Paleontological Contributions

Number 6

A Field-Based Analysis of the Accuracy of Niche Models Applied to the Fossil Record

Bradley J. Walls and Alycia L. Stigall

KU PALEONTOLOGICAL INSTITUTE
The University of Kansas

April 12, 2012
Lawrence, Kansas, USA
ISSN 1946-0279 (online)
paleo.ku.edu/contributions
http://hdl.handle.net/1808/8950

Paleontological Contributions

April 12, 2012 Number 6

A FIELD-BASED ANALYSIS OF THE ACCURACY OF NICHE MODELS APPLIED TO THE FOSSIL RECORD

Bradley J. Walls[1‡] and Alycia L. Stigall[1,2*]

[1]Department of Geological Sciences and [2] OHIO Center for Evolutionary Studies, Ohio University, Athens, Ohio 45701, USA, bradley.walls@weatherfordlabs.com, stigall@ohio.edu

ABSTRACT

The use of ecological niche modeling (ENM) to estimate the geographic ranges of species is widely employed with modern fauna and is becoming more widespread in paleontology. Herein, field validation is utilized to assess the predictive accuracy of ENM methods for Paleozoic brachiopod species. This study represents the first field validation analysis of ENM methods in the fossil record. Previously published species distributions models for 8 Late Ordovician brachiopod species from the Cincinnati, Ohio region (United States) developed using GARP (Genetic Algorithm using Rule-set Prediction) were assessed for accuracy by comparing species occurrence data from a newly available set of 18 localities with the original species distribution models. Based on this data, the statistical significance of the original model set was assessed; 18 of the 22 original models were demonstrated to be statistically significant, based on field validation. Of the 140 individual species occurrences assessed, 60.8% were accurately predicted, 9.2% exhibited over prediction, and 30% exhibited under prediction. Accurate results were more common for species modeled from the greatest number of species occurrence points. The least accurate species models developed were for eurytopic species or those for which taxonomic affinities are unclear. Results indicate that with ample outcrop, well-defined stratigraphy, and sufficient fossil occurrence data, ENM methods could be successfully applied to many intervals in Earth history.

Keywords: brachiopod, field validation, GARP, Late Ordovician, paleobiogeography

INTRODUCTION

Ecological niche modeling (ENM) is used commonly in biogeographic studies to predict the geographic ranges of taxa. A series of ENM programs have been developed (see reviews in Elith & others, 2006). Although algorithms and data input differ, all ENM methods reconstruct the niche of a species based on a set of known species occurrence points and the environmental conditions at those locations (Stockwell & Noble, 1992; Stockwell & Peters, 1999; Peterson, 2001). The accuracy of the niche models produced is assessed using internal tests as part of the model creation or by retaining a subset of species occurrence points not used for model creation for model validation (e.g., Guisan & others, 2007; Walls & Stigall, 2011). Much less commonly, models have been validated through secondary field work (e.g., Feria & Peterson, 2002; Raxworthy & others, 2003). Although the accuracy of various ENM methods has been well established for analyses with modern organisms (e.g., Elith & others, 2006; Guisan & others, 2007; Elith & Graham 2009), paleontological analyses have largely relied on the internal tests generated during model creation to determine model validity (e.g., Stigall Rode & Lieberman, 2005; Maguire & Stigall, 2009; Dudei & Stigall, 2010; Walls & Stigall, 2011). Sequestering data for secondary validation is often not possible when models are generated with small sample sizes, such as is typically the case when applying

*Corresponding author.
‡Current address: Weatherford Laboratories, Houston, Texas 77086, USA.

Copyright © 2012, The University of Kansas, Paleontological Institute

Table 1. Brachiopod species modeled per time slice and number of species occurrence data points utilized in the creation of original GARP prediction models in Walls and Stigall (2011); †, models that are not statistically significant; *Vinlandostrophia* = *Platystrophia* King, 1850, in part (new).

Species	Time slice		
	Early	Middle	Late
Dalmanella meeki	6	7	5
Hebertella occidentalis	10	5	6
Rafinesquina alternata	17†	25	22†
Gnamptorhynchos auburnensis	–	11†	11
Vinlandostrophia cypha	–	5	5†
Vinlandostrophia laticosta	5	5	6
Vinlandostrophia ponderosa	16	48	45
Zygospira modesta	7	11	14

ENM models to the fossil record; and secondary field validation of paleontological ENM studies has not previously been undertaken. Herein, we assess the results of a previously published ENM analysis of Late Ordovician brachiopods (Walls & Stigall, 2011) using a set of species occurrence data acquired from a new donation of museum specimens and secondary field work.

Specifically, this analysis examines the validity of the GARP (Genetic Algorithm for Rule-set Prediction) modeling system (Stockwell & Peters, 1999) for generating paleontological ecological niche models using sedimentological proxies for environmental variables. Although taxon sampling is often comparable between modern and fossil ENM analyses, environmental data sets differ substantially in their generation. Modern environmental data is typically acquired via download from online databases, whereas paleoenvironmental conditions must be inferred from field investigation, analysis of published stratigraphic columns, or derived from climate models (useful for Quaternary analyses only). All paleontological applications have reported results of high statistical significance based on internal model tests, but none of these analyses have validated in the field the method's efficacy in the fossil record. In this study, the ability of GARP to accurately predict species distributions in the fossil record is assessed by testing the Walls and Stigall (2011) model results against data collected *post hoc* from targeted field and museum collections. These results provide the first direct test of the utility of ENM methods for predicting niches and modeling geographic ranges of fossil taxa.

GARP has been used extensively with modern fauna. Applications include predicting effects of climate change on species distributions, assessing invasive species risks, testing evolutionary hypotheses, and guiding conservation efforts, such as selecting habitat for endangered species (e.g., Peterson, Soberón, & Sánchez-Cordero, 1999; Araújo & Williams, 2000; Peterson, 2003; Thomas & others, 2004; Thuiller & others, 2005; Chen, Wiley, & McNyset, 2007). Although GARP is widely employed, only a handful of prior studies (e.g., Feria & Peterson, 2002; Raxworthy & others, 2003; Stockman, Beamer, & Bond, 2006; Kostelnick & others, 2007; DeVaney & others, 2009) have attempted to validate model results in the field. Field validation for each of these analyses recorded high rates of predictive accuracy, with reported accuracy levels ranging from 75% to 96% prediction success. Notably, omission levels (false negatives) and commission levels (false positives) were generally also low, typically beneath 10 or 20%.

GARP is the most commonly used modeling program in published paleo-ENM studies (e.g., Martínez-Meyer, Peterson, & Hargove, 2004; Stigall Rode & Lieberman, 2005; Martínez-Meyer & Peterson, 2006; Maguire & Stigall, 2009; Dudei & Stigall, 2010; Walls & Stigall, 2011). GARP is designed to analyze species occurrence data sets that are assembled by sampling data that are neither uniform nor designed for statistical tests (such as data collected from museums or randomly selected outcrops) and where environmental data consist of poorly structured domains (Stockwell & Peters, 1999). These features are critical for analyses of fossil species, because geographic occurrence data can rarely be collected within a statistical sampling regime. The GARP system has been tested extensively and has been shown to achieve high accuracy with low numbers of species occurrence data, even when there are as few as five environmental parameters (Peterson & Cohoon, 1999; Peterson, 2001; Anderson, Peterson, & Gómez-Laverde, 2002; Stockwell & Peterson, 2002; Hernandez & others, 2006). Effectiveness with small sample sizes is critical in studies of fossil taxa where species occurrence and environmental data may be limited. Furthermore, GARP requires only presence data (Peterson, Papeş, & Eaton, 2007). Because absence in the fossil record is often due to undersampling or lack of outcrop availability rather than true absence, this is a key attribute required for paleontological applications. Although other modeling systems, notably Maxent (Phillips, Anderson, & Schapire, 2006), are gaining traction in paleontological analyses, their use in deep-time analyses with sedimentary proxies for environmental data has been limited to date (i.e., Malizia & Stigall, 2011; Stigall, 2011). We, therefore, restrict the analyses in this paper to a test of the GARP modeling system.

METHODS

The field validation presented in this paper examines GARP output models published in Walls and Stigall (2011). In that study, GARP was utilized to model the geographic ranges of eight brachiopod species during three time slices within the Late Ordovician C3 depositional sequence (Table 1, Fig. 1) of the Cincinnati, Ohio region (Walls & Stigall, 2011). This stratigraphic interval is characterized by gradual shallowing of sea level of the Cincinnatian shallow marine ramp system, due to the basin infilling with sediments weathered from the Taconic Highlands (Brett & Algeo, 2001). In order to study how species were affected by the relative sea level fall, the C3 sequence was divided into early, middle, and late time slices to provide a framework for assessing how the geographic range of each species changed through time.

The primary results of the biogeographic analysis of Walls and Stigall (2011) indicated that the amount of relative sea level fluctuation impacted the stability of niche parameters of brachiopod taxa differently between time slices and taxa. Data for the original analysis in Walls and Stigall (2011) were obtained from field collection, published literature, and unpublished references, including theses and field guides. From these data, five environmental basemaps were created to model the niche of each species (inferred

water depth, limestone bedding style, limestone thickness, biofacies, and lithology). Between 5 and 48 spatially unique points were used to model species niches in each time slice (Table 1). Models were generated using Desktop GARP (www.nhm.ku.edu/desktopgarp/), as detailed in Walls and Stigall (2011). The 50 best subset models were retained. In the final maps, values of 50 indicate regions of the research area where all of the 50 best models predicted the species occurrence, while values of 0 indicate areas where none of the best models predict species occurrence, and so forth.

Results from Walls and Stigall (2011) indicate a clear link between relative sea level and species ranges. As relative sea level fell from the early to middle time slice, species exhibited high levels of niche stability (average of 75% niche overlap between early to middle time slices). Because the decline in relative sea level decreased the maximum potential geographic extent of specific niches, species inhabiting these niches also experienced a decrease in geographic range. This is demonstrated by a reduction in average geographic range for species in the middle versus the early time slice (75% versus 89% of the map area occupied). From the middle to late time slice, however, the average range of brachiopod species increased (to 80% of the map area) as five of the eight species modeled exhibited greater niche evolution (average of only 49% niche overlap with middle to late time slices) in response to the fall in relative sea level. Subsequent analyses by Malizia and Stigall (2011) that used the methods of Warren, Glor, and Turelli (2008) to examine niche conservatism generated with the Maxent algorithm produced congruent results.

The present study tests the significance of the 22 geographic range models produced by Walls and Stigall's (2011) GARP analysis of the C3 sequence using subsequent field validation. A total of 18 new sites were identified to test the original GARP models (Fig. 2, Table 2, Appendix 1). Species occurrence data for these new sites were based on 10 sites identified from a new donation of specimens to the Ohio University Paleontology Collections and investigations of 8 new field locations, which were selected from regions where the previously generated GARP models suggested species occurrence but that were previously unsampled. Field validation sites were strategically selected to sample portions of the map area where 0 to 5 of the 50 best models predicted species occurrence, where 26 to 45 of the best models predicted occurrence, and where 45 to 50 best models predicted species occurrence. For the 8 new field sites, field work consisted of identification of brachiopod species *in situ* and construction of stratigraphic columns (see Appendix 2). Stratigraphic columns were constructed to provide a basis for correlation of the new field sites with the stratigraphic framework detailed in Walls and Stigall (2011). Each site was inspected for several hours to exhaustively determine presence or absence of brachiopod species. Additional sites were obtained by canvassing the Ohio University Paleontological Collection for species occurrence information. Absence data were only counted for sites visited during the course of field work for this study. These absences, which persisted after directed searches for generally common species, are considered to be true absences. Absences within the museum data set, however, cannot be distinguished from undersampling and are thus excluded from this study. For each species, a minimum of 5 total sites were investigated (Table 1).

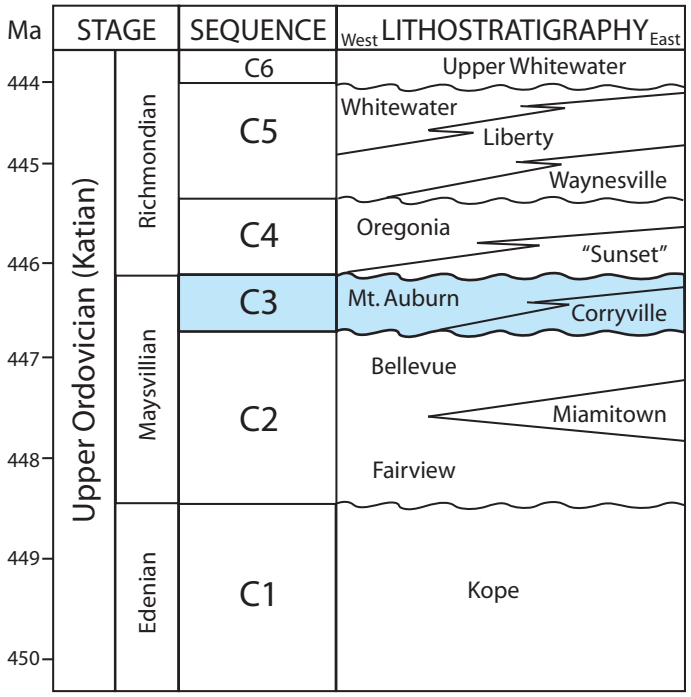


Figure 1. Generalized stratigraphic column of Upper Ordovician strata in the Cincinnati region; this study focuses on the C3 depositional sequence, highlighted in *blue* (adapted from Holland & Patzkowsky, 1996).

Accuracy was assessed relative to whether a species was predicted to occur at a site in more than half of the best subset models. This threshold was selected for analysis, because it is neither too broad to dampen the results nor so specific that no pattern can be assessed. For the predicted distribution models examined, there is strong agreement among the majority of the range predictions, resulting in a sharply defined region of predicted occurrence for the 26 or more of the 50 best subset maps (Fig. 3). However, inclusion of all 50 of the best subset maps would result in a substantial expansion of the predicted range that results in much of the predicted range being supported by less than 20% of the best subset models. Because of this pattern, a threshold of greater than 50% of the best subset models is more informative than the more commonly used threshold of all of the best subset models (e.g., Kostelnick & others, 2007; DeVaney & others, 2009). Alternate thresholding methods, such as calculation of the E parameter as proposed by Peterson, Papeş, and Soberón (2008), have been proposed in the literature. Incorporation of these methods, however, required generation of detailed AUC (area under the curve) analyses to determine model validity in concert with the original ENM generation. Since the intent of this study is a *post hoc* analysis of the accuracy of the original Walls and Stigall (2011) results, these methods are not feasible for the current project.

Model accuracy at field validation sites was analyzed, both statistically and by directly calculating accuracy, omission (underprediction), and commission (overprediction) rates. For statistical analysis,

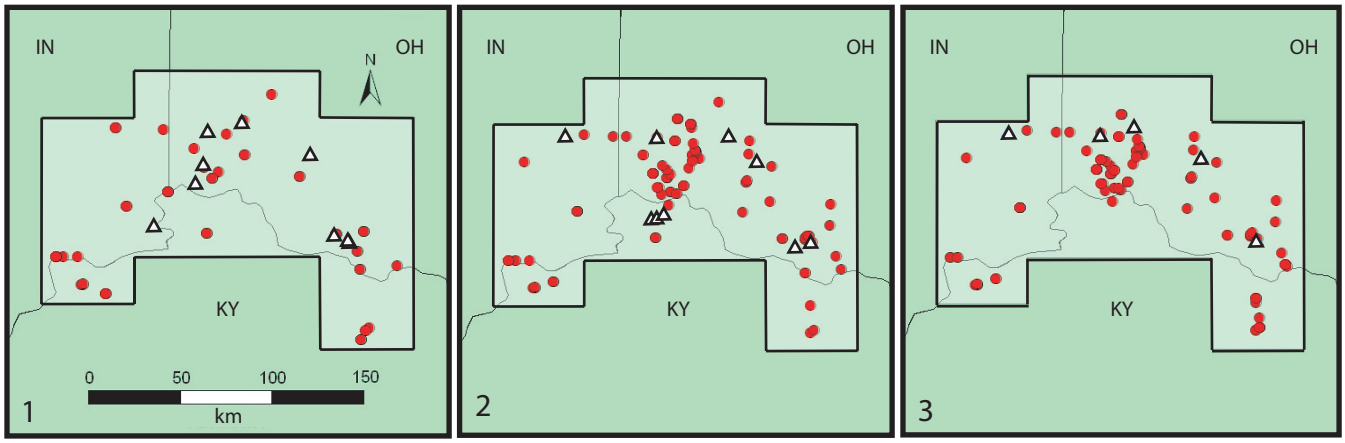


Figure 2. Location of field validation sites; *circles,* species occurrence localities utilized for original GARP models of Walls and Stigall (2011); *triangles,* sites utilized for field validation; *1,* early time slice; *2,* middle time slice; *3,* late time slice; *IN,* Indiana; *OH,* Ohio; *KY,* Kentucky (new).

the probability of model accuracy was adjusted for the geographic area of the predicted distribution, which is similar to the method employed by Raxworthy and others (2003). Probability values were calculated by multiplying the probability of random capture of a species at a location, based on percentage of the model region a species was predicted to occupy (for verified species occurrences) or to be absent (for verified species absences) for all accurately predicted species occurrences. This method, therefore, calculates the probability that species occurrence accuracy is due to random distribution within the field area adjusted by the niche breadth of each species. Fischer's Exact Test was employed for data aggregated by time slice. Fischer's Exact Test was not employed for individual models, however, because the sample sizes for individual species are too small for that test to overcome type II errors. Statistical methods specific to ENM validity assessment, such as calculation of partial receiver operator curve AUCs (Peterson, Papeş, & Soberón, 2008) and Pearson's Low Sample Test (Pearson, Papeş, & Eaton, 2007), have been developed. However, these techniques are conducted in concert with the original niche model generation and/or operate on the raw data files from the original modelling and, therefore, are not able to be applied this *post hoc* analysis. Although these ENM specific methods would provide sensitive tests of model accuracy, the methods employed herein have a history in the literature and been used previously to successfully investigate accuracy of GARP models (e.g., Raxworthy & others, 2003).

RESULTS

A total of 140 species occurrence sites were investigated across the 18 localities to test the original 22 GARP models (Table 3). Of these, 54 sites of predicted species occurrences were verified to be correctly predicted. Overprediction, where the species was predicted to occur but was not found, was identified at 13 sites. For 42 sites, species occurrences were verified where that species was predicted to be absent, which indicates model underprediction. Finally, 31 sites of predicted species absence were confirmed to be correct. In summary, 85 sites (60.8%) correctly verified the original GARP prediction maps, with 13 sites (9.2%) of underprediction and 42 (30%) sites of overprediction. The Fisher's Exact Test result of the combined validation data for all 22 models is statistically significant ($p = 0.002$), supporting high accuracy of aggregate model results.

Statistical analyses also demonstrated high levels of predictive accuracy within the original GARP models. Statistical significance of the original models was assessed, based on the probability of a species occurring by random chance in the map area versus the probability of locating the species only where it was predicted to occur. Results

Table 2. Sites investigated for field validation of GARP models by time slice. *, sites based on Ohio University Paleontology Collection; other locations are sites of field investigation; additional details in Appendices 1 and 2 (new).

Early time slice	Middle time slice	Late time slice
Faber Ave., Cincinnati, Ohio	Brookville Dam, Indiana	Brookville Dam, Indiana
Georgetown, Ohio	Eagle Creek, Ohio	Eagle Creek, Ohio
Rising Sun, Indiana*	Kentucky Drive, Carrolton, Kentucky*	Lick Run, Osceola, Ohio
Carrolton, Kentucky*	Lebanon, Ohio*	Middletown, Ohio
South Rd., Mack, Ohio*	Lick Run, Osceola, Ohio	Wal-Mart, Hamilton, Ohio
U.S. 68, Red Oak, Ohio*	Power Line Dr., Florence, Kentucky*	
Sheits Rd., Dornbusch, Ohio	S. Loop Dr., Florence, Kentucky*	
Wal-Mart, Hamilton, Ohio	T. Moore Pky., Florence, Kentucky*	
Red Oak North, Ohio*	U.S. 68, Red Oak, Ohio*	
	Wal-Mart, Hamilton, Ohio	

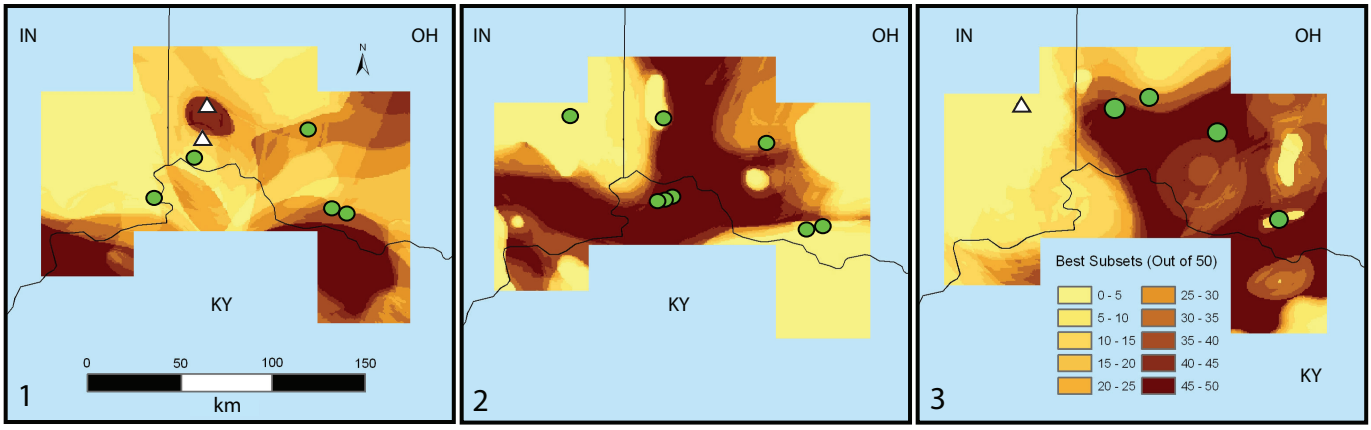


Figure 3. Field validation sites overlain on original GARP prediction models from Walls and Stigall (2011); *darker shading,* region where 45–50 of the best 50 models predicted species occurrence; *light shading,* region where 0–5 of the best 50 models predicted species occurrence; *circles,* validation sites where the species was identified as present; *triangles,* validation sites where the species was identified as absent. *1,* Prediction map for *Hebertella occidentalis* for early time slice; *2,* prediction map for *Rafinesquina alternata* for middle time slice; *3,* prediction map for *Vinlandostrophia ponderosa* for late time slice; *IN,* Indiana; *OH,* Ohio; *KY,* Kentucky (new).

indicate that 18 out of the 22 species models are significant at the 95% confidence level (Table 4).

DISCUSSION

Statistical results of field validation analysis demonstrate the original GARP models were accurate at predicting species occurrences. Eighteen out of the 22 GARP models were shown to be statistically significant at the 95% confidence level (Table 4), with 61% of all the sites investigated correctly predicting species occurrences or absences (Table 3). Some species, however, have higher predictive accuracy than others. Of the 8 species models that were field validated, all models for the species *Dalmanella meeki, Hebertella occidentalis, Vinlandostrophia laticosta, V. ponderosa,* and *Zygospira modesta* were demonstrated to be statistically significant. Conversely, one model each for *V. cypha, Gnamptorhynchos auburnensis,* and two models of *Rafinesquina alternata* were not statistically different from random. Model accuracy can be explained primarily by 4 parameters: data density, niche breadth, species abundance, and taxonomic accuracy.

The primary control on model accuracy was the number of species occurrence points available for model generation. Within our tests, models generated from larger numbers of spatially distinct species occurrences exhibited higher levels of accuracy in field validation tests ($r = 0.42$, $p = 0.05$). Of species with statistically significant models in in all three time slices, the percentage of occurrences accurately predicted was higher for *Vinlandostrophia ponderosa* (average 80%) and *Zygospira modesta* (average 63%) than the other three species (averages 45%–60%) (Table 3). The original models for these 2 species were constructed with the greatest amount of species occurrence data of the five most accurately modeled species (Table 1). *Vinlandostrophia ponderosa* had the highest number of species occurrence points for the original models (16 to 48), whereas *Zygospira modesta* models were based on 7 to 14 species occurrence points. In comparison, the original model for *Vinlandostrophia cypha* was constructed with 5 species occurrence points in each time slice and was not statistically significant for the late time slice (Table 1). Computer modeling has shown GARP to be able to achieve high predictive accuracy with as few as 5 species occurrence points (Peterson & Cohoon, 1999), although it performs most reliably with greater than 10 species occurrences (Hernandez & others, 2006). In our field validation, 4 of the 5 niche models constructed with only 5 species occurrence points were statistically significant (Tables 1, 4); however, they are prone to presence/absence accuracy errors (Table 3).

Although models with more data produce more accurate results in general, the minimum number of discrete data points required for successful models is related to niche breadth. Notably, 2 of the 3 models generated for *Rafinesquina alternata* were not statistically significant, even though they included more discrete species occurrence points than many other models in this study that were statistically significant (Table 1). *Rafinesquina alternata* is a very common species and was collected at every site visited. This species has previously been interpreted as an ecological generalist with wide environmental tolerances (Holland, 2005; Stigall, 2010, 2011), and *R. alternata* occupies a larger percentage of the study region in each time slice than any other species (Table 4). This flexibility to inhabit most paleoevironments in the study area makes it difficult for the ENM program to model a constrained set of conditions that encompass the ecological niche of the species. Even though *R. alternata* is represented by between 17 and 25 species occurrence points per time slice, the inaccurate models included the least data, with 17 and 22 species occurrence points. Even though accurate range models were produced for specialist taxa (e.g., *Dalmanella meeki*) with as few as 5 species occurrence points, it is possible that generalist species may require at least 20 species occurrence points for the niche model to be accurate within the framework of this analysis. Analyses of modern taxa (e.g., Raxworthy & others, 2003; Hernandez & others, 2006; Guisan & others, 2007) also demonstrated difficulty in modeling broadly distributed species, compared to narrowly adapted species.

Another factor influencing model accuracy is the relative abundance of each species within the ecosystem and, consequently, at the outcrop. With abundant species, the effects of sampling bias are reduced compared to uncommon or rare species (Westrop & Adrain, 2001; Finnegan & Droser, 2005; Tarver, Braddy, & Ben-

Table 3. Field validation data and error rates; predicted presence or absence is relative to the area where 26 or more of the 50 best subset model predictions of Walls and Stigall (2011) predict species occurrence; additional details in Appendix 3 (new).

Species	Sites of predicted presence		Sites of predicted absence		Error rate		Accuracy
	Present (accurate)	Absent (comission)	Present (omission)	Absent (accurate)	Comission	Omission	
Early time slice							
Dalmanella meeki	1	0	2	2	0%	40%	60%
Hebertella occidentalis	3	2	2	1	25%	25%	50%
Rafinesquina alternata	4	0	4	0	0%	50%	50%
Vinlandostrophia laticosta	2	2	4	0	25%	50%	25%
Vinlandostrophia ponderosa	3	0	2	2	0%	29%	71%
Zygospira modesta	3	0	5	0	0%	63%	38%
Total	**16**	**4**	**19**	**5**	**9%**	**43%**	**48%**
Middle time slice							
Dalmanella meeki	3	0	3	2	0%	38%	63%
Hebertella occidentalis	2	2	1	1	33%	17%	50%
Rafinesquina alternata	6	0	3	0	0%	33%	67%
Gnamptorhynchos auburnensis	0	2	0	3	40%	0%	60%
Vinlandostrophia cypha	1	1	0	4	17%	0%	83%
Vinlandostrophia laticosta	3	0	2	2	0%	29%	71%
Vinlandostrophia ponderosa	6	0	1	1	0%	13%	88%
Zygospira modesta	4	0	2	1	0%	29%	71%
Total	**25**	**5**	**12**	**14**	**9%**	**21%**	**70%**
Late time slice							
Dalmanella meeki	0	1	3	1	20%	60%	20%
Hebertella occidentalis	2	0	1	2	0%	20%	80%
Rafinesquina alternata	3	0	2	0	0%	40%	60%
Gnamptorhynchos auburnensis	0	1	2	2	20%	40%	40%
Vinlandostrophia cypha	0	1	1	3	20%	20%	60%
Vinlandostrophia laticosta	2	1	0	2	20%	0%	80%
Vinlandostrophia ponderosa	3	0	1	1	0%	20%	80%
Zygospira modesta	3	0	1	1	0%	20%	80%
Total	**13**	**4**	**11**	**12**	**10%**	**28%**	**63%**
Combined	**54**	**13**	**42**	**31**	**9%**	**30%**	**61%**

ton, 2007). If a species is rare, sites of known species occurrence usable for niche modeling may comprise very few locations. These locations could represent chance collections within what may be a much larger range, geographically and ecologically. Therefore, it is likely that analyses pertaining to rare species will underpredict species occurrence due to undersampling of the ecological variation that exists within the full geographic range of the species. In our analysis, underprediction was documented in 30% of species models investigated (Table 3). Species with fewer discrete occurrence points in the original GARP models more commonly exhibit underprediction (compare Tables 1 and 3). For example, field validation of *Dalmanella meeki* range models recovered sites of underprediction in all three time slices (Table 3).

Conversely, common species should be expected to have fewer sampling biases and have more accurate range models, which is congruent with our model results. For example, *Vinlandostrophia ponderosa* and *Zygospira modesta* are species that are both easily identifiable and are abundant in the stratigraphic layers in which they occur (Holland, 2005). As noted above, these species models exhibit the highest accuracy rates and were verified to occur at all sites of predicted species presence, regardless of time slice (Table 3). Furthermore, all range models for these species were shown to be statistically significant. However, these species also occurred at some sites where they were predicted to be absent in 26 or more of the 50 best models. Of the 40 sites investigated for these 2 species, the species occurred at 34 sites. Twenty-eight sites visited verified the predictions of the original GARP models, whereas 12 sites yielded specimens in locations where they were not predicted to occur. This suggests that for very common species (much like rare species), underprediction may be more of a concern than overprediction of species ranges for ecological niche models.

An additional factor contributing to model validity is the taxonomic accuracy. If species included in this analysis are paraphyletic or polyphyletic groups rather than discrete biological entities, then no coherent niche exists for the set of organisms represented by the species occurrence data set, and poor model results should be expected. This likely contributed to the poor model results of *Vinlandostrophia cypha* and *Gnamptorhynchos auburnensis*. Each species exhibited one statistically invalid model: *G. auburnensis* in the middle time slice and *V. cypha* in the late time slice (Table 4). Although the poor predictive ability of these models is due partly to low sample size,

as discussed above, it is also likely a factor of taxonomic confusion. *Vinlandostrophia cypha* and *G. auburnensis* are both problematic species. In fact, the entire genus *Vinlandostrophia* Zuykov & Harper, 2007 (=*Platystrophia* King, 1850, in part) is taxonomically unstable. Cummings and Mauck (1902) determined there is no single character or combination of characters that can be utilized to separate specimens referred to *G. auburnensis, V. laticosta,* and *V. dentata* into distinct species. Alberstadt (1979) determined that many species within the genus *Vinlandostrophia*, including *V. cypha* and *V. laticosta*, were so similar morphologically that a complete reevaluation of the genus was necessary. Zuykov and Harper (2007) also indicated the need for revision of species assigned to the genus *Vinlandostrophia.* This taxonomic confusion can obfuscate attempts at creating accurate niche models by polluting the data input to the models.

CONCLUSIONS

The utility of ecological niche modeling, specifically the GARP modeling system, is confirmed as a useful tool for paleobiogeographic studies through field validation of brachiopod species in the Late Ordovician units of the Cincinnati Arch. Eighteen out of the original 22 species range prediction models of Walls and Stigall (2011) yielded statistically significant results based on validation by secondary field work. Over 60% of field sites investigated for species occurrences were correctly predicted by the niche models, and model errors were more commonly underprediction errors (30% of sites) than overprediction errors (9% of sites). Several factors promoted higher accuracy among niche models. First, most of the species utilized in the study are easily identifiable and fairly widespread, such as *Zygospira modesta* and *Vinlandostrophia ponderosa.* During secondary field validation, these species were correctly identified at every site where the species were predicted to occur in the GARP models. Another key factor contributing to model success for particular species is the number of species occurrence data points used for constructing the original GARP models. Models produced with a larger amount of original occurrence data are more likely to be statistically significant.

Three species, *Vinlandostrophia cypha, Gnamptorhynchos auburnensis*, and *Rafinesquina alternata,* were shown to have significant model error. The poor predictive results of *V. cypha* and *G. auburnensis* can be explained by taxonomic confusion and undersampling. The validity of these taxa as biological species has been questioned, and they are likely to be locally prolific subspecies or community variations. Also, results of this study, as well as Raxworthy and others (2003), Hernandez and others (2006), and Guisan and others (2007), demonstrate that species with broad environmental tolerances can yield poorer statistical results than specialist species. *Rafinesquina alternata* is a widely distributed species for which two of the three models were not statistically significant. This may be explained by too limited a set of initial data. Widely distributed species may require more thorough sampling to yield statistically significant results. In this study, a sample size of 25 localities yielded significant results, whereas the models based on 17 and 22 locations were not significant.

Table 4. Statistical significance of original GARP models based on field validation. Area of prediction represents percentage of map area covered by 26 or more of the 50 best subset model predictions of Walls and Stigall (2011); *bold,* values that are not statistically significant (new).

Species	Area of prediction (%)	p value
Early time slice		
Dalmanella meeki	29%	0.012
Hebertella occidentalis	36%	0.002
Rafinesquina alternata	70%	**0.058**
Vinlandostrophia laticosta	42%	0.002
Vinlandostrophia ponderosa	36%	0.002
Zygospira modesta	29%	<0.001
Middle time slice		
Dalmanella meeki	32%	0.000
Hebertella occidentalis	44%	0.015
Rafinesquina alternata	52%	0.003
Gnamptorhynchos auburnensis	39%	**0.084**
Vinlandostrophia cypha	45%	0.023
Vinlandostrophia laticosta	25%	0.001
Vinlandostrophia ponderosa	63%	0.015
Zygospira modesta	46%	0.005
Late time slice		
Dalmanella meeki	33%	0.016
Hebertella occidentalis	43%	0.026
Rafinesquina alternata	59%	**0.071**
Gnamptorhynchos auburnensis	48%	0.032
Vinlandostrophia cypha	17%	**0.081**
Vinlandostrophia laticosta	56%	0.027
Vinlandostrophia ponderosa	54%	0.039
Zygospira modesta	47%	0.026

This study is the first direct test of the accuracy of paleontological environmental niche models. Even though ENM methods have been widely used with extant taxa and are increasingly utilized in paleontology, this study is the first to perform a *post hoc* validation of ENM models that utilized sedimentary data to infer environmental parameters. Results of this analysis validate ecological niche modeling in general and the GARP modeling system in particular as useful tools for paleontological investigation. It clearly demonstrates that utilizing fossil occurrence data (sourced from new field work, museum collections, or databases) with environmental parameters estimated from sedimentary parameters can provide a detailed quantitative assessment of paleobiogeographic patterns. These methods can provide a framework for researchers to study the biogeography of fossil organisms and assess how environmental changes have affected the evolution of life through time.

ACKNOWLEDGMENTS

We thank Miguel Araújo, Town Peterson, Greg Nadon, Greg Springer, and two anonymous reviewers for comments on an earlier version of this paper. B.J.W.'s research was funded by grants from the Geological Society of America, Paleontological Society, and the Ohio University Department of Geological Sciences. A.L.S.'s research was supported by NSF EAR-0922067, a grant from the donors of the American Chemical Society's Petroleum Research Fund, and an Ohio University Baker Fund Award.

REFERENCES

Alberstadt, L. P. 1979. The brachiopod genus *Platystrophia*. U.S. Geological Survey Professional Paper 1066-B:1–20.

Anderson, R. P., A. T. Peterson, & M. Gómez-Laverde. 2002. Using niche-based GIS modeling to test geographic predictions of competitive exclusion and competitive release in South American pocket mice. Oikos 98:3–16.

Araújo, M. B., & P. H. Williams. 2000. Selecting areas for species persistence using occurrence data. Biological Conservation 96:331–345.

Brett, C. E., & T. J. Algeo. 2001. Sequence stratigraphy of Upper Ordovician and lower Silurian strata of the Cincinnati Arch Region. *In* T. J. Algeo & C. E. Brett, eds., Sequence, cycle, & event stratigraphy of upper Ordovician & Silurian strata of the Cincinnati Arch region. Kentucky Geological Survey, Field Trip Guidebook, Series 12, Guidebook 1:34–46.

Chen, P., E. O. Wiley, & K. M. McNyset. 2007. Ecological niche modeling as a predictive tool: Silver and bighead carps in North America. Biologic Invasions 9:43–51.

Cummings, E. R., & A. V. Mauck. 1902. A quantitative study of variation in the fossil brachiopod *Platystrophia lynx*. American Journal of Science 14:9–16.

DeVaney, S. C., K. M. McNyset, J. B. Williams, A. T. Peterson, & E. O. Wiley. 2009. A tale of four "carp:" Invasion potential and ecological niche modeling. PLoS ONE 4:e5451.

Dudei, N. L., & A. L. Stigall. 2010. Using ecological niche modeling to assess biogeographic and niche response of brachiopod species to the Richmondian Invasion (Late Ordovician) in the Cincinnati Arch. Palaeogeography, Palaeoclimatology, Palaeoecology 296:28–43.

Elith, J., & C. H. Graham. 2009. Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models. Ecography 32:66–77.

Elith, J., C. H. Graham, Robert P. Anderson, Miroslav Dudík, Simon Ferrier, Antoine Guisan, Robert J. Hijmans, Falk Huettmann, John R. Leathwick, Anthony Lehmann, Jin Li, Lucia G. Lohmann, Bette A. Loiselle, Glenn Manion, Craig Moritz, Miguel Nakamura, Yoshinori Nakazawa, Jacob McC. M. Overton, A. Townsend Peterson, Steven J. Phillips, Karen Richardson, Ricardo Scachetti-Pereira, Robert E. Schapire, Jorge Soberón, Stephen Williams, Mary S. Wisz, & Niklaus E. Zimmermann. 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29:129–151.

ESRI Inc. 2006. ArcGIS 9.2 Release. Redlands, California.

Feria, T. P., & T. A. Peterson. 2002. Prediction of bird community composition based on point-occurrence data and inferential algorithms: A valuable tool in biodiversity assessments. Diversity and Distributions 8:49–56.

Ferrier, S., G. Watson, J. Pearce, & M. Drielsma. 2002. Extended statistical approaches to modeling spatial pattern in biodiversity in northeast New South Wales. I. Species-level modeling. Biodiversity and Conservation 11:2275–2307.

Fiaboe, K. K. M., R. F. Fonseca, G. J. Moraes, C. K. P. O. Ogol, & M. Knapp. 2006. Identification of priority areas in South America for exploration of natural enemies for classical biological control of *Tetranychus evansi* (Acari: Tetranychidae) in Africa. Biological Control 38:373–379.

Finnegan, S., & M. L. Droser. 2005. Relative and absolute abundance of trilobites and rhynchonelliform brachiopods across the Lower/Middle Ordovician boundary, eastern Basin and Range. Paleobiology 31:480–502.

Graham, C. H., S. R. Ron, J. C. Santos, C. J. Schneider, & C. Moritz. 2004. Integrating phylogenies and environmental niche models to explore speciation mechanisms in dendrobatid frogs. Evolution 58:1781–1793.

Guisan, A., N. E. Zimmerman, J. Elith, C. H. Graham, S. Phillips, & A. T. Peterson. 2007. What matters for predicting the occurrences of trees: Techniques, data, or species' characteristics? Ecological Monographs 77:615–630.

Hernandez, P. A., C. H. Graham, L. L. Master, & D. L. Albert. 2006. The effect of sample size and species characteristics on performance of different species distribution modeling methods. Ecography 29:773–785.

Holland, S. M. 2005. The signatures of patches and gradients in ecological ordinations. PALAIOS 20:573–580.

Holland, S. M., & M. E. Patzkowsky. 1996. Sequence stratigraphy and long-term oceanographic change in the Middle and Upper Ordovician of the eastern United States. *In* B. J. Witzke, G. A. Ludvigson, & J. E. Day, eds., Paleozoic Sequence Stratigraphy: Views from the North American Craton. Geological Society of America Special Paper 306:117–129.

Iverson, L. R., & A. M. Prasad. 1998. Predicting abundance of 80 tree species following climate change in the eastern United States. Ecological Monographs 68:465–485.

King, W. 1850. A monograph of the Permian Fossils of England. Palaeontographical Society (London), Monograph 3:1–258.

Kostelnick, J. C., D. L. Peterson, S. L. Egbert, K. M. McNyset, & J. F. Cully. 2007. Ecological niche modeling of the Black-tailed prairie dog habitats in Kansas. Transactions of the Kansas Academy of Science 110:187–200.

Maguire, K. C., & A. L. Stigall. 2009. Distribution of fossil horses in the Great Plains during the Miocene and Pliocene: An ecological niche modeling approach. Paleobiology 35(4):587–611.

Malizia, R. W., & A. L. Stigall. 2011. Niche stability in Late Ordovician articulated brachiopod species before, during, and after the Richmondian Invasion. Palaeogeography, Palaeoclimatology, Palaeoecology 311:154–170.

Martínez-Meyer, E., & A. T. Peterson. 2006. Conservatism of ecological niche characteristics in North American plant species over the Pleistocene-to-Recent transition. Journal of Biogeography 33:1779–1789.

Martínez-Meyer, E., A. T. Peterson, & W. W. Hargrove. 2004. Ecological niches as stable distributional constraints on mammal species, with implications for Pleistocene extinctions and climate change projections for biodiversity. Global Ecology and Biogeography 13:305–314.

Pearson, R. G., C. J. Raxworthy, M. Nakamura, & A. T. Peterson. 2007. Predicting species' distributions from small numbers of occurrence records: A test case using cryptic geckos in Madagascar. Journal of Biogeography 34:102–117.

Peterson, A. T. 2001. Predicting species' geographic distributions based on ecological niche modeling. The Condor 103:599–605.

Peterson, A. T. 2003. Predicting the geography of species' invasions via ecological niche modeling. Quarterly Review of Biology 78:419–433.

Peterson, A. T., & K. P. Cohoon. 1999. Sensitivity of distributional prediction algorithms to geographic data completeness. Ecological Modelling 117:159–164.

Peterson, A. T., M. Papeş, & M. Eaton. 2007. Transferability and model evaluation in ecological niche modeling: A comparison of GARP and MAXTENT. Ecography 30:550–560.

Peterson, A. T., M. Papeş, & J. Soberón. 2008. Rethinking receiver operating characteristic analysis applications in ecological niche modelling. Ecological Modelling 213:63–72.

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

Phillips, S. J., R. P. Anderson, & R. E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. Ecological Modelling 190(3–4):231–259.

Raxworthy, C. J., E. Martínez-Meyer, N. Horning, R. A. Nussbaum, G. E. Schneider, M. A. Ortega-Huerta, & A. T. Peterson. 2003. Predicting distributions of known and unknown reptile species in Madagascar. Nature 426:837–841.

Stigall, A. L. 2010. Using GIS to assess the biogeographic impact of species invasions on native brachiopods during the Richmondian Invasion in the type-Cincinnatian (Late Ordovician, Cincinnati region). Palaeontologia Electronica 13:5A, http://palaeo-electronica.org/2010_1/207/index.html.

Stigall, A. L. 2011. Application of niche modeling to analyze biogeographic patterns in Palaeozoic brachiopods: Evaluating niche stability in deep time. Memoirs of the Association of Australasian Palaeontologists 41:229–255.

Stigall, A. L., & B. S. Lieberman. 2006. Quantitative paleobiogeography: GIS, phylogenetic biogeographic analysis, and conservation insights. Journal of Biogeography 33:2051–2060.

Stigall Rode, A. L., & B. S. Lieberman. 2005. Using environmental niche modeling to study the Late Devonian biodiversity crisis. *In* D. J. Over, J. R. Morrow, & P. B. Wignall, eds., Understanding Late Devonian and Permian-Triassic Biotic and Climatic Events: Towards an Integrated Approach. Developments in Paleontology and Stratigraphy Series. Elsevier. Amsterdam. p. 93–180.

Stockman, A. K., D. A. Beamer, & J. E. Bond. 2006. Predicting the distribution of non-vagile taxa: A response to McNyset and Blackburn (2006) and re-evaluation of Stockman et al. (2006). Diversity and Distributions 12:787–792.

Stockwell, D. R. B., & I. R. Noble. 1992. Induction of sets of rules from animal distribution data: A robust and informative method of data analysis. Mathematical and Computer Simulation 33:385–390.

Stockwell, D. R. B., & D. Peters. 1999. The GARP modeling system: Problems and solutions to automated spatial prediction. International Journal of Geographic Information Systems 13:143–158.

Stockwell, D. R. B., & A. T. Peterson. 2002. Effects of sample size on accuracy of species distribution models. Ecological Modelling 148:1–13.

Tarver, J. E., S. J. Braddy, & M. J. Benton. 2007. The effects of sampling bias on Palaeozoic faunas and implications for macroevolutionary studies. Palaeontology 50:177–184.

Thomas, C. D., A. Cameron, R. E. Green, M. Bakkenes, L. J. Beaumont, Y. C. Collingham, B. F. N. Erasmus, M. Ferreira de Siquira, A. Grainger, L. Hannah, L. Hughes, B. Huntley, A. S. van Jaarsveld, G. F. Midgley, L. Miles, M. A. Ortega-Huerta, A. T. Peterson, O. L. Phillips, & S. E. Williams. 2004. Extinction risk from climate change. Nature 427:145–148.

Thuiller, W., D. M. Richardson, P. Pysek, G. F. Midgley, G. O. Hughes, & M. Rouget. 2005. Niche-based modeling as a tool for predicting the global risk of alien plant invasions at a global scale. Global Change Biology 11:2234–2250.

Walls, B. J., & A. L. Stigall. 2011. Analyzing niche stability and biogeography of Late Ordovician brachiopod species using ecological niche modeling. Palaeogeography, Palaeoclimatology, Palaeoecology 299:15–29.

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

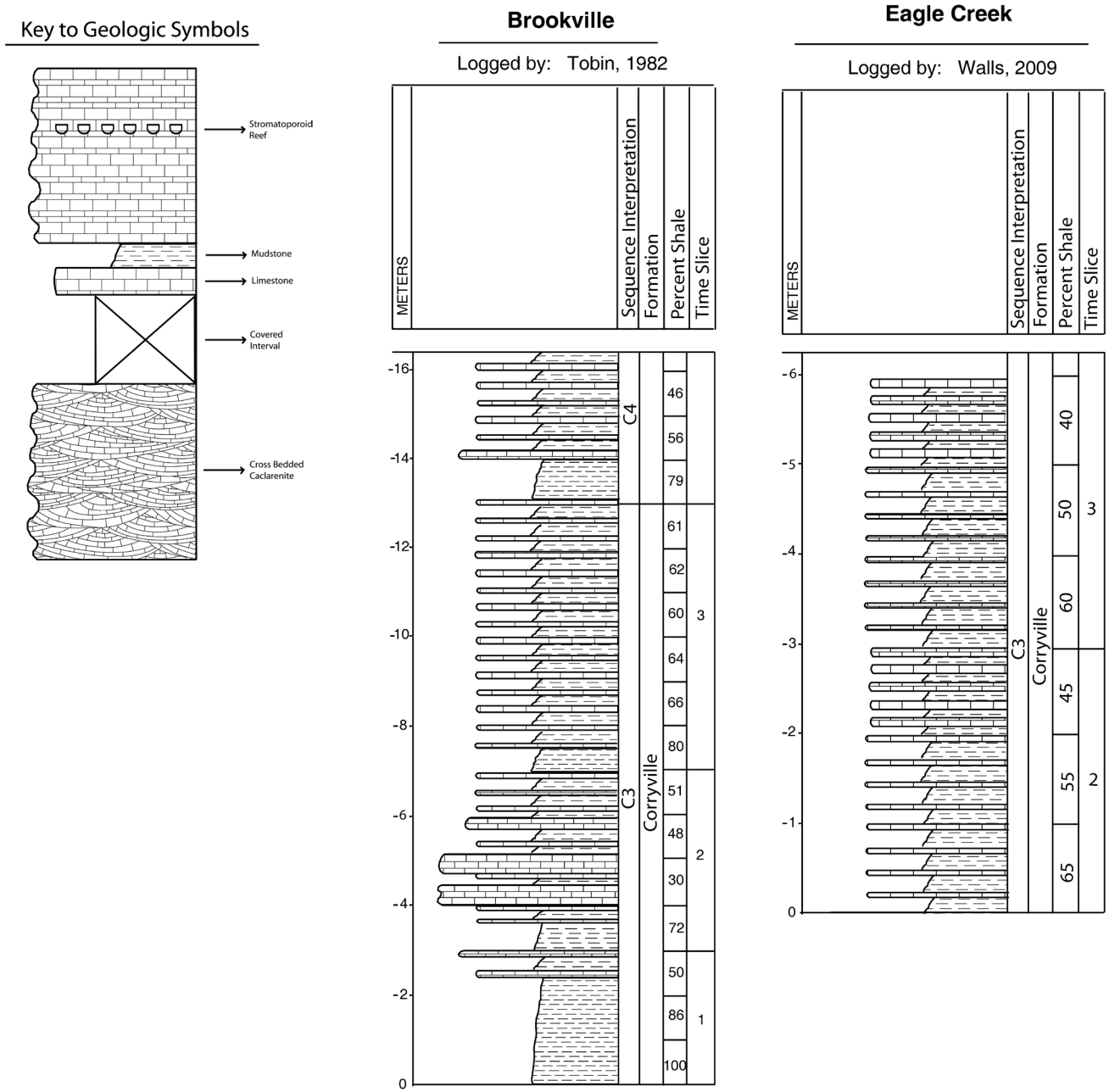
Westrop, S. R., & J. M. Adrain. 2001. Sampling at the species level; impact of spatial biases on diversity gradients. Geology 29:903–906.

Zuykov, M. A., & D. A. T. Harper. 2007. *Platystrophia* (Orthida) and new related Ordovician and Early Silurian brachiopod genera. Estonian Journal of Earth Science 56:11–34.

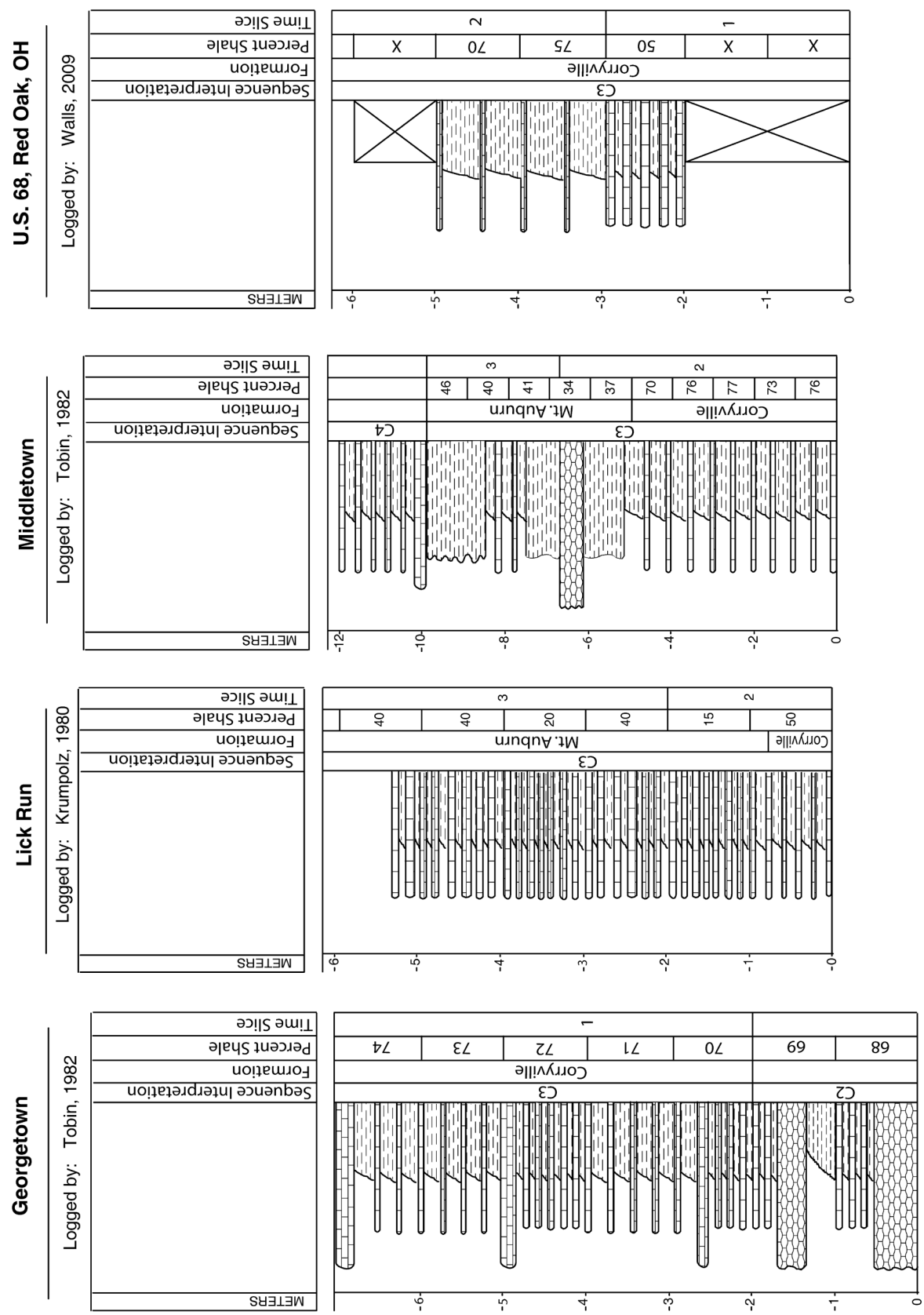
Appendix 1. Field validation locality directory; *, sites based on Ohio University Paleontology Collection; *IN,* Indiana; *KY,* Kentucky, *OH,* Ohio; other locations are sites of field investigation.

Site	Longitude	Latitude	Location
Brookville Dam, IN	-85.10	39.44	Spillway below Brookville Dam, north of Brookville, IN
Carrolton, KY*	-85.20	38.68	Two cuts on US 42, west of Carrolton
Eagle Creek, OH	-83.75	38.85	Exposures on unnamed tributary of West Fork Eagle Creek, about 415 m (1700 ft) south of junction of Ohio Route 125 and Dr. Faul Road, Russellville, OH
Faber Ave., Cincinnati, OH	-84.58	39.14	Faber Ave
Georgetown, OH	-83.92	38.87	Road cuts along both sides of Ohio Route 125, on both sides of White Oak Creek Valley, 0.8 mi (1.3 km) west of intersection with US Route 68 in Georgetown, OH
Kentucky Drive, Carrolton, KY*	-84.63	38.98	US 42, west of Carrolton, second cut
Lebanon, OH*	-84.20	39.44	Stream exposure in creek along Ohio State Route 63, west of Ohio 741, east of Monroe, OH
Lick Run, Osceloa, OH	-84.05	39.30	Stream exposures along Lick Run west of Ohio Route 132 bridge and 760 m (2500 ft) north of Blackhawk OH, at junction of Ohio Routes 132 and 123
Middletown, OH	-84.42	39.48	Road cuts along both sides of Ohio Route 4, just south of bridge across Dicks Creek, about 0.5 km (0.3 mi) south of junction with Ohio Route 74, Middletown, OH
Power Line Drive, Florence, KY*	-84.60	38.98	Series of cuts near Foundation Drive off Industrial Rd., off US 42, east of exit 180, cut at dead end
Red Oak North, OH*	-83.85	38.84	US 68 approx 2 mi south of US Route 125 from Georgetown, OH
Rising Sun, IN*	-84.89	38.92	IN 56 West out of Rising Sun to intersection with IN 156, stay on IN 156 (right turn) about 1 mi to exposure near top of hill
S. Loop Road, Florence, KY*	-84.56	39.01	Thomas Moore Parkway, ~200 yards north of intersection with Med Village Dr., ~ 100 yards south of chancellor drive
Sheits Road, Dornbusch, OH	-84.62	39.25	Exposures along two branches of an unnamed stream: branch that follows Sheits Road, just west of intersection of stream and I-275 near Dornbush, OH
South Road, Mack, OH*	-84.67	39.15	1/4 mile from South Rd. near Mack, OH
Thomas Moore Parkway, Florence, KY*	-84.55	39.01	Intersection of Thomas Moore Parkway and Dudley
US 68, Red Oak, OH*	-83.84	38.83	2.8 miles south of St. Rt. 125 from Georgetown, OH
Wal-Mart, Hamilton, OH	-84.60	39.43	Cut behind Wal-Mart. Located on Ohio 177 just SE of intersection with Ohio 130, northwest of Hamilton, OH

Appendix 2. Stratigraphic columns for field validation sites. The stratigraphic columns in this appendix consist of sites visited for field validation of original ecological niche models. These sites are a combination of newly identified C3 localities and sites of previously known C3 outcrops, for which either no previously published section existed or published sections lacked the detail required for this project.



Appendix 2. *Continued.*



Appendix 3. Presence/absence data by time slice; *P,* predicted species occurrence; *N,* Predicted species absence; *I,* Identified; *A,* Absent; *X,* museum site, no occurrence data; *IN,* Indiana; *KY,* Kentucky, *OH,* Ohio.

Early time slice

Locality	*Dalmanella meeki*	*Hebertella occidentalis*	*Rafinesquina alternata*	*Gnamptorhynchos laticosta*	*Vinlandostrophia ponderosa*	*Zygospira modesta*
Carrolton, KY	X	P, I	P, I	N, I	P, I	P, I
Faber Ave, Cincinnati, OH	X	X	N, I	X	X	X
Georgetown, OH	N, I	P, A	X	N, I	X	N, I
Hamilton, Wal-Mart, OH	N, I	P, A	N, I	P, A	N, A	N, I
Red Oak North, OH	X	P, I	P, I	N, I	P, I	N, I
Rising Sun, IN	X	N, I	P, I	P, I	N, I	N, I
Sheits Road, Dornbusch, OH	N, A	N, A	N, I	P, A	N, A	P, I
South Road, Mack, OH	P, I	N, I	N, I	P, I	N, I	P, I
US 68, Red Oak, OH	N, A	P, I	P, I	N, I	P, I	N, I

Middle time slice

Locality	*Dalmanella meeki*	*Hebertella occidentalis*	*Rafinesquina alternata*	*Gnamptorhynchos auburnensis*	*Vinlandostrophia cypha*	*Vinlandostrophia laticosta*	*Vinlandostrophia ponderosa*	*Zygospira modesta*
Brookville Dam, IN	N, I	N, A	N, I	N, A	N, A	N, A	N, A	N, I
Eagle Creek, OH	N, A	N, I	N, I	P, A	N, A	N, I	P, I	N, A
Hamilton, Wal-Mart, OH	N, I	P, A	N, I	N, A	N, A	N, A	N, I	N, I
Kentucky Drive, Carrolton, KY	N, A	P, I	P, I	P, A	N, A	N, I	P, I	P, I
Lebanon, OH	N, I	X	X	X	X	X	X	X
Lick Run, OH	X	X	P, I	X	X	X	P, I	X
Power Line Drive, Florence, KY	P, I	P, A	P, I	P, A	P, A	P, I	P, I	P, I
South Loop Road, Florence, KY	P, I	P, I	P, I	X	P, I	P, I	P, I	P, I
Thomas Moore Parkway, Florence, KY	X	X	P, I	X	X	P, I	P, I	X
US 68, Red Oak, OH	N, A	P, I	P, I	P, A	N, A	N, I	P, I	P, I

Late time slice

Locality	*Dalmanella meeki*	*Hebertella occidentalis*	*Rafinesquina alternata*	*Gnamptorhynchos auburnensis*	*Vinlandostrophia cypha*	*Vinlandostrophia laticosta*	*Vinlandostrophia ponderosa*	*Zygospira modesta*
Brookville Dam, IN	N, I	N, A	N, I	N, A	N, A	N, A	N, A	N, I
Eagle Creek, OH	N, I	N, I	N, I	P, A	N, A	N, A	N, I	N, A
Hamilton, Wal-Mart, OH	N, A	N, A	P, I	P, A	P, A	P, I	P, I	P, I
Lick Run, OH	P, A	P, I	P, I	N, I	N, I	P, I	P, I	P, I
Middletown, OH	N, I	P, I	P, I	N, I	N, A	P, A	P, I	P, I