Geospatial modeling approach to monument construction using Michigan from A.D. 1000-1600 as a case study

Howey, M.C.L.; Palace, M.W.; McMichael, C.H.

DOI
10.1073/pnas.1603450113

Publication date
2016

Document Version
Final published version

Published in
Proceedings of the National Academy of Sciences of the United States of America

License
Other

Citation for published version (APA):
Geospatial modeling approach to monument construction using Michigan from A.D. 1000–1600 as a case study

Meghan C. L. Howey\textsuperscript{a,b,1}, Michael W. Palace\textsuperscript{b,c}, and Crystal H. McMichael\textsuperscript{d,e}

\textsuperscript{a}Department of Anthropology, University of New Hampshire, Durham, NH 03824; \textsuperscript{b}Earth Systems Research Center, Institute for the Study of Earth, Oceans, and Space, University of New Hampshire, Durham, NH 03824; \textsuperscript{c}Department of Earth Sciences, University of New Hampshire, Durham, NH 03824; \textsuperscript{d}Paleoecology and Landscape Ecology, Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, 1090 GE, Amsterdam, The Netherlands; and \textsuperscript{e}Department of Biological Sciences, Florida Institute of Technology, Melbourne, FL 32901

Edited by James A. Brown, Northwestern University, Evanston, IL, and approved May 17, 2016 (received for review March 1, 2016)

Building monuments was one way that past societies reconfigured their landscapes in response to shifting social and ecological factors. Understanding the connections between those factors and monument construction is critical, especially when multiple types of monuments were constructed across the same landscape. Geospatial technologies enable past cultural activities and environmental variables to be examined together at large scales. Many geospatial modeling approaches, however, are not designed for presence-only (occurrence) data, which can be limiting given that many archaeological site records are presence only. We use maximum entropy modeling (MaxEnt), which works with presence-only data, to predict the distribution of monuments across large landscapes, and we analyze MaxEnt output to quantify the contributions of spatioenvironmental variables to predicted distributions. We apply our approach to co-occurring Late Precontact (ca. A.D. 1000–1600) monuments in Michigan: (i) mounds and (ii) earthwork enclosures. Many of these features have been destroyed by modern development, and therefore, we conducted archival research to develop our monument occurrence database. We modeled each monument type separately using the same input variables. Analyzing variable contribution to MaxEnt output, we show that mound and enclosure landscape suitability was driven by contrasting variables. Proximity to inland lakes was key to mound placement, and proximity to rivers was key to sacred enclosures. This juxtaposition suggests that mounds met local needs for resource procurement success, whereas enclosures filled broader regional needs for intergroup exchange and shared ritual. Our study shows how MaxEnt can be used to develop sophisticated models of past cultural processes, including monument building, with imperfect, limited, presence-only data.

Significance

Monumental construction was one way that past societies reconfigured their landscapes in response to social and environmental factors. Given that these factors changed through space and time, societies often constructed multiple types of monuments. We developed a maximum entropy modeling (MaxEnt) framework for identifying the spatioenvironmental factors that mattered most in the placement of different monument types. We turned to MaxEnt, given its ability to model spatial distributions from presence-only data. With modern development continuing to destroy archaeological sites across the globe, archaeological datasets, including monument locations, are evermore restricted to being presence-only data. Our framework shows how archaeologists can harness MaxEnt to develop robust models of past cultural processes, including monumentality, even when faced with limited, presence-only data.

Author contributions: M.C.L.H. and M.W.P. designed research; C.H.M. performed research; M.W.P. analyzed data; and M.C.L.H. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

\textsuperscript{1}To whom correspondence should be addressed. Email: meghan.howey@unh.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1603450113/-/DCSupplemental.
to the distribution of mounds and enclosures. Doing this analysis, we identified key differences between the spatioenvironmental variables that mattered most to mound and enclosure landscape placement. These differences extend the view that precontact indigenous peoples in Michigan used these two types of monuments to answer different but simultaneously pressing social, economic, and ideological demands: mounds for local ones and enclosures for regional ones.

**MaxEnt Framework for the Study of Monumentality**

Geospatial technologies have made connecting cultural activities and environmental variables at large scales more feasible. Archaeologists have increasingly looked toward advanced geospatial modeling to investigate past landscapes (6, 7). Many available geospatial modeling approaches, however, are not designed for presence-only (occurrence) data and therefore, present problems for archaeologists. Archaeologists often know where some but not all sites were, and they rarely know for certain where past sites are absent. Given that many archaeological site records are occurrence data, including monument locations, archaeologists need modeling approaches that can tie environmental and geospatially derived data layers together with presence-only data. MaxEnt is one such approach.

The principle of maximum entropy originates in statistical mechanics (8), but its application has been expanded to many disciplines. MaxEnt has been especially developed as a species distribution modeling approach in ecology used for predicting habitat suitability of plant and animal species from presence-only records (9–13). MaxEnt allows for any number of continuous or discrete spatioenvironmental factors to be used as predictor variables for the species of interest (11–13). Predictor variable values are extracted from occurrence locations and randomly located background points across the sampling area to develop a statistical model that extrapolates a species habitat probability map across the landscape (15). Paleodistribution modeling, where floral and faunal species distributions are hindcast, has also been explored during the past decade (14, 15).

Such paleodistribution modeling has recently been appreciated as a potentially powerful tool for understanding human prehistory (14). MaxEnt emerges as a particularly useful approach for archaeological applications, because it commonly outperforms other modeling techniques that use presence-only data, especially when dealing with small sample sizes or nonrandom sampling (16). Accordingly, MaxEnt has increasingly been used in creating predictive models of the distribution of past human activities (16–20). In this research, MaxEnt has been used to identify shared spatioenvironmental parameters among the locations of known archaeological sites/features and extrapolate other geographic locales where these parameters are present. Archaeologists have used these outputs to examine and compare the total ranges of culture groups adaptations, identify ecocultural niches (17, 18), and build predictive models of archaeological features (16, 19, 20).

These recent uses of MaxEnt modeling have proven to be robust and informative. We explore how MaxEnt distribution modeling can be harnessed in additional ways for archaeological questions about past cultural processes, including monument construction and placement. Extending the ecological metaphor from MaxEnt’s use in species distribution modeling, monument construction can be seen as a cultural response to the specific mix of environmental, spatial, and social variables in a given landscape. Archaeologists can input archaeological monument occurrence data and any number of relevant spatioenvironmental variables into MaxEnt to create a probability map of monument location (in essence, a “monument habitat” map). The cultural process of monument building then becomes the “species” whose distribution is predicted. This framework shares some logic with the recent work by Saatchi et al. (21) to estimate aboveground biomass, wherein biomass classes were broken into presence data, and each biomass species was modeled across the landscape. After a monument habitat distribution is established, the relative contribution of each included spatioenvironmental predictor variable in determining overall distribution can be statistically analyzed to determine which variables mattered most in monument landscape placement.

We suggest that using MaxEnt modeling in this way is powerful for examining the frequently occurring cases where regions host multiple types of monuments. Each type of monument in the landscape can be separated into its own occurrence dataset and run through the model using the same set of spatioenvironmental variables. The contributions of the input variables for each type of monument can be quantified. Similarities and differences in the contributing factors to the geographic distribution of each type of monument can then examined; this offers a robust means of understanding how monuments were built in response to varied socioecological demands.

**Mounds and Earthwork Enclosures of Michigan**

Mounds and earthworks are ubiquitous features across the eastern United States, and they have long been the focus of popular speculation, archaeological investigation, and post–Euro-American destruction (22, 23). These earthen constructions represent some of the most interesting examples of monument building anywhere in the world (2, 5, 24–27). Although overshadowed by striking sites in neighboring parts of the eastern United States, like Cahokia in Illinois and the Newark Earthworks in Ohio, earthen constructions are still a major part of Michigan’s archaeological record.

Michigan displays two different kinds of earthen constructions: burial mounds and earthwork enclosures (Fig. 1). Notable burial mound complexes were built in southwestern Michigan as part of Middle Woodland Hopewellian traditions (28). Less elaborate, normal use conical burial mounds became geographically widespread features across Michigan’s Lower Peninsula during the Late Woodland Period and remained so through Late Prehistory (after ca. A.D. 800–1000 to A.D. 1600) (29) (Fig. 1). Earthwork enclosures, large-scale geometric earthen features, were less common than mounds, reaching their highest density in southern Michigan and continuing into northern Indiana and Ontario (30–37) (Fig. 1). These constructions almost always date to Late Prehistory (ca. A.D. 1200–1600). Michigan’s enclosures share a common form of being circular ditch and embankment enclosures with planned entryways. Depending on geographic and temporal factors, they range from 30 to 120 m in diameter (38).

The later Late Woodland/Late Prehistoric Period (referred to here as Late Precontact; ca. A.D. 1000–1600) saw the co-occurrence
of these two types of earthen constructions. A variety of perspectives has been developed about these constructions, and debate continues, particularly about the nature of enclosures (30, 35, 38, 39). A striking aspect of the research on mounds and enclosures is that, despite being shared features across Michigan’s forager-fisher-horticulturalists during Late Precontact, they are not typically investigated in conjunction with each other. However, it is the presence of both together that suggests that monument building was a particularly important part of life in Late Precontact Michigan.

Monuments are not superfluous to socioeconomic developments; they are intertwined with every aspect of community organization. These mounds and enclosures represent then a reconfiguration of the Late Precontact landscape tied to social, economic, and ideological developments (5). Here, we ask whether mounds and enclosures were placed in contrasting positions in the landscape in response to different spatioenvironmental variables and what insights these contrasts provide about their roles during Late Precontact.

Results and Discussion

The geospatial database of mounds and enclosures that we created (by working with archival archaeological files in Michigan) has 321 records: 60 enclosure locations and 261 mound locations (multiple enclosures and mounds are often found in one locale) (Fig. 1). Today, only 13 of these enclosure locations and 45 of these mound locales have a high certainty of still existing (that is, a destruction rate—since their recording in the late 1800s and early 1900s—of 78.33% for enclosures and 82.76% for mounds). These staggering rates of destruction show that it is imperative to use archives of early archaeological site records.

We conducted MaxEnt with 13 spatioenvironmental input variables (covariates) relevant to earthen construction location suitability and presence—only occurrence records of Late Precontact mounds/enclosures created through archival research. Covariates included temperature and precipitation regimes, elevation, and distance to key hydrographic features, representing a range of factors that impacted the subsistence practices, social developments, and settlement patterns of the region’s mixed fisher–forager–horticulturalists. Using modern temperature and precipitation covariates carries limitations for archaeological modeling. In recognition of these limitations, we did not rely on these as actual values for Late Precontact but instead used them to understand how areas across the landscape would have varied in temperature and precipitation relative to each other (i.e., warmer, wetter areas today would have been warmer, wetter areas during Late Precontact, even if absolute temperature and precipitation values have changed). The output produced a mound “habitat” suitability map and an enclosure habitat suitability map for all of Michigan (Fig. 2).

Visual inspection of these maps highlights that mound and enclosure features occupied, to extend the ecological metaphor, distinct “niches” in the landscape (Fig. 2). To quantify and compare the relative importance of input spatioenvironmental variables to the suitability of places for mounds and enclosures across the study area, we: (i) computed a table of variable overall contribution to MaxEnt model fit, (ii) ran jackknife tests for the training and testing datasets, and (iii) created individual variable response curves (40) (SI Text has expanded explanations and results of these analyses). Note that we also conducted an independent analysis of the discriminating power of input variables with nominal logistic regression and computed a multivariate correlation table for all input variables (presented in SI Text).

For mounds, two variables dominated the overall contribution to MaxEnt model fit: mean temperature of the warmest quarter (BIO10) and distance to an inland lake (61.9% together) (Table 1). In test and training jackknife tests, annual mean temperature (BIO1) very closely followed by mean temperature of the warmest quarter (BIO10) show the highest gains in isolation (SI Text). These two variables have the most useful information when used by themselves for predicting mound distribution (note that they are strongly positively correlated) (SI Text). The variable that decreases the gain the most when it is omitted from the model is distance to an inland lake; this variable contains the most information about mound distribution not present in the other variables (test results in SI Text). The response curve of distance to an inland lake shows that it is mound proximity to an inland lake that is the key to predicted suitability (Fig. 3).

For enclosures, annual mean temperature (BIO1) and distance to a river made the largest contributions to MaxEnt model fit (48.5% together), with topography (Shuttle Radar Topography Mission, SRTM) a notable third as the only other double-digit contributor (Table 1). In both test and training jackknife tests, the environmental variable with the highest gain when used in isolation is annual mean temperature (BIO1). The variable that decreases the gain the most when it is omitted is distance to a river (test results in SI Text). The response curve of distance to a river shows clearly that it is enclosure proximity to a river that is the key to predicted suitability (Fig. 3).

The juxtaposition of the importance of inland lakes to mounds and rivers to enclosures is striking. Michigan’s 11,037 inland lakes are localized hydrographic features positioned in single ecosystems covering small to moderate areas and having relatively circumscribed watersheds. Rivers, in contrast, are regional hydrographic features, traversing substantial distances, crossing ecosystems, and having large watersheds. These features seem then positioned to have distinct geographic reaches—local and regional, respectively. These results can be understood within the context of Late Precontact.

Although there is no consensus interpretation for Late Precontact, it is generally agreed that, after ca. A.D. 1000, the region saw an intensification of a mixed economy involving the variable adoption of maize horticulture, targeted fishing of seasonal spawns, and intensified use of wild plant foodstuffs (particularly those amenable to storage), such as acorns, berries, maple syrup, and wild rice (41–43). Spatial proximity became increasingly important with these economic shifts, and across Michigan, communities found themselves engaging in their annual mobility rounds within more restricted territories. With the (wild and cultivated) resource base of the region being both seasonally variable and unpredictable year to year, the reduction in the spatial range that communities could exploit to harvest food heightened the risk of experiencing resource failure and scarcity.

We have argued that, within this context, dual needs arose to (i) maximize and extend the use life of local resources and (ii) increase interaction with other communities occupying different territories and resource bases (5, 39, 43, 44). As introduced above, we have suggested that the construction of mounds filled local needs and that the construction of enclosures filled more
Mean temperature of the warmest quarter (BIO10) 41.9* 0.1
Distance to an inland lake 19.6* 3.8
Mean temperature of the driest quarter (BIO9) 11.6 7.9
Distance to a river 10.9 22.9*
SRTM topography 3.7 13.5
Mean temperature of the wettest quarter (BIO8) 3.6 4.1
Annual mean temperature (BIO1) 3.2 25.6*
Precipitation of the wettest month (BIO13) 3 8.4
Temperature seasonality (BIO4) 1.1 1.3
Mean diurnal range (BIO2) 0.6 4.6
Temperature annual range (BIO7) 0.4 0
Annual precipitation (BIO12) 0.4 3.9
Precipitation seasonality (BIO15) 0.2 3.9

*Top two contributors.

Table 1. The relative contribution of each included spatioenvironmental predictor variable in determining the overall fit of the MaxEnt model for mounds and enclosures (normalized to percentages)

Enclosures and Rivers. Despite being the largest features in Michigan’s archaeological landscape, enclosures have been subject to surprisingly little systematic investigation. They were first assumed to be fortifications, but work across northern and southern located enclosures has shown this to be an insufficient explanation. Today, instead, many agree that enclosures were special use, sacred spaces, but whether these were local or regional in nature is debated (30, 34–39).

Across our MaxEnt variable contribution analyses, we found that proximity to regional hydrographic features—rivers— mattered in enclosure distribution (Figs. 2 and 3 and Table 1). Canoe travel along Michigan’s major rivers enabled people to transport themselves and goods long distances (48). Enclosures were placed where there was access to regional transportation corridors. Looking at enclosures and mounds simultaneously helps highlight the fact that enclosure placement was meant to extend beyond the local. Looking at each monument type alone, we would miss their juxtaposition in the landscape. The fact that enclosures’ prime habitat proximate to rivers indicates a regional orientation becomes a more compelling conclusion when we see that the prime habitat for mounds is very different, oriented around local hydrographic features (inland lakes).

What was the purpose of such enclosure positioning? As we discussed above, Late Precontact communities across Michigan faced heightened risks of resource scarcity. Rivers, then, formed important avenues through which communities could increase their interaction with groups occupying different territories and resource bases. Such nonlocally circumscribed interaction would create opportunities for the trade of physical foodstuffs, providing immediate scarcity relief, while also securing future relief through the furthering of social relationships. In small-scale societies, like those of Late Precontact, the pooling of risk through sharing and storing social obligations, often through the creation and/or extension of kin relationships among dispersed communities, is a well-documented risk-buffering strategy (49). By placing enclosures (sacred monumental constructions), near rivers, communities provided designated, attractive, and accessible spaces for interaction, trade, and relationship building.

We have argued for the regional ceremonial importance of enclosures previously (5, 39). Some recent research emphasizes the importance of enclosures as locally embedded features, noting that, although they may share a circular form, they each have their own distinct layout and material signatures, reducing prospects that they had broader regional orientations (30, 35). In both our previous research and these recent arguments, the regional and/or local import of enclosures has been presented as an “either/or” scenario, and this is unrealistic. Enclosures, even as features visited by neighboring communities, were still built and remained perpetually present in a local landscape.

We found, beyond river proximity, that other, more localized variables also mattered in enclosure distribution. Enclosures were positioned in places that were generally warmer than surrounding locales (Fig. 3 and Table 1). Topography was also a notable predictor of enclosure suitability (and not mound) (Table 1). Looking at the response curve, enclosures were not just built on higher ground, but rather, they were built where the topography offered a kind of “sweet spot” higher than surrounding locales but not too high (Fig. 3). This kind of nuanced topographic placement may have been driven by how local communities wanted to connect these sacred constructions to their own local production of place, histories, and memories (35), which may, perhaps, have given local groups a level of “ownership” (spiritual and/or territorial) over specific enclosure locales, even as these were part of broader regional trends.

Conclusions

Monument building was one way that past societies reconfigured their landscapes. Understanding how such reconfigurations affect
We used MAXENT, v3.3.3 (www.cs.princeton.edu/~schapire/maxent/) to model mound and enclosure distributions. Spatioenvironmental data were acquired or derived from the state of Michigan and, if necessary, resampled to a 90-m spatial resolution (in ArcGIS 10.2). Ten WorldClim Bioclimatic variables, which include precipitation (centimeters) and temperature (degrees Celsius) parameters across the globe at a 1-km resolution, were used as climatic predictors (www.worldclim.org/bioclim). Topography (SRTM 30 m), distance to an inland lake, and distance to a river were other input variables (derived from lake and hydrography layers from Michigan’s GIS Open Data Portal: gis.michigan.opendata.arcgis.com).

We performed 25-fold cross-validation of the model for mounds and 6-fold cross-validation of the model for enclosures. Folds were related to the total numbers of occurrence records. Mounds, with more occurrences, were split randomly into 25 equal-sized groups called “folds,” and the model was replicated 25 times, each time leaving out 1-fold, to test the performance of the model (enclosures with 6-fold). This process was advantageous, because all of the data points were used for model training and also, performance testing. The evaluated area under the curve (AUC) statistic, which indicates the predictive capacity of a model, and how much better (or worse) a given model performed compared with a random model. A model with AUC value greater than 0.75 can be considered as predicting the distribution of test points accurately (40). For mounds, the average across all runs AUC = 0.758, and for enclosures, the average across all runs AUC = 0.852, indicating that the model fit the data well and had a high predictive capacity (SI Text).

We conducted analyses to assess the relative importance of input spatioenvironmental variables to the predicted geospatial suitability of each monument type (40). An estimate of relative contribution of the spatioenvironmental variables to MaxEnt model fit was derived. In each of the training algorithm, the increase in regularized gain was added to the contribution of the corresponding variable or subtracted from it if the change to the absolute value of $\lambda$ was negative (Table 1) (normalized to percentages). Jackknife tests for training and testing datasets were done. Training gain values were derived from points used to train the model, and test gain values were derived from occurrence points used to test the predictive capacity of the model. In running jackknife tests, each variable is excluded in turn, and a model was created with the remaining variables. Then, a model is created using each variable in isolation. Using these jackknife tests, we identified for each monument type which input variable produced the largest gain value when it was excluded and which one produced the largest gain value when it was used as a single predictor (SI Text). A response curve using each input variable alone for the MaxEnt model for mounds.

Fig. 3. Response curves of key variables for mounds and enclosures (all variable curves are in SI Text). These curves were created using each input variable alone for the MaxEnt model for mounds and enclosures, which shows the dependence of predicted suitability as it patterns within each model. In running jackknife tests, each variable is excluded in turn, and a model was created with the remaining variables. Then, a model is created using each variable in isolation. Using these jackknife tests, we identified for each monument type which input variable produced the largest gain value when it was excluded and which one produced the largest gain value when it was used as a single predictor (SI Text). A response curve using each input variable alone for the MaxEnt model for mounds.

Howey et al.

Materials and Methods

Enclosure and Mound Database. We digitized mound and enclosure locations for the entire state of Michigan (which are found only in the Lower Peninsula) (Fig. 1) working through three archival sources: (i) US Geological Survey maps, which locate sites (housed at the Michigan State Archaeologist Office); (ii) hanging state archaeological site files for each of 83 counties in Michigan at the University of Michigan Museum of Anthropology’s Great Lakes Range with associated historic documents, such as letters, maps, and photographs from concerned Michigan residents about mound/enclosure destruction; and (iii) Hindsdale’s (32) very own annotated version of a copy of his Archaeological Atlas of Michigan (an invaluable resource available only at the University of Michigan Museum of Anthropology’s Great Lakes Range). Working across these three sources, we produced a geospatial database of reported mounds and enclosures without duplication. The database includes only constructions with some certainty of existence as determined from available historic records and other associated information in site files. Site locations were recorded by the quarter section and heads up digitized using the Michigan Geographic Framework available at Michigan’s Geographic Information Systems (GIS) Open Data Portal to reference section locations (work done in ArcGIS 10.0). Mounds and enclosures in this database with clear non-Late Precontact Period affiliations were excluded from analyses.

MaxEnt. We used MAXENT, v3.3.3 (www.cs.princeton.edu/~schapire/maxent/) to model mound and enclosure distributions. Spatioenvironmental data were acquired or derived from the state of Michigan and, if necessary, resampled to a 90-m spatial resolution (in ArcGIS 10.2). Ten WorldClim Bioclimatic variables, which include precipitation (centimeters) and temperature (degrees Celsius) parameters across the globe at a 1-km resolution, were used as climatic predictors (www.worldclim.org/bioclim). Topography (SRTM 30 m), distance to an inland lake, and distance to a river were other input variables (derived from lake and hydrography layers from Michigan’s GIS Open Data Portal: gis.michigan.opendata.arcgis.com).

We performed 25-fold cross-validation of the model for mounds and 6-fold cross-validation of the model for enclosures. Folds were related to the total numbers of occurrence records. Mounds, with more occurrences, were split randomly into 25 equal-sized groups called “folds,” and the model was replicated 25 times, each time leaving out 1-fold, to test the performance of the model (enclosures with 6-fold). This process was advantageous, because all of the data points were used for model training and also, performance testing. The evaluated area under the curve (AUC) statistic, which indicates the predictive capacity of a model, and how much better (or worse) a given model performed compared with a random model. A model with AUC value greater than 0.75 can be considered as predicting the distribution of test points accurately (40). For mounds, the average across all runs AUC = 0.758, and for enclosures, the average across all runs AUC = 0.852, indicating that the model fit the data well and had a high predictive capacity (SI Text).

We conducted analyses to assess the relative importance of input spatioenvironmental variables to the predicted geospatial suitability of each monument type (40). An estimate of relative contribution of the spatioenvironmental variables to MaxEnt model fit was derived. In each of the training algorithm, the increase in regularized gain was added to the contribution of the corresponding variable or subtracted from it if the change to the absolute value of $\lambda$ was negative (Table 1) (normalized to percentages). Jackknife tests for training and testing datasets were done. Training gain values were derived from points used to train the model, and test gain values were derived from occurrence points used to test the predictive capacity of the model. In running jackknife tests, each variable is excluded in turn, and a model was created with the remaining variables. Then, a model is created using each variable in isolation. Using these jackknife tests, we identified for each monument type which input variable produced the largest gain value when it was excluded and which one produced the largest gain value when it was used as a single predictor (SI Text). A response curve using each input variable alone for the MaxEnt model for mounds.

Our research also shows MaxEnt modeling as a robust tool for working with archival archaeological site databases. Because of long experienced intensive rates of destruction, the mounds and earthworks of the eastern United States may have experienced pressures of archaeological site destruction earlier than other parts of the globe, many places have been subject to similar levels of development, and many more places are increasingly subject to such pressure; this means that archaeologists will increasingly have to turn to already established, presence-only site records, with their inherent limitations, to understand the past in their study areas. As they do, MaxEnt will continue to offer a means of developing sophisticated models of myriad past human activities, including monument building, from imperfect, limited, presence-only data.

We performed 25-fold cross-validation of the model for mounds and 6-fold cross-validation of the model for enclosures. Folds were related to the total numbers of occurrence records. Mounds, with more occurrences, were split randomly into 25 equal-sized groups called “folds,” and the model was replicated 25 times, each time leaving out 1-fold, to test the performance of the model (enclosures with 6-fold). This process was advantageous, because all of the data points were used for model training and also, performance testing. The evaluated area under the curve (AUC) statistic, which indicates the predictive capacity of a model, and how much better (or worse) a given model performed compared with a random model. A model with AUC value greater than 0.75 can be considered as predicting the distribution of test points accurately (40). For mounds, the average across all runs AUC = 0.758, and for enclosures, the average across all runs AUC = 0.852, indicating that the model fit the data well and had a high predictive capacity (SI Text).

We conducted analyses to assess the relative importance of input spatioenvironmental variables to the predicted geospatial suitability of each monument type (40). An estimate of relative contribution of the spatioenvironmental variables to MaxEnt model fit was derived. In each of the training algorithm, the increase in regularized gain was added to the contribution of the corresponding variable or subtracted from it if the change to the absolute value of $\lambda$ was negative (Table 1) (normalized to percentages). Jackknife tests for training and testing datasets were done. Training gain values were derived from points used to train the model, and test gain values were derived from occurrence points used to test the predictive capacity of the model. In running jackknife tests, each variable is excluded in turn, and a model was created with the remaining variables. Then, a model is created using each variable in isolation. Using these jackknife tests, we identified for each monument type which input variable produced the largest gain value when it was excluded and which one produced the largest gain value when it was used as a single predictor (SI Text). A response curve using each input variable alone for the MaxEnt model for mounds.

societal development is a grand challenge for archaeology (1). We have presented MaxEnt as an analytic approach with much utility for improving our understanding of monument building across past landscapes. MaxEnt allowed us, working with an archival presence-only dataset, to establish the contribution of multiple spatioenvironmental variables to the distribution of mounds and enclosures in Michigan. MaxEnt proved particularly useful for looking at these two types of monuments together, allowing us to disentangle the ways that they were differentially positioned to help answer diverse social needs and environmental constraints. Our approach offers an especially powerful tool then for other cases where multiple types of monuments were integral in the spatial and material reconfiguration of past cultural landscapes. Our use of MaxEnt can also be expanded beyond the cultural process of monument building to other key cultural processes. Archaeologists can use MaxEnt to create maps, wherein the cultural process of interest becomes the species whose distribution is predicted across the landscape, and the variable contribution of input variables to the distribution of these cultural processes is likewise assessed.

Our research also shows MaxEnt modeling as a robust tool for working with archival archaeological site datasets. Because of long experienced intensive rates of destruction, the mounds and earthworks of the eastern United States make up a particularly compelling case where scholars must turn to early site record databases. During the 1800s and early 1900s, mounds and earthworks were recorded by scholars as well as citizens who recognized the significance of their rapid destruction. The records from these early efforts represent significant scholarly but also, public investment in archaeology. Although archival records have imperfections that make them difficult to work with, they contain irreproducible information about the spatial distribution of these monuments, and it is imperative to incorporate presence-only archival data into current research. Although the
and enclosures was created (Fig. 3 and SI Text). These curves show the dependence of predicted suitability for the selected variable (40).

ACKNOWLEDGMENTS. We thank the University of Michigan Museum of Anthropology for access to state site files; Emma Pankey for her time in the archives; and Joyce Marcus, two anonymous reviewers, and the PNAS Editor for providing invaluable feedback on the paper. Research was supported by NASA Space Archaeology NNH11ZDA001N SAP Grant NNX13AO40G [to M.W.P. (principal investigator)]. An ACLS Digital Innovation Fellowship provided time for synthesis and writing (M.C.L.H.).