Satellite-based remote sensing rapidly reveals extensive record of Holocene coastal settlement on Madagascar

Dylan S. Davis a,*, Vanillah Andriankaja b, Tahirisoa Lorine Carnat b, Zafy Maharesy Chrisostome b, Christophe Colombe b, Felicia Fenomanana b, Laurence Hubertine b, Ricky Justome b, François Lahiniriko b, Harson Leonce b, George Manahira b, Briand Venance Pierre b, Razafimagnéfa Roi b, Patricia Soafiavy b, Faralahy Victorien b, Vavisoa Voahirana b, Barthelemay Manjakahery c, Kristina Douglass a,b,d

a Department of Anthropology, The Pennsylvania State University, University Park, PA 16802, USA
b Morombe Archaeological Project, Andavadoaka, Madagascar
c Centre de Documentation et de Recherche sur l’Art et les Traditions Orales a Madagascar, Université de Toliara, Toliara, Madagascar
d Institutes of Energy and the Environment, The Pennsylvania State University, University Park, PA 16802, USA

A R T I C L E  I N F O

Keywords:
Settlement patterns
Remote sensing
Predictive modeling
GIS
Human behavioral ecology
Madagascar
Ideal free distribution

A B S T R A C T

Despite decades of archaeological research, roughly 75% of Madagascar’s land area remains archaeologically unexplored and the oldest sites on the island are difficult to locate, as they contain the ephemeral remains of mobile hunter/forager campsites. The known archaeological record is therefore biased toward later sites, especially sites dating to the second millennium AD, following the expansion of Indian Ocean trading networks. Systematic archaeological investigations are required to address these biases in the known archaeological record and clarify the island’s early human history, but funding limitations, logistical and time constraints in surveying large areas and a relatively small number of active field archaeologists present substantial barriers to expansive aerial survey coverage. Using theoretical models derived from human behavioral ecology (i.e., ideal free distribution, optimal foraging theory) in conjunction with freely available remote sensing data, we illustrate how archaeological survey of Madagascar’s landscapes can be rapidly expanded, more effectively target early archaeological deposits, and address questions about the island’s settlement. This study illustrates the potential for theoretically-driven satellite-based remote sensing analysis to improve our understanding of the archaeological record of the world’s fourth largest island.

1. Introduction

The human history of Madagascar, the world’s fourth largest island, is complex and involves the movement and dynamic interaction of people, plants, animals, and ideas from around the Indian Ocean (Radinilahy and Crossland, 2015; Dewar and Richard, 2012; Fuller et al., 2011). To-date, archaeological, genetic, and linguistic research have revealed the earliest known evidence of Madagascar’s far-reaching connections; the island lies at the westernmost reach of the Austronesian expansion (Growthber et al., 2016) and multiple lines of evidence testify to the migration of Bantu peoples from the African mainland to Madagascar (Parker Pearson et al., 2010; Pierron et al., 2017; Sussman et al., 1994). Important questions, however, regarding Madagascar’s human past remain poorly resolved. The timing and nature of Madagascar’s human colonization continue to generate intense debate in archaeology (Douglass et al., 2019a), and our understanding of subsequent social, economic, political, and ecological processes is limited, both temporally and spatially (Douglass and Zinke, 2015; Dewar and Wright, 1993).

Research into Madagascar’s early history requires new approaches to overcome existing barriers to our understanding. These include the poorly understood remains of ancient foraging and fishing communities, and the relationship between archaeological settlement patterns, environmental conditions, and climate change (e.g., Kull, 2005; Parker Pearson et al., 2010; Wright, 2007; Wright and Rakotoarisoa, 2003). Landscape-level approaches are critically needed to address these

* Corresponding author.
E-mail address: dds40@psu.edu (D.S. Davis).

https://doi.org/10.1016/j.jas.2020.105097
Received 22 October 2019; Received in revised form 23 December 2019; Accepted 27 January 2020
0305-4403/© 2020 Elsevier Ltd. All rights reserved.
research lacunae. To date, landscape-level approaches are mostly absent from archaeological studies on Madagascar (for exceptions see Dewar and Wright, 1993; Mille, 1976; Verin, 1986; Wright, 2007; Parker Pearson et al., 2010). This is partly because ground-based landscape investigations require large investments of time and resources in the field to generate sufficient information; funding, logistics and a small number of active field archaeologists have proven to be barriers to extensive areal coverage. Innovative approaches are critically needed to expand archaeological survey coverage and document cultural heritage, particularly considering that Madagascar is experiencing increasing impacts from climate change. Climate-driven threats on the island extend to both its people and their histories (IDMC, 2019; Lemahieu et al., 2018; USAID, 2016).

Here we present the first satellite-based remote sensing archaeological survey of the Velondriaka Marine Protected Area of southwest coastal Madagascar. Using freely-available satellite imagery, image processing algorithms, predictive modeling derived from human behavioral ecology (HBE) theory and ground-truthing survey, our approach successfully identifies cultural deposits throughout a ∼1400 km² area. The Velondriake (Fig. 1) case study demonstrates how the development of a predictive model to analyze satellite imagery can rapidly expand the known record of archaeological settlements on Madagascar, filling both temporal and spatial gaps at the landscape level. Systematically documenting ephemeral components of the archaeological record at the landscape scale is essential for answering longstanding questions in archaeology surrounding human-environment interactions, social complexity, resilience, and mobility (e.g., Kintigh et al., 2014). On Madagascar, specifically, little attention has been paid to internal migration within the island, but rather focused on arrival events and migration between Madagascar and surrounding regions (Allibert, 2008; Anderson et al., 2018; Beaujard, 2011; Dewar and Richard, 2012; Douglass et al., 2019b; Hansford et al., 2018; Mitchell, 2019; Verin et al., 1969). By conducting landscape scale surveys of the island, we will be able to address how communities moved throughout the landscape, and how such mobility was related to environmental, political, and social developments.

Our case study also highlights the importance of integrating theoretical models with remote sensing methods in African archaeology more broadly (Davis and Douglass, in press). Drawing on lessons from research conducted using HBE and related theoretical models from other regions (e.g., Baja California (Coddington and Jones, 2013), the Channel Islands (Winterhalder et al., 2010), Australia (O’Connell and Alien, 2008), and New York State (Adovasio, 1994)), we use archaeological deposits and historical record to inform our predictive models.
2012), Polynesia (DiNapoli and Morrison, 2017), we demonstrate that it is possible to use satellite-based remote sensing to test the nature of past human-environment interaction and drivers of settlement mobility. We further demonstrate that the integration of theoretical models and satellite-based remote sensing methods holds great potential for rapidly locating previously unrecorded archaeological deposits at vast geographical scales, even when these deposits are ephemeral in nature.

2. Previous landscape-level investigations on Madagascar

Most landscape-level archaeological investigations on Madagascar focus on periods from 900 B.P. to the present and highlight important demographic and political processes, including the overall increase in size and number of settlements and the rise of centers of political power (Parker Pearson et al., 2010; Wright, 2007). Although the reliability of chronometric determinations for early periods has been questioned (e.g., Anderson et al., 2018; Mitchell, 2019), evidence of the earliest occupations exists on Madagascar, extending the island’s human record as far back as ~10,000 B.P. during the Early Holocene (Hansford et al., 2018). Recent systematic assessment of Madagascar’s radiocarbon chronology supports the possibility of an Early Holocene human presence on Madagascar, despite a lack of contextual information on the nature of such an early presence (Douglass et al., 2019a). Given the evidence for Early Holocene human activity on Madagascar and the taphonomic and sampling challenges inherent in studying Madagascar’s ephemeral early forager sites (Douglass and Zinke, 2015), new approaches are urgently needed to record and assess over 90% of the span of time for which human presence has been recorded on the island. Landscape-level approaches, in particular, will be critical to understanding the evolution of settlement patterns and human-environment dynamics during early periods of human occupation. A diversity of landscape-level approaches has proven useful for understanding the interplay between human behaviors and environmental contexts in other parts of the world (e.g., Codding and Jones, 2013; DiNapoli and Morrison, 2017; Jazwa et al., 2017; Winterhalder et al., 2010).

Despite the critical temporal gaps in landscape-level archaeological investigations to-date, understandings of landscapes from the 10th century onwards (Crossland, 2001; Pearson, 1992; Sussman et al., 1994; Tucker, 2004; Wallace et al., 2016) have illuminated connections between humans and their environmental surroundings. Theory from human behavioral ecology (HBE) and historical ecology have been integrated in ethnographic work (e.g., Tucker, 2004; Tucker et al., 2010), but such approaches are still scarce in archaeological contexts (e.g., Douglass et al., 2019b). For example, Tucker (2004) demonstrates how Mikea foragers’ food-sharing practices are dictated by economic factors, reciprocity, kin selection, and tolerated theft. In another study by Tucker et al. (2010), HBE is used to understand risk mitigation via the practice of mixed subsistence strategies. Recent advances in remote sensing methods and datasets (e.g., Davis et al., 2019; LaRocque et al., 2019; Thabeng et al., 2019) offer important insights into environmental changes and human settlement patterns. It is possible to use satellite-based remote sensing to test the nature of past human-environment interaction and drivers of settlement mobility. We further demonstrate that the integration of theoretical models and satellite-based remote sensing methods holds great potential for rapidly locating previously unrecorded archaeological deposits at vast geographical scales, even when these deposits are ephemeral in nature.

3. Methods

Here, we outline a preliminary study that combines HBE modeling with remote sensing survey to predict the distribution of archaeological sites on southwest Malagasy landscapes. In this discussion, “site” refers to any area containing two or more artifacts during ground surveys. Sites thus encompass artifact clusters, settlements, and any other cultural materials present in an area. The approach is based on ideal free distribution (IFD) models (see Fretwell and Lucas, 1969). These models assume that individuals settle areas with the best overall suitability (with regards to available resources) and that, as population density and resource consumption increase, settlements shift to areas with lower resource suitability. Because the current study lacks absolute temporal control, the assumption is made that the earliest sites will be located in “high” suitability areas. Confirmation of this hypothesis requires further testing. Here we focus on the density and variability of cultural materials present in different suitability locations. Furthermore, we assess whether ethnographically and historically important resources (e.g., coral reefs, vegetatively productive land, distance to the coast, etc.) are good predictive variables for locating archaeological sites in southwest coastal Madagascar.

3.1. Ideal free distribution modeling

Within HBE, there are a series of different optimality models (optimal foraging theory, OFT) which try to predict decision making of individuals based on costs and benefits of different actions (e.g., Blurton Jones, 1986; Charnov 1976; Fretwell and Lucas, 1969; MacArthur and Pianka, 1966; O’Connell and Hawkes, 1981). Such modeling approaches have proven useful in exploring the rationale behind observed phenomena in anthropology, including archaeological evidence of behavior and choice (e.g., Bird et al., 2016; Codding and Bird, 2015; Jazwa et al., 2017; Robinson et al., 2019; Tucker et al., 2010). Despite criticisms of OFT (see Zeder, 2012), the explicit framework offered by such approaches provides a heuristic device for exploring factors that may influence settlement choice in human populations (e.g., Stiner and Kuhn, 2016).

IFD models, a type of OFT model developed by Fretwell and Lucas (1969), have been applied in various settings around the world for identifying temporal and ecological trends in population settlement distribution (see Winterhalder et al., 2010; Codding and Jones, 2013; Yaworsky and Coddin, 2018; Hanna and Giovas, 2019). IFD stems from the work of Fretwell and Lucas (1969) and operates on the principle of negative density dependence (Winterhalder et al., 2010; Yaworsky and Coddin, 2018). As population pressures increase, the overall resource quality of that area will degrade, thereby lowering the suitability of that habitat and its likelihood of being settled.

The IFD model, however, is simplistic, and there are biological principles that often violate its assumptions. For example, the Allee effect accounts for temporary improvements in habitat suitability caused by immigrating populations, community aggregation, and habitat modification (Fretwell and Lucas, 1969, 19). One example of Allee-effect IFD comes from Neolithic farmers who modified their landscapes to increase agricultural production by clearing forestland (McClure et al., 2009). IFD-Allee models predict that individuals settling lower ranking habitats attract others to follow, thereby abandoning higher suitability areas (Winterhalder et al., 2010, 473). As such, the highest suitability areas will have a slightly lower population than medium suitability locations.

There is also a variant of IFD for when access is restricted, and people establish certain controls over resources – ideal despotic distribution...
(IDD). IDD accounts for differences in competitive ability and resource control (Jazwa et al., 2017). In an IDD model, the opposite pattern of population distribution is expected from IFD, wherein the highest density of individuals will inhabit lower suitability habitats.

Since we currently have limited information about the resource management and land-use practices of these communities or changes in their demography at a fine resolution, we cannot definitively assess whether land-use practices led to degradation of environments as the IDD model posits. IFD models, therefore, are used as a theoretically-framed starting point, rather than IDD, so that we may begin to address this information in a theoretically sufficient manner (sensu Lewontin, 1974).

3.2. Remote sensing and predictive modeling

For this study we use freely downloadable satellite imagery from the European Space Agency Sentinel-2 satellite (https://www.copernicus.eu/en/access-data). This satellite has proven useful for a wide range of disciplines, including archaeology (Agapiou et al., 2014), but its medium-to-low resolution (10 m visual and near infrared (NIR), 20 m NIR and short-wave infrared (SWIR)) constrains its applicability, including for the documentation and preservation of cultural heritage. Because archaeological deposits on Madagascar’s southern coasts are often subtle artifact scatters, Sentinel-2 data do not have the spatial resolution necessary to directly identify these features. However, its resolution is conducive to developing a predictive model of site locations using the theoretical assumptions of IFD. While similar predictive measures have been used by other scholars (e.g., Agapiou et al., 2014; Bennett et al., 2012; Kirk et al., 2016; Lasaponara and Masini, 2007), most rely on interpreting vegetative indices for soil and vegetative anomalies, and do not always utilize explicit theoretical models from anthropology.

If our method is successful – and the data conform to an IFD – we expect: 1) that high value areas will contain the greatest proportion, density, and variety of artifacts; 2) these amounts will decrease steadily in Medium, Low, and Null probability areas; and 3) that the settlements located in high probability areas will be older than those in other locations. The third hypothesis is beyond the scope of the current paper and will be the focus of future research once sufficient temporal data become available.

3.3. Processing steps of predictive modeling analysis

1. Based on a review of available archaeological and ethnographic data, we developed a list of important resources and landscape features for communities of the southwest coast (e.g., Douglass, 2016; Douglass et al., 2018; Gommeroy et al., 2011; Pearson, 1992, 1997; Tucker, 2004; Tucker et al., 2010). These data include locations of coastal archaeological sites identified by surface survey and excavation (Douglass, 2016; MAP 2011–2017). Important variables that influence human settlement include: distance from the sea shore; distance from offshore coral reefs; distance from paleo-onudes; and the vegetative productivity of specific locations.

2. Training samples were created using the 2-D scatterplot function in ENVI to develop a total of 6 landscape classes (Fig. 2): water, coral, bare soil, shrubs, paleodunes, and dense vegetation (i.e., mangrove forests). This method was used for training sample collection to ensure a minimal amount of spectral overlap between each land class. An initial assessment of the spectral properties of the study area led to the decision to use the NIR, Red, and Green bands (RGB 843) in order to capture the most information pertaining to vegetative health and moisture properties for landscape classification.

3. Sentinel-2 images were classified using a support vector machine (SVM) classifier in ENVI 4.7 (Exelis Visual Information Solutions, 2009). SVM is a non-parametric classification technique that has gained popularity due to its ability to produce highly accurate classifications using limited training datasets (Mountrakis et al., 2011). The method works by identifying optimal separations between classes and can handle multiple classes simultaneously (Pal and Mather, 2005).

4. Coral reefs in some instances were not reliably classified using pixel-based methods (i.e., SVM). We therefore used an object-based image analysis (OBIA) approach with threshold classification (see Davis, 2019a; Sevara et al., 2016). Unlike pixel-based methods, object-based methods take shape, texture, and morphology into account to classify image components (Blaschke, 2010; Davis, 2019a; Hay and Castilla, 2008). This same procedure was used to classify the locations of offshore islands, which serve to extend fishing grounds and offer safe-havens for coastal fishers during periods of political instability (Cripps, 2009; Douglass, 2016:72). OBIA was used to generate shapefiles of offshore island and coastal locations using eCognition 9.0.1 (Trimble, 2014). Multiresolution segmentation was conducted using a scale parameter of 60, shape parameter of 0.7, and compactness factor of 0.6. These parameters were chosen following trial-and-error, wherein the chosen parameters resulted in the greatest accuracy. Following this step, pixel brightness thresholds were used to extract all image objects located in areas covered by or immediately adjacent to water (as identified by SVM) that matched threshold values for coral or offshore island features. Corals within this region contained brightness values between 600 and 1150 and offshore islands contained values of 1200 or greater. The OBIA results were then assessed manually to eliminate the few errors present throughout the study region.

5. Data generated from the SVM and OBIA classifications were imported into ArcGIS 10.6.1 (ESRI, 2018) and underwent several processing steps (Fig. 3). The water and paleodune classes were extracted into their own raster layers in ArcMap and subjected to Euclidean distance tests. Euclidean distance produces a raster of distance measurements between the input (i.e., water and paleodunes) and the surrounding pixels in an image. Euclidian distance is appropriate, as opposed to a cost-distance analysis, because of the gradual landscape elevation changes in this region. While hills and other topographic features are present, there are no extreme elevation changes within the study region.

6. The final variable incorporated is vegetative productivity. To
measure vegetative productivity a SAVI (soil adjusted vegetative index) was used, which takes into consideration soil properties, including moisture content (Huete, 1988). Given the extreme variance in soil reflectance characteristics on Madagascar (see Clark et al., 1998), SAVI was chosen as the most appropriate index, as opposed to NDVI (Normalized Difference Vegetation Index) and others (e.g., simple ratio, leaf area index, etc.; see Jensen, 2007:384–385) that decrease in accuracy over large geographic areas with high vegetative diversity (Jensen, 2007). SAVI is calculated using the formula:

\[
SAVI = \frac{NIR - red}{NIR + red + L}
\]

where the NIR and red bands are used, and L represents the soil adjustment factor. The best soil adjustment has been demonstrated around L = 0.5 (Huete, 1988; also see Jensen, 2007) and was chosen for this study. Once calculated, SAVI indices that contained values associated with the presence of shrubs and other vegetation were extracted and we conducted another Euclidean distance function to produce a distance raster of vegetative areas.

7. With all these variables together, we used the following formula to calculate overall probability of early forager settlements in ArcGIS using the raster calculator:

\[
P_{arc} = \frac{1}{\sum d_{i}}\left(\frac{1}{d_{i}}\right)^{100}
\]

Where \(P_{arc}\) is the probability of archaeological deposits, \(d_{i}\) is distance from water, \(d_{c}\) is the distance from coral beds, \(d_{b}\) is the distance from paleodunes, \(d_{l}\) is the distance from land with SAVI index values of 0.35 or better (this value represents the minimum value for shrubland), and \(d_{o}\) represents the distance to offshore islands. Each distance raster was inversed to produce the highest values for the lowest distance from each resource type. Once the index was calculated, we used inverse-distance weighting (IDW) interpolation to fill in gaps in the probability raster up to 100 m using the elevation void fill function in ArcMap 10.6.1 (ESRI, 2018).

8. Following the development of our predictive model, we assessed the model’s ability to detect prerecorded archaeological sites (MAP 2011–2017) and established a sampling strategy for field tests to assess the model’s ability to predict the location of previously unrecorded cultural deposits. To accomplish this, we first compared the locations of prerecorded sites to the probability values generated by the algorithm. Then, to assess the algorithm’s ability to detect previously unrecorded materials, we created a grid of 50 m × 50 m squares throughout the entire study area (~1400 km²). Each grid was assigned a unique identification number by ArcMap. In Excel, we randomly selected 600 ID numbers using the “randbetween” function. These 600 areas were then checked to ensure they were accessible on foot. Ultimately, a total of 145 areas were selected on the basis of proximity to other points, accessibility, and feasibility of visitation during the 2019 field season. Among the randomly selected grids, 73 contained “high” probability zones, 31 contained “medium probability”, and 27 had “low” probability. Surveyors did not have any prior knowledge of the probability of locating sites to ensure an unbiased recording of materials. Table 1 shows the quantitative breakdown of these qualitative categories. The remaining

<table>
<thead>
<tr>
<th>Quantitative Values</th>
<th>Qualitative Ranking Equivalent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Null/Blank</td>
<td>Null</td>
</tr>
<tr>
<td>0–5.5</td>
<td>Low</td>
</tr>
<tr>
<td>5.5–11.4</td>
<td>Medium</td>
</tr>
<tr>
<td>&gt; 11.5</td>
<td>High</td>
</tr>
</tbody>
</table>
14 areas had null probability values and were chosen to assess false negative results.

Together, these methods produced information needed to calculate overall habitat suitability, and by extension, probability of settlement for coastal communities. Based upon the expectations of IFD, the highest suitability (and hence probability) locations will hold the greatest number of archaeological deposits, with lower suitability areas containing fewer archaeological assemblages.

4. Results

SVM resulted in 93.6% accuracy (KIA = 0.931) (see Tables 2 and 3) and OBIA attained an overall accuracy of 97.7% (KIA = 0.914) for the classification of the chosen environmental land-types (Tables 4 and 5).

4.1. Prediction of pre-recorded site locations

Within the entire dataset of prerecorded archaeological deposits (n = 756), we find that only five previously surveyed deposits do not fall within areas identified by the algorithm (Fig. 4). All of these deposits are located on paleodune features, however, suggesting a strong relationship between this environmental context and human settlement. The model thus reliably predicts the location of pre-recorded deposits.

4.2. Prediction of previously unrecorded site locations

To assess the ability of the model to locate previously unrecorded cultural materials, ground surveys were carried out on 71 of the 145 selected sites during the summer of 2019 (Supplemental File 1). A variety of different materials were recovered during surveys, ranging from ceramics and beads to elephant bird eggshell and marine shells (Table 6). When assessing these artifacts, we distinguished between “definitive” and “possible” human presence, with “definitive” referring to materials that were clearly made or altered by people (e.g., ceramics, modified shell, etc.) and “possible” referring to materials that are in direct association with other cultural artifacts or contexts, such as burning activity. The results largely fit the hypothesis that high suitability areas will contain the greatest proportion, density, and variety of artifacts and that these amounts will decrease steadily in Medium, Low, and Null probability areas. However, there is a slight increase in the density of artifacts within medium probability zones, suggesting a possible fit with an IFD with Allee effect model (see Fretwell and Lucas, 1969, Fig. 5).

Grid probability values are an average of raster pixel data. As such, even “high” probability grids may contain values that are lower in likelihood. Therefore, we investigated the precise locations where materials are present to see if individual materials are also accurately predicted by the model. When examining the locations of individual artifacts, we once again find a strong clustering in high probability areas (Table 7; Fig. 6).

5. Discussion

The distribution of materials recovered during pedestrian survey

<table>
<thead>
<tr>
<th>Table 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confusion Matrix for SVM classification. Numbers are total ground-testing points.</td>
</tr>
<tr>
<td>Class</td>
</tr>
<tr>
<td>Mangroves</td>
</tr>
<tr>
<td>Dunes</td>
</tr>
<tr>
<td>Coral</td>
</tr>
<tr>
<td>Water</td>
</tr>
<tr>
<td>Forest</td>
</tr>
<tr>
<td>Shrubs</td>
</tr>
<tr>
<td>Sand/Bare Ground</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

suggests that the immediate coastline is the most densely inhabited area of the study region (Fig. 1), with material culture abundance (i.e., ceramics, beads, modified shells, etc.) steadily decreasing as one moves inland (Fig. 7). As indicated in Table 6, the diversity of artifacts also decreases in lower suitability areas, suggesting possible limitations for resource acquisition. For example, in high suitability areas, metal and coral artifacts, in addition to beads and other artifact types, are present. However, in medium, low, and null suitability areas metal and coral were not recovered, and the amount of beads and other artifacts decreases drastically. This raises the possibility that resource access may have been controlled, as the variety of materials is not even across the study region.

This preliminary landscape analysis illuminates several possibilities to understand settlement patterns. Archaeological deposits identified here all fall within the expectations of optimal foraging and IFD modeling frameworks. Populations settling the coast of southwest Madagascar appear to have prioritized shoreline ecosystems with ready access to resources that are still valued today. Cultural deposits found further inland are often within flood-zones during the wet season. This pattern coupled with the dearth of cultural materials exceeding 5 km from the modern coastline suggest that settlements are greatly influenced by marine resources.

Furthermore, Allee’s principle (Allee and Bowen, 1932) argues that community formation can produce increased fitness for a population’s survival. Social ties can act as Allee effects because they can foster cooperation between individuals, thereby allowing for resource
Fig. 4. Results of Archaeological Probability Index and select pre-recorded sites. A. Sites represent those recorded by recent survey work between 2017 and 2018 and earlier surveys recorded by the MAP between 2011 and 2018 (see Douglass, 2016). B. Shows closeup of a cluster of sites found in map A. C. Shows another cluster of sites in map A. Both clusters (B and C) fall primarily on high and medium values.

Table 6
Sum of survey results by grid probability value. Note, the amount of material increases between each level, with the greatest increase for high probability locations. Lower probability areas still produce artifacts, which is expected, but at lower densities and amounts, as predicted by an ideal free distribution model. The variety of artifacts also decreases with probability.

<table>
<thead>
<tr>
<th>Grid Probability</th>
<th>Types of artifacts</th>
<th>Charcoal</th>
<th>Eggshell</th>
<th>Marine Shell</th>
<th>Faunal Shell</th>
<th>Ceramics</th>
<th>Beads</th>
<th>Lithics</th>
<th>Botanicals</th>
<th>Glass</th>
<th>Metal</th>
<th>Burnt Stones</th>
<th>Coral</th>
<th>Total Materials Collected</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td></td>
<td>9</td>
<td>6</td>
<td>186</td>
<td>204</td>
<td>35</td>
<td>193</td>
<td>26</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>Medium</td>
<td></td>
<td>8</td>
<td>1</td>
<td>88</td>
<td>102</td>
<td>6</td>
<td>129</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Low</td>
<td></td>
<td>5</td>
<td>1</td>
<td>54</td>
<td>41</td>
<td>5</td>
<td>39</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Null</td>
<td></td>
<td>3</td>
<td>0</td>
<td>22</td>
<td>13</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
acquisition to be shared and limiting the burden on smaller groups or individuals. An Allee effect distribution would suggest that coastal foraging populations: a) actively changed and improved the suitability of the areas they inhabited (which has been documented elsewhere: see Freeman and Baggio, 2017; McClure et al., 2009; Quintus and Cochrane, 2018; Quiros et al., 2017); and/or b) that social networks were strong unifying factors that led to significant population movements as environmental resources shifted.

Additional support for an idealized distribution comes from Kolmogorov-Smirnov (K–S) distribution tests which reveal distinct differences between the archaeological data and other continuous distribution functions (Table 8; Fig. 8; Supplemental File 2). The data are not normally distributed, nor do they conform to Gamma or Poisson patterns. Gamma distributions are often used to evaluate skewed datasets with positive values (Hogg et al., 2005) and Poisson distributions measure spacing between randomly occurring events (Haight, 1967). The results indicate that archaeological distributions are statistically different from these patterns but are closest to a uniform dispersal.

If comparing a uniform dispersal to an IFD, we can expect that in high suitability areas IFD will have greater densities than uniform distributions, but moderate distributions should be about equal (i.e., densities are distributed rather evenly across moderate suitability spaces). When looking at the comparison between the archaeological and simulated uniform distribution data, the archaeological data mostly matches this expectation. The conformity of the archaeological data to an IFD remains a hypothesis, however, as temporal information is needed. Nonetheless, the results of the K-S tests provide evidence that justifies further research into this question.

Consistent with findings by Douglass (2016), this study also suggests a relationship in the Velondriake area between possible elephant bird (Aepyornithidae) nesting grounds and human settlement locations. Elsewhere on Madagascar associations between cultural contexts and elephant bird eggshell have also been noted (e.g. Parker Pearson et al., 2010; Radimilahy, 2011; Battistini and Verin, 1972). Aepyornis eggshell remains are often located in ancient paleodunes which are present along the coasts of southern Madagascar and are easily visible from medium-to-course resolution satellite imagery (Clark et al., 1998). The identification and survey of paleodune features is likely to yield exciting new information regarding the interaction between humans and these large avifauna, and the investigation of paleodunes should be prioritized to better understand the processes that contributed to the birds’ decline.

### 5.1. Future work

While the results are highly positive, there is room to improve the predictive power of the algorithm developed here. Future work will look to improve the method by incorporating additional ethnographic and environmental variables that were potentially overlooked, such as groundwater levels. Looking at the results, the greatest density clusters of materials seem to occur in areas closest to offshore islands and on coastlines that contain coves sheltered by rocky coastal barriers. Conducting spatial-statistical tests can reveal the most significant variables for predicting archaeological material and will be the focus of future work.

Furthermore, the results of the surveys carried out under the direction of this remote sensing model will be used to address larger questions concerning human-environmental interaction through time. In particular, future work will integrate settlement pattern data with high resolution paleoenvironmental and paleoclimate records, to enable modeling of human response to climate and environmental change. This will enable researchers to understand settlement and migratory patterns and their connection to environmental conditions. As fieldwork continues, temporal data will become available for many of these newly identified deposits. To date, we know that several previously excavated sites dating to ~2500 B.P (see Douglass, 2016), were re-identified as “high” likelihood by our algorithm. This suggests that other contemporaneous –

---

**Fig. 5.** Density distribution of artifacts recovered from different probability locations. Black lines represent the IFD-Allee curves, with the top line representing population thresholds for the best habitats, and the second line showing thresholds for less suitable habitats. Temporal data is still needed to confirm conformity to an ideal distribution.

**Table 7**

Descriptive statistics of probability values for specific points within survey grids where materials were recovered. The average and most frequently occurring values are high probability, thus coinciding with the grid-level data. Note, these calculations ignore null values.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Quantitative Value</th>
<th>Qualitative Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum</td>
<td>1.69</td>
<td>Low</td>
</tr>
<tr>
<td>Maximum</td>
<td>22.96</td>
<td>High</td>
</tr>
<tr>
<td>Mode</td>
<td>16.33</td>
<td>High</td>
</tr>
<tr>
<td>Mean</td>
<td>12.71</td>
<td>High</td>
</tr>
</tbody>
</table>
and possibly earlier sites – will emerge as our ground surveys continue.

Based on preliminary analysis of ceramic decorative attributes (see Douglass, 2016) recovered during ground surveys, many of the sites identified using this predictive model date to at least 800–1000 B.P. High probability sites had the greatest range in ceramic styles, signifying longer occupational durations, while low probability sites had fewer ceramics and less variation. There were also many sites with undecorated ceramics, and an absence of ceramics, which tend to signify earlier occupations than those with decorated ceramics (Douglass, 2016). Furthermore, these surveys contain only surface deposits, meaning that these dates likely represent the latest materials on sites that were occupied during earlier periods. Seeing as present sea levels were similar to those 3000–6000 B.P., with a 2–3 m rise between 1000 and 3000 B.P. (Virah-Sawmy et al., 2009), identified sites that are near the modern coastline and lack ceramics may yield information about human settlement as early as 6000–7000 B.P. As radiocarbon dates become available, this hypothesis will be assessed.

6. Conclusions

The method developed here is already revealing important information pertaining to settlement patterns in Southwest Madagascar. We now have evidence that coastal foragers in the past placed importance on similar environmental resources as contemporary communities. We also have additional evidence of human-megafauna interactions, which will prove useful for understanding extinction patterns and the role of terrestrial resources in coastal community lifeways. Furthermore, we have a systematic dataset that can be used to test hypotheses regarding internal mobility and migration.

Our case study illustrates the utility of HBE theory for framing predictive remote sensing analysis. The protocol described here evaluated the probability of cultural activity at an average rate of ~50 km² per hour of processing time.¹ With greater processing power, this rate can be increased further, saving time, money, and resources by targeting high probability areas for ground survey. Additionally, all the analyses conducted here use freely available satellite imagery and can be analyzed using open-source software, including QGIS (QGIS Development Team, 2018) and R (R Core Team, 2018). With greater access to geo-spatial and statistical training, this work can be greatly expanded by other researchers, particularly in regions that are understudied in archaeology.

The acquisition of remote sensing datasets at higher spatial and spectral resolutions will allow researchers to directly identify archaeological deposits on Madagascar, rather than assign general probabilities for where these features are located (e.g. Calleja et al., 2018; Davis et al., 2019; De Laet et al., 2007; Guyot et al., 2018; Lasaponara and Masini, 2007; LaRocque et al., 2019; Traviglia and Cottica, 2011; Trier et al., 2019).

¹ This ratio was calculated on the basis of the average time allocation for each section of the study area. The region was divided into 3 parts totaling ~1400 km², with each section requiring approximately 6–8 h of computer processing time for the SVM classification and another 2 h of manual processing time to create the final probability map. Total, this procedure can be achieved with high levels of time- and cost-efficiency which can be cut down even further depending upon computing power and processing speeds. Computer used for analysis had an Intel® Core™ i7-4790 CPU @ 3.60 GHz Processor with 32.0 GB of RAM.
Because remote sensing surveys can often only identify locations of the largest-scale features – and thereby bias our understanding towards specific activities, the use of theoretical models can help to direct ground survey efforts in conjunction with remote sensing data to reduce some of these biases by identifying a greater variety of cultural activities. The method developed here makes it possible to identify early deposits on Madagascar which are currently at risk of disappearing due to erosion and sea-level change. We must act quickly to uncover the fragile remains of the earliest settlers of Madagascar, as these components represent an actively disappearing cultural landscape. Threats to cultural heritage from environmental factors such as erosion and sea level rise, are exacerbated by urban development and other anthropogenic factors (Douglass, 2016; Parker Pearson et al., 2010; Wright, 2007; Wright and Rakotoarisona, 2003).

Uncovering and preserving these data requires an expansion of remote sensing surveys – via satellites, drones, and other instruments – to rapidly and systematically survey vast geographic space. There have been calls in recent years to expand systematic survey of Madagascar’s landscape (Parker Pearson et al., 2010; Douglass and Zinke, 2015), including the often-neglected areas inland from the immediate coastline (Douglass et al., 2018). While our study looks at coastal areas in the

Table 8
Results of K-S tests between archaeological probability distribution and randomly generated probability distributions. All tests were run in R (R Core Team, 2018) using the stats package (see Supplemental File 2).

<table>
<thead>
<tr>
<th>Compared Distribution</th>
<th>D-value</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>0.83673</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Gamma</td>
<td>0.80823</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Poisson</td>
<td>0.81229</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Uniform</td>
<td>0.13469</td>
<td>0.02438</td>
</tr>
</tbody>
</table>

Fig. 7. Shows the locations of artifacts (and clusters of artifacts) recovered from grids visited throughout the study area during July and August of 2019. Definitive human presence is signified by beads, ceramics, and burnt/worked marine shells. Possible human presence is signified by the presence of shells and faunal remains that are burned, but not worked or modified, and an absence of ceramics or beads.
Southwest, the method can easily be expanded to inland regions of Madagascar.

Funding

This research was supported by the Institute for Computational and Data Sciences (formerly the Institute for CyberScience) at Penn State, a Hill Fellowship Award through Penn State, and the National Aeronautics and Space Administration under Grant No. NNX15AK06H issued through the PA Space Grant Consortium.

Data availability

All data necessary for replicating the analyses are available through Penn State’s ScholarSphere repository at https://doi.org/10.26207/1a47-pw11.

Declaration of competing interest

None.

Acknowledgements

A version of this paper was presented at the 14th Madagascar Workshop at the University of Toronto, October 25–26, 2019. The authors wish to thank the entire Morombe Archaeology Project team, in particular, Tsivery, and Hira for assisting with transportation and survey. We are also indebted to Mayor Adolphe Ediery of the Commune of Befandefala and the following villages and their presidents for granting us permission to conduct fieldwork on their lands: Presidents Eugene Radafinely (Andavadoaka), Vincent Andre (Akitambagna), Joseph Abesola (Befandefala), Meridary Elson (Belavenoike), and Agto Davison. We also wish to thank the three anonymous reviewers whose feedback greatly improved the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jas.2020.105097.

References


