–200 m) pellets may form under locally reduced conditions; large concretions of glauconite only in shallow shelf environments with slow sedimentation rates and starved of siliciclastics.</td>
<td>Prothero and Schwab 2004</td>
</tr>
<tr>
<td>Calcareite</td>
<td>100% carbonate</td>
<td>&gt;50% sand-sized grains of carbonate.</td>
<td>Neuendorf et al. 2005</td>
</tr>
<tr>
<td>Calcisiltite</td>
<td>100% carbonate</td>
<td>&gt;50% silt-sized grains of carbonate.</td>
<td>Neuendorf et al. 2005</td>
</tr>
<tr>
<td>Clay Ironstone</td>
<td>10% clay, 90% carbonate</td>
<td>Mixture of up to 30% argillaceous material and iron-carbonate (siderite).</td>
<td>Neuendorf et al. 2005</td>
</tr>
</tbody>
</table>

Inferred water depth

<table>
<thead>
<tr>
<th>Inferred water depth</th>
<th>Coding</th>
<th>Comments</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shoreface</td>
<td>1–3</td>
<td>Zone between seaward limit of shore and near-horizontal surface of offshore zone; typically seaward of shore to storm wave base (~10 m).</td>
<td>Neuendorf et al. 2005</td>
</tr>
<tr>
<td>Upper shoreface</td>
<td>2</td>
<td>Below low tide line; above fair-weather wave base; lower intertidal.</td>
<td>Prothero and Schwab 2004</td>
</tr>
<tr>
<td>Middle shoreface</td>
<td>2</td>
<td>Below low tide line; at fair-weather wave base; lower intertidal.</td>
<td>Prothero and Schwab 2004</td>
</tr>
<tr>
<td>Lower shoreface</td>
<td>3</td>
<td>Below low tide line; below fair-weather wave base; subtidal.</td>
<td>Prothero and Schwab 2004</td>
</tr>
<tr>
<td>Delta plain/Marsh</td>
<td>0–2</td>
<td>Delta plain: landward part of a delta complex; characterized by distributary channels and interdistributary flood basins. Marsh: intermittently to permanently water-covered area. Subaerial to lower intertidal.</td>
<td>Neuendorf et al. 2005</td>
</tr>
<tr>
<td>Delta front</td>
<td>2–3</td>
<td>Area of most-active sediment deposition; within effective depth of wave erosion, &lt;10 m water; subtidal to offshore.</td>
<td>Neuendorf et al. 2005</td>
</tr>
<tr>
<td>Prodelta</td>
<td>3–4</td>
<td>Fully submarine; below the effective depth of wave erosion; subtidal to offshore.</td>
<td>Neuendorf et al. 2005</td>
</tr>
<tr>
<td>Depositional setting</td>
<td>Coding</td>
<td>Comments</td>
<td>Reference</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>--------</td>
<td>--------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Estuarine/Marginal Marine/</td>
<td>1</td>
<td>Peritidal; beach/channel deposits, high sedimentation deltaic environments, shallow estuarine.</td>
<td>Kauffman 1969; Sepkoski 1988; Prothero and Swab 2004; Neuendorf et al. 2005; Stigall Rode and Lieberman 2005a</td>
</tr>
<tr>
<td>Delta Plain</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lagoons/Delta Front</td>
<td>2</td>
<td>Nearshore, protected subtidal including shelf lagoons, delta platform, and delta front; frequently heterolithic fine-grained lithofacies with storm deposits; wave-agitated environments including bars, oolite shoals, biohermic areas; above wave base, may or may not be steep.</td>
<td>Kauffman 1969; Sepkoski 1988; Prothero and Swab 2004; Neuendorf et al. 2005; Stigall Rode and Lieberman 2005a</td>
</tr>
<tr>
<td>Inner Shelf/Prodelta</td>
<td>3</td>
<td>Dominated by sand and silt deposits; shallow open shelf and prodelta environments; below fair-weather wave base, but evidence of storm deposits.</td>
<td>Kauffman 1969; Sepkoski 1988; Prothero and Swab 2004; Neuendorf et al. 2005; Stigall Rode and Lieberman 2005a</td>
</tr>
<tr>
<td>Midshelf</td>
<td>4</td>
<td>Dominated by dark clay muds; deeper open shelf and fore-delta environments; fine-grained sediments, low frequency of storm reworking.</td>
<td>Kauffman 1969; Sepkoski 1988; Prothero and Swab 2004; Neuendorf et al. 2005; Stigall Rode and Lieberman 2005a</td>
</tr>
<tr>
<td>Outer Shelf</td>
<td>5</td>
<td>Dominated by impure clayey carbonate muds; below storm wave base.</td>
<td>Kauffman 1969; Sepkoski 1988; Prothero and Swab 2004; Neuendorf et al. 2005; Stigall Rode and Lieberman 2005a</td>
</tr>
<tr>
<td>Basin</td>
<td>6</td>
<td>Dominated by carbonate muds; deep water; black shales; less oxygenated.</td>
<td>Kauffman 1969; Sepkoski 1988; Prothero and Swab 2004; Neuendorf et al. 2005; Stigall Rode and Lieberman 2005a</td>
</tr>
</tbody>
</table>

Examples of lithological descriptors

<table>
<thead>
<tr>
<th>Coding</th>
<th>Comments</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silty</td>
<td>10% silt</td>
<td>Potter et al. 2005</td>
</tr>
<tr>
<td>Muddy</td>
<td>5% silt, 5% clay</td>
<td>Applied to non-mudstones</td>
</tr>
<tr>
<td>Shaley</td>
<td>5% silt, 5% clay</td>
<td>Having the character of shale/fissility, may be synonymous with “argillaceous”.</td>
</tr>
<tr>
<td>Clayey</td>
<td>10% clay</td>
<td>&gt;10% clay.</td>
</tr>
<tr>
<td>Sandy, Pebble, etc.</td>
<td>10% sand or pebble</td>
<td>&gt;10% sand or pebble, etc.</td>
</tr>
<tr>
<td>Calcareous</td>
<td>10% carbonate</td>
<td>&gt;10% CaCO₂ forams, nannofossil, etc.</td>
</tr>
<tr>
<td>Siliceous</td>
<td>10% SiO₂</td>
<td>&gt;10% SiO₂ diatoms, radiolarians, etc.</td>
</tr>
<tr>
<td>Carbonaceous</td>
<td>1% carbonate</td>
<td>&gt;1% organic carbon.</td>
</tr>
<tr>
<td>Argillaceous</td>
<td>10% clay</td>
<td>Appreciable amount of clay (“argillaceous limestone” has significant, but &lt;50% clay).</td>
</tr>
<tr>
<td>Arenaceous</td>
<td>10% sand</td>
<td>Consisting in some part of sand-sized fragments or having a sandy texture.</td>
</tr>
<tr>
<td>Pyritiferous, Ferruginous</td>
<td>3%</td>
<td>Typically 1–5%;</td>
</tr>
<tr>
<td>Micaceous, Phosphatic, etc.</td>
<td>3%</td>
<td>Typically 1–5%;</td>
</tr>
</tbody>
</table>
Figure 3, two units contain evidence of bioturbation. The sum of these units is 1.14 m, which constitutes 11.2% of the section and the coded value for this location. Coding for bedding style is calculated as described for grain sizes above.

As discussed in Stigall Rode and Lieberman (2005a), the variable “inferred water depth” is a measure of water depth relative to tides, storm-, and fair-weather wave bases. This variable relates to oxygenation in addition to wave energy in a given marine environment (Boucot 1981; Brenchley and Harper 1998; Patzkowsky and Holland 2012). “Depositional environment” is related to distance from the shoreline and relative water depth. Characterization of depositional environments is modified from Stigall Rode and Lieberman (2005a) in conjunction with the methods of other authors (Kauffman 1969; Sepkoski 1988; Prothero and Schwab 2004; Neuendorf et al. 2005). These variables are calculated as abundance-weighted averages within a given sediment package (Table 2 and example in Fig. 3).

As an example from the literature, Owen et al. (2005: pp. 222–224) provided the following lithostratigraphic descriptions of members of the Dakota Sandstone in the Chama Basin of New Mexico:

“The Encinal Canyon of the Chama Basin is far enough east to show abundant evidence of deposition in a marginal-marine environment, perhaps in somewhat protected estuaries, bays, and tidal flats along the western shoreline of the Western Interior seaway during early Cenomanian time…. The Oak Canyon was deposited in an offshore marine environment…. Both Cubero parasequences were deposited as shoreface marine sands, mostly in the middle shoreface zone, but outer shoreface silty sand is more prominent in the lower parasequence. The middle shaley zone was deposited in the adjacent offshore muddy environment…. The Paguate was deposited in a middle and outer shoreface environment that was well populated with burrowing organisms.”

From this information, the depositional environment of the Encinal Canyon Member is estuarine or tidal flats, which, following the coding rules provided in Tables 1 and 2, is coded as 1. The Oak Canyon Member depositional environment is offshore marine. This description is less specific, and so is coded as 3–5 to include the potential contribution of all three offshore shelf marine depositional environments (i.e., inner shelf, mid-shelf, and outer-shelf environments). The Cubero Sandstone Tongue and Paguate Sandstone Tongue represent middle to outer shoreface environments, coded as 3’s. According to the stratigraphic column provided in the text, the Encinal Member makes up 12% of the section, the Oak Canyon Member makes up 15.5%, the Cubero Sandstone Tongue is 50.7%, and the Paguate Sandstone Tongue is 21.8%. Thus, the depositional environment for this sedimentary package is coded as: 1*0.12 + 3*0.052 + 4*0.052 + 5*0.052 + 3*0.507 + 3*0.218 = 2.92. Inferred water depth would be coded in the same fashion.
The environmental layer “oxygenation” describes the relative oxygen content at the sediment-water interface (modified from Sageman and Binna 1997; Brenchley and Harper 1998; Stigall Rode and Lieberman 2005a). This variable is also an abundance-weighted average of a sediment package based on detailed reading of literature sources and/or direct field observation. Bioturbation and oxygenation, as well as depositional environment and inferred water depth, have the potential to be correlated. Generally, it is more prudent to remove correlated variables from the analysis. However, if this is not possible (e.g., there are too few environmental layers to remove any) a PCA could be used to produce new environmental layers composed of principal components (for more on variable autocorrelation see the supplemental text; Guisan and Zimmerman 2000; Guisan and Thuiller 2005; Peterson et al. 2011).

Finally, geochemical proxies such as TOC, $\delta^{13}$C, $\delta^{18}$O may be measured from field samples and/or data provided in the literature. These techniques and their relationships to specific environmental parameters are evolving: TOC may be a proxy for nutrients, oxygenation, and sedimentation rate, whereas $\delta^{13}$C and $\delta^{18}$O are potential proxies for water temperature and salinity under certain conditions (e.g., lack of diagenetic alteration and when analyzed using clumped isotope methods) (Boucot 1981; Johnson Ibach 1982; Creaney and Passey 1993; Fürsich 1993; Brenchley and Harper 1998; Tyson 2001). Currently, sedimentological variables may provide more robust estimates of past environmental conditions; however, as methods and data sampling improve (e.g., clumped isotope studies: Eiler 2011; Dennis et al. 2013), increasingly refined geochemical proxies are likely to become useful tools for estimating important direct variables such as temperature, pH, and oxygenation.

## Applications of PaleoENM

Once species occurrence data have been collected and stratigraphic correlations and paleoenvironmental layers have been reconstructed, a wealth of hypotheses can be tested with ENM to better understand the relationships among ecology, evolution, and the environment. Of particular interest are the effects of changing environments on abiotic niche stability within species, the influence of niche breadth on extinction and speciation rates among species, and the prevalence of phylogenetic niche conservation and its evolutionary consequences. Understanding the accuracy and generality of niche stability, breadth, and conservation under periods of environmental change is significant because these properties limit the geographic expansion of species, which mediates allopatric speciation, extinction resistance, patterns of species richness, community structure, and the spread of invasive species (e.g., Kammer et al. 1997; Peterson 2003; Peterson et al. 2005; Wiens and Graham 2005; Araújo and Rahbek 2006; Kozak and Wiens 2006, 2010; Rangel et al. 2007; Tingley et al. 2009; Wiens et al. 2010; Heim and Peters 2011; Stigall 2012a; Myers and Saupe 2013; Saupe et al. 2014). Thus far, PaleoENM techniques have been applied to studies of Paleozoic marine invertebrates and Neogene terrestrial vertebrates. We briefly describe some of these studies to highlight the types of macroevolutionary hypotheses that can be tested with this approach.

**Survivorship across the Late Devonian Biodiversity Crisis.**—PaleoENM techniques were first applied to the deep-time fossil record by Stigall Rode and Lieberman (2005a), who used PaleoENM methods to assess controls on survivorship of 32 brachiopod and bivalve species in three conodont zones just before and across the Frasnian/Famennian Biodiversity Crisis. Occurrences, stratigraphic correlation, and paleoenvironmental data were collected through a combination of fieldwork, analysis of museum collections, and literature survey; PaleoENMs were developed using the GARP algorithm. Eleven paleoenvironmental layers were constructed with 0.5° grid cell spatial resolution. PaleoENMs were trained within the extent of the Northern Appalachian Basin in eastern North America. This study determined that species with larger areas of predicted suitable habitat were more likely to survive the Devonian Biodiversity Crisis, and that surviving species experienced increases in suitable habitat across the extinction interval. A following investigation considered...
the importance invasive species played in mediating this biodiversity crisis (Stigall and Lieberman 2006). These results provide support and a potential causal mechanism (i.e., large area of suitable environment) for the many previous studies finding a correlation between range area and species longevity (e.g., Kammer et al. 1997; Liow 2007; Stigall 2012b).

Niche Stability in the Ordovician.—A series of analyses in the Late Ordovician have addressed the question of niche stability in a variety of marine invertebrates from the Cincinnati Basin (Fig. 2 provides an example of environmental layer generation and modeling of crinoid Ectenocrinus simplex). Malizia and Stigall (2011) investigated responses of brachiopod species to cyclical sea-level rise and fall across nine time intervals spanning three million years. Occurrences, stratigraphic correlation, and paleoenvironmental data were collected through a combination of fieldwork, analysis of museum collections, and literature survey with a spatial resolution of 15’ grid cells; PaleoENMs were developed using both the GARP and Maxent algorithms. PaleoENMs were trained within the extent of the Cincinnati Basin, covering portions of Indiana, Kentucky, and Ohio, and projected forward in time to facilitate environmental comparisons across species durations. Niche stability was assessed by using percent geographic overlap of niches reconstructed in one time and projected to the subsequent interval, as well as through direct comparison of environmental characteristics using the Schoener’s D statistic in ENMTools (Warren et al. 2008, 2010). This study found that under conditions of gradual environmental change (here cyclical sea-level change), species tracked preferred habitat through the study region, demonstrating niche stability through time. However, when environmental change was coupled with biotic pressure during the Richmondian Invasion, species became more likely to demonstrate changes in niche dimensions (primarily contraction of occupied environmental space). These results were recently expanded to 11 genera of marine invertebrates (adding bryozoans, trilobites, crinoids, anchozoans, bivalves, and gastropods), for which the same pattern was recovered using similar PaleoENM methods and analysis (Brane and Stigall 2014). The congruent results among these taxa suggest that increased competition during invasive regimes may be an important driver of biodiversity patterns on evolutionary timescales (Stigall 2014).

Speciation in the Miocene.—In the first application of PaleoENM techniques to the terrestrial fossil record, Maguire and Stigall (2009) investigated the mid-Miocene radiation of horses in the subfamily Equinae. Occurrences for 30 equid species, stratigraphic correlation, and paleoenvironmental layers were constructed from literature survey at a spatial resolution of 1° grid cell size. PaleoENMs were trained in the Great Plains region and the GARP algorithm was used to reconstruct suitable vs. unsuitable habitat for two time slices: the middle to early late Miocene during the peak of the equid radiation, and the late Miocene to early Pliocene as equid speciation rates declined. High rates of speciation were correlated with statistically greater patchiness of projected suitable habitat during the first and second time slices. This increased patchiness occurred during the initial period of Miocene climate cooling and increased aridity. Other studies (e.g., Vrba 1985; Abe and Lieberman 2012) have also linked increased habitat patchiness with high speciation rates, as isolated populations are more likely to experience speciation by vicariance. However, as climate continued to cool and aridity decreased towards the end of the Miocene, speciation rates declined. This was paired with a decrease in patchiness of predicted suitable habitat, supporting the vicariance model (Stigall 2013).

Conclusions

PaleoENM is an effective and powerful tool for elucidating the relationships between species and their environments. A plethora of previous work has shown that evolution is highly dependent on Earth processes (e.g., Hallam 1981; Cracraft 1982; Raup and Sepkoski 1982; Vrba 1985; Knoll 1989; Allmon and Ross 1990; Raup 1994; Knoll et al. 1996; Carroll 2000; Lieberman 2000, 2003; Barnosky 2001; Rothschild and Lister 2003; Stigall Rode and Lieberman 2005a,b; Lieberman et al. 2007; Maguire and Stigall 2008). Thus, PaleoENM may be used to quantitatively
test hypotheses regarding the effect of a dynamic planet on species evolution. There is also a broad agreement that large-scale, independent events have significantly affected evolutionary history by causing major mass extinctions (e.g., Gould 1985, 2002; Jablonski and Raup 1995; Jablonski 2001; Congreve 2013). However, an oft-overlooked corollary is the role (and to what degree) abiotic variables play in initiating evolutionary change (see discussion in Lieberman et al. 2007; Knoll 2012; Stigall 2012b, 2013; Myers and Saupe 2013). PaleoENM can be used to investigate this issue in addition to other macroevolutionary phenomena, such as niche stability, the evolutionary effect of niche breadth, and phylogenetic niche conservation. In particular, phylogenetic niche conservatism is a pattern increasingly observed among sister species of modern biota (e.g., Peterson et al. 1999; Graham et al. 2004; Pearman et al. 2008; Weins et al. 2010; Peterson 2011). Investigations in the deep-time fossil record provide an important temporal perspective for evaluating the generality and macroevolutionary effect of these phenomena. PaleoENM analyses may also be useful in distinguishing between evolutionary radiations driven by adaptive vs. non-adaptive processes (e.g., natural selection vs. geographic complexity, exaptation, species-level selection) (Abe and Lieberman 2012; Lieberman 2012).

PaleoENM is similar in theory and in most aspects of its application to ENM analyses of modern taxa. Uniquely, however, it requires the acquisition of detailed taxonomic and geographic species occurrence data within a stratigraphic context. We have provided information on how to develop environmental layers for this type of analysis, and also described the various caveats necessary to apply this approach successfully. In particular, we have tried to present a possible standard for some of the best practices for this technique. We hope this will serve as a guide for future paleobiologists interested in applying PaleoENM to quantitatively test hypotheses of species-environment interactions in the deep-time fossil record.

Acknowledgments

We are very grateful to A. Knoll, E. Saupe, M. Kowalewski, and two anonymous reviewers for thoughtful comments and discussion of earlier versions of this manuscript. This research was supported by the NASA Astrobiology Institute Postdoctoral Program (C.E.M.), Madison and Lila Self Graduate Fellowship (C. E.M.), and a National Science Foundation’s Advancing Digitization of Biodiversity Collections grant (no. 1206757 to B.S.L. and A.L.S.).

Literature Cited


Petit, J. R., J. Jouzel, D. Raynaud, N. I. Barkov, J.-M. Barnola, I. Basile, M. Bender, J. Chappellaz, J. Davis, G. Delaygue,


