Open-Access Archaeological Predictive Modeling Using Zonal Statistics: A Case Study from Zanzibar, Tanzania

WOLFGANG ALDERS

ABSTRACT

This paper presents a case study using zonal statistical analysis for archaeological predictive modeling with open-access software and free geospatial datasets. The method is applied to the rural landscapes of Zanzibar, Tanzania on the Swahili Coast. This study used QGIS (version 3.28) to perform zonal statistical analyses of environmental datasets weighted by settlement classes digitized from a 1907 historical map, to create predictive models for settlement across the island. These models were compared against the locations of major precolonial archaeological sites on the island and site data from a random stratified archaeological survey in an environmentally diverse region of northern Zanzibar. The results show strong correspondences between larger permanent site locations and areas of high likelihood for site detection in the predictive model. Additionally, there were correspondences between areas of lower likelihood for site detection and smaller, ephemeral sites related to agricultural production in swidden field plots. These results attest to Swahili rural complexity and enable an understanding of the specific environmental affordances that structured settlement and land use over the last millennium, in ways that shaped colonial contact in rural areas and altered the sociopolitical development of Zanzibar and the East African coast. The methods described here may be applicable for researchers and heritage managers in Africa and the Global South, where funding for large-scale field projects, expensive satellite imagery, or software licensing is limited.

CORRESPONDING AUTHOR:
Wolfgang Alders
University of Arkansas, US
walders@uark.edu

1. INTRODUCTION

The African continent is a key site for the expansion of geospatial archaeological methods. Geospatial and remote sensing approaches to archaeological research in Africa have increased in visibility and importance in the last decade, and this has coincided with calls to increase the accessibility of these methods for researchers with limited funding (Davis and Douglass 2020; Klehm and Gokee 2020). One avenue for increased geospatial accessibility has been the development free, low-cost, and open-access tools for archaeological remote sensing and geospatial research (Casana 2020; Cerasoni et al. 2022; Davis and Sanger 2021; Fisher et al. 2021; Khalaf and Insoll 2019; Sadr 2016; Rayne et al. 2020). This paper contributes by developing an archaeological predictive model for Zanzibar, Tanzania, using free geospatial datasets and open-access software. The model was created by digitizing a historical map and performing zonal statistical analyses of these features across weighted environmental raster images in QGIS 3.28. Summing these weighted zonal raster images produced two predictive models showing zones of probability for future site detection. These models were ground-truthed with archaeological field survey data from an inland region of Zanzibar, Tanzania, and were also compared against the locations of known major precolonial sites.

Zanzibar is an island region in Tanzania that was centrally important to the sociopolitical development of the Swahili Coast and the western Indian Ocean social system over the last two millennia (Crowther et al. 2016; Fitton 2018). The island of Zanzibar is environmentally diverse, with two major ecological zones: an agriculturally fertile northwestern region with deep soils and above-ground streams, and a rocky, agriculturally marginal karstic limestone landscape in the south and east, where water does not persist above ground (Alders 2023). Archaeologists have increasingly investigated ecological relationships between Swahili people and their landscapes on Zanzibar (Faulkner et al. 2022; Fitton et al. 2023; Kotarba-Morley et al. 2022; Prendergast et al. 2017; Quintana Morales et al. 2022). This study builds on this recent research, contributing to an understanding of the environmental affordances that structured long-term settlement and social change. Modeling human-environment relationships may enable future archaeological prospection on the island where settlement patterns are poorly understood, especially with regard to sites that do not possess standing stone architecture (Alders 2023; Fitton 2018; Horton and Clark 1985). In doing so, this research contributes to a long-standing orientation toward uncovering "hidden majorities" (Fleisher and LaViolette 1999) of Swahili non-elites, who created complex and independent rural societies beyond the boundaries of monumental stone-built towns (Kusimba et al. 2013; LaViolette and Fleisher 2018; LaViolette et al. 2023). Given its dynamic precolonial and colonial history and diverse environmental conditions, Zanzibar is a well-suited context for investigating the relationships between processes of urbanism, colonialism, and environmental factors, and as a case study for testing the suitability of an open-access method for predictive modeling.

1.1 PREDICTIVE MODELING, ZONAL STATISTICS, AND ARCHAEOLOGICAL PROSPECTION IN AFRICA AND BEYOND

Archaeological predictive modeling has continued to develop in relevance and sophistication since its inception, incorporating post-processual critiques related to environmental determinism, agency, and the interplay between data-driven and theory-laden approaches (Castiello 2022; Magnini and Bettineschi 2021; Verhagen and Whitley 2020). An assumption shared by all archaeological predictive models is that archaeological features were not randomly produced by humans in the past, but that a confluence of social and environmental factors conditioned their spatial location. Modeling the relationship between known archaeological features and their material and spatial environments can give insights into the locations of presently unknown features, aiding in archaeological prospection and survey and informing an understanding of human-environment relationships. Recent studies have evaluated the predictive power of different statistical approaches and sampling strategies (e.g., Castiello and Tonini 2021; Comer et al. 2023; Kelly et al. 2023; Yaworsky et al. 2020) and developed techniques for raster imagery analysis using machine learning and object-based imagery analysis (e.g., Magnini and Bettineschi 2021).

In African archaeology, recent studies have used a combination of remote sensing and spatial analysis for site detection, predictive modeling, and for understanding archaeological landscapes (Biagetti 2017; Davis and Douglass 2021; Fitton et al. 2023 Harrower et al. 2020; Klehm et al. 2019; Ochungo et al. 2022; Pawlowicz et al 2020; Reid 2016, 2020; Thabeng et al. 2020). Creating models for site detection with multispectral imagery is increasingly accessible, with resources like Landsat, Planet, and Sentinel 2 imagery becoming available at increasingly high spatial, spectral, and temporal resolutions. While recognizing the utility of multispectral remote sensing for site detection, one limitation in rainy tropical regions like Zanzibar is the low availability of consistently cloud-free imagery, and the academic licensing required to access multispectral imagery.

To work around this problem, this paper presents a case study for an alternative method: predictive modeling through zonal statistical analyses of environmental raster images in comparison to training data, which are used to the weight a final summative model (see Fitton et al. 2023 for another regional study integrating multiple geospatial and legacy datasets on Zanzibar). This

method is based around quantifying the most suitable spatial zones for the occurrence of specific phenomena by summing weighted raster datasets using raster calculations (e.g., Behr et al. 2017; Kuria et al. 2011). It is a predictive model that relies on environmental zonal raster images (for instance, published or archived maps) and factors for weighting these raster images; in this case, the primary factor is known site location and site sizes for a region of northern Zanzibar during the late colonial period, digitized from a historical map.

Zonal statistical analysis of training features with the Majority statistic is the method chosen to weight environmental raster images, because it is a simple, accessible, and powerful tool that is built into the functionality of open-access geospatial software like QGIS. In keeping with the theme of accessibility, zonal statistical analysis can be carried out by researchers with limited resources, lack of experience in computer programming languages, and limited experience with complex statistical modeling. An innovation of this paper is the use of the coefficient of variation statistic to further weight favored zones.

Sources used for zonal statistical analysis in this paper include a digital elevation model derived from free SRTM satellite imagery (D'Andrea 2008; Harrower 2010; Harrower et al. 2012; Hritz 2010), spatial-environmental datasets from published sources (Colbert et al. 1987; Hardy et al. 2015; Khamis et al. 2017), and a historical map of Zanzibar that was published in 1907. The historical map was georeferenced using methods developed for 19th-century Survey of India maps (Garcia et al. 2019; Green et al. 2019; Petrie et al. 2019). It contains a wealth of information about pre-modern rural Zanzibar, showcasing landscapes and features that have disappeared due to urban and agricultural development. This is the first time this map has been considered in detail, and in addition to the conclusions of this paper, it is hoped that the data will be a valuable resource for archaeologists and heritage managers in the future who are concerned with Zanzibar's colonial history. Heritage management and conservation on Zanzibar is constrained by limited resources and many archaeological sites may be in danger of destruction as urban growth and agricultural development continues (Mansab 2021, interview with Mariam Mansab, director of Zanzibar's Department of Museums and Antiquities). The digitization of this map and the creation of archaeological predictive models may aid in site conservation and stewardship in areas where development is proceeding.

Workflows for historical map digitization, zonal statistical analyses, and raster calculation are modeled in QGIS 3.28, a free and open-source GIS. The archaeological predictive model was ground-truthed using a random stratified survey across an environmentally and socio-politically diverse zone in rural northern Zanzibar (Alders 2023), and also compared to major known precolonial sites on the island (Horton and Clark 1985; Fitton 2018).

This paper advances the development of open-source geospatial applications for archaeological prospection, especially in the tropical, forested environments of sub-Saharan Africa. In line with other recent examples, it draws on the availability of free geospatial datasets to help understand archaeological landscapes. The sections below outline the methods and results of this case study in Zanzibar, Tanzania. Results inform a discussion of the environmental affordances which structured Swahili social development over the last millennium.

2. METHODS

Table 1 below presents the workflow for the methods used in this paper, outlining basic steps for producing an

1	**Prepare the Map and Raster Datasets:** Assemble and normalize raster datasets, define Area of Interest (AOI), and set map coordinate reference system.
2	**Create Zonal Raster Images:** Create raster images with unique values for each zone, which can be uniformly queried in relation to training features.
3	**Prepare the Training Features:** Import or digitize training set datapoints, and buffer them as polygons to reflect their real-world area.
4	**Calculate Zonal Statistics:** Quantify the distribution of training features for each zone across each raster image.
5	**Determine Zone Favorability:** Find zones with the highest density of training points for each zonal raster. This will determine which zone will be weighted as part of the predictive model for each zonal raster image.
6	**Determine the (Un)Evenness of Training Feature Distribution:** Find the coefficient of variation (CV) for training features across all zones. This value will be used to weight zones.
7	**Create and Assign Weight Classes:** Determine coefficient of variation (CV) threshold values and assign these thresholds to weight classes, which will be used to weight and add zonal raster images.
8	**Create Weighted Zonal Raster Images:** Create zonal raster images with a value for the most favored zone (weighted by CV class), and a value of 0 for all other zones.
9	**Sum Weighted Zonal Raster images into a Predictive Model:** Create a raster image that reflects site detection probability for each pixel, and reclassify it into a zonal raster image with categories for site detection probability.

Table 1 Workflow of methods for this study.

archaeological predictive model using zonal statistical analysis. A more detailed version of this table with step-by-step processing instructions for QGIS 3.28 is located in Supplementary Materials. Sections 2.2 through 2.5 describe the application of these steps for the study area.

2.1 STEPS 1–2: PREPARING THE MAP AND RASTER DATASETS AND CREATING ZONAL RASTER IMAGES

Table 2 lists the geospatial datasets used for this study. Converting these datasets into zonal raster images required digitizing geospatial datasets through a two-step process. The first step was to normalize the colors within all zones in an image editing software and to remove text, arrows, and boundary lines which would give different values when attempting to classify zones. This process decreases accuracy at boundaries to some extent but is necessary for producing raster images with a single pixel value per zone. Next, the maps were georeferenced and clipped to an outline of Zanzibar in QGIS 3.28. Then, the Reclassify by Table tool in QGIS 3.28 was used to turn the digitized maps into zonal raster images with one unique pixel value per zone. This study produced eight zonal raster images from previously published geospatial data and free SRTM imagery acquired from USGS. Figure 1 shows an example of a zonal raster image: soil type zones on the island, derived from Khamis et al. (2017: 120). Maps of the rest of the zonal raster images can be found in the Supplementary Materials section, as Figures A through H.

Care must be taken to avoid choosing raster images with zones that co-vary significantly because of dependencies between them. Dependent environmental factors summed together as raster images would over-weight certain zones. For instance, two other zonal raster images were considered for use in this model: a map depicting areas inside and outside of the historical clove plantation zone (Sheriff et al. 2016: 20), and a map of soil infiltration zones (Hardy et al. 2015). These images were

discarded because they co-vary with elevation and soil types in a dependent way, meaning that their inclusion would bias the model more heavily toward specific zones. Geology and soil type raster images co-vary to some extent, but they were both included because they describe different independent categories: the former describes geological categories, while the latter map of soil types refers to indigenous Swahili topsoil categories that relate to organic composition, soil color, and soil depth.

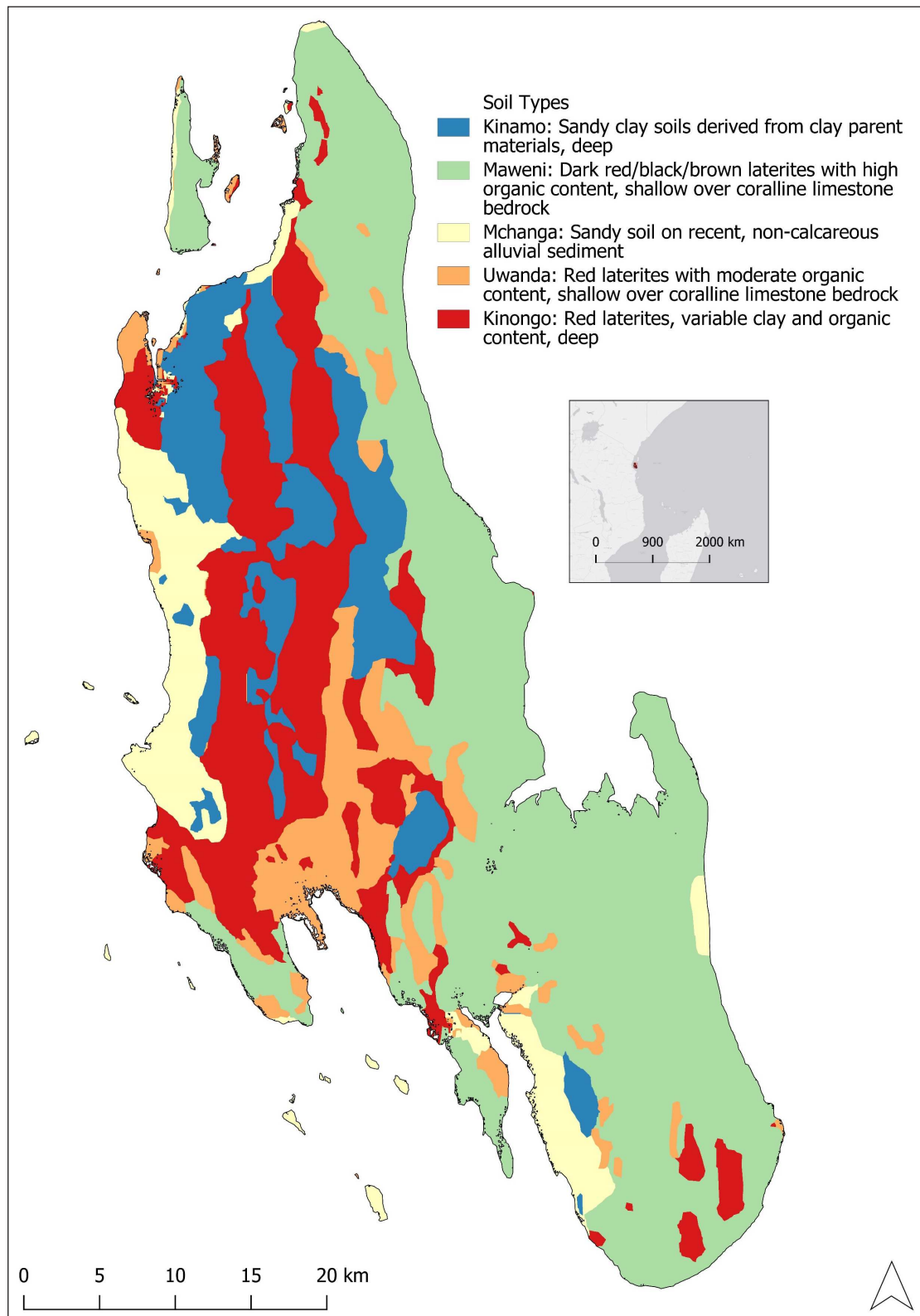## 2.2 STEP 3: PREPARING THE TRAINING FEATURES

Training features for the predictive model were derived from settlement classes digitized from a historical map of Zanzibar, since these settlement locations likely reflect environmental affordances that may have conditioned the spatial patterns of archaeological sites over the last millennium in Zanzibar. Stanford's Geographical Establishment in London published a map of Zanzibar, showing villages, landforms, and other features recorded on the island during the 1890s (A Map of Zanzibar Island, 1907). Figure 2 shows this map, referred to from here on as the 1907 Zanzibar map. The inset is shown in higher resolution in Figure I in the Supplementary Materials section.

The map is not an official British Survey of India map, but the legend names the mapmaker as Imam Sherif Khan Bahadur, a surveyor of the British Survey of India. The survey that produced the map likely occurred between 1892 and 1894, when Imam Sherif Khan Bahadur was stationed in Zanzibar (National Archives of India, 1894). However, this date is complicated by some details on the map. Marahubi Palace (built by Sultan Barghash in 1880) is listed on the map as a ruin, suggesting that Imam Sherif Khan Bahadur or someone else surveyed that region after Marahubi Palace was destroyed by a fire in 1899 (Rhodes et al. 2015: 350). Some other aspects of the map also suggest different

NAME	DESCRIPTION	SOURCE
Aspect	Eight aspect (orientation) zones across the island	Derived from a DEM from free 30 m SRTM imagery, from USGS
Elevation	Elevation zones across the island	Derived from a DEM from free 30 m SRTM imagery, from USGS
Geology	Geological zones	Colbert et al. 1987; Hardy et al. 2015
Rainfall	Rainfall zones across the island	Colbert et al. 1987
Reef Buffers	Buffer zones showing distance intervals from offshore coral reefs	Khamis et al. 2017: 120
Soil Types	Local soil types classified by Swahili categories	Khamis et al. 2017: 120
Slope Degree	Three slope degree zones across the island	Derived from a DEM from free 30 m SRTM imagery, from USGS
Stream Buffers	Buffer zones showing distances intervals from streams depicted on the 1907 map	Digitized from the 1907 map of Zanzibar

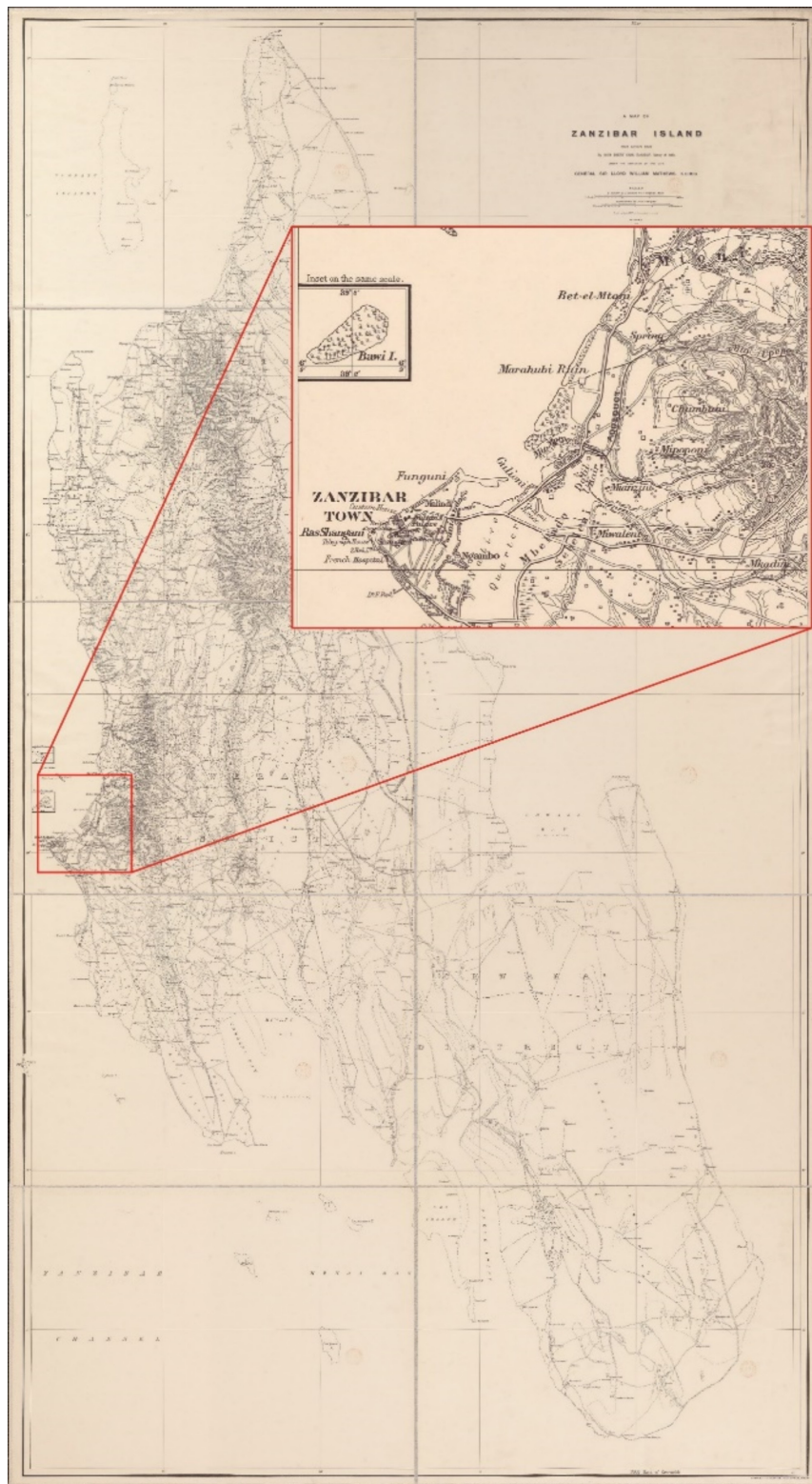**Table 2** List of geospatial datasets used in this study.

Figure 1 Example of a zonal raster image, a soil type map derived from Khamis et al. (2017: 120).

dates. Frazer's Sugar Mill and Frazer's House are not listed as ruins, even though Fitzgerald (1898) described them as such during his travels through Zanzibar in 1898. Fitzgerald wrote then that the house and sugar mill were active around 25 years prior, in 1873 (Fitzgerald 1898: 521–523). The depiction of Frazer's sugar mill and house would seemingly contradict the depiction of Marahubi as a ruin if the map were representative of a single snapshot in time. Imam Sherif Khan Bahadur likely did not carry out surveys prior to 1892, but these details show that this map was made from a composite of surveys which were only finished in 1899 at the earliest.

Figure 2 1907 map of Zanzibar, with inset showing detail. The inset is shown in high resolution in the Supplementary Materials section in Figure I.

The map was the main cartographic source for Zanzibar prior to a more recent map created in 1984–85 (Horton, pers. comm.). However, the map has been overlooked by both historians and archaeologists of the colonial period in Zanzibar, despite the wealth of information it contains regarding settlement, land use, and geography on the island during the late 19th and early 20th centuries. Because

of the map's status as a tool of British imperial dominion in Zanzibar, it records invaluable data for understanding the composition of rural areas on the island. Preserved in the map is a settlement system that was surveyed while the plantation system was still fully developed, around the same time that slavery was being abolished, from 1896 to the early 1900s (Cooper 1977: 122). It contains

detailed information regarding Zanzibar's villages, wells, hydrology, roads, and other features at the end of the 19th century. As such, the map is a unique source of information for understanding the spatial patterns of the 19th-century plantation system, preserving many features that have changed through urban and agricultural development in the 20th and 21st centuries.

2.2.1 Georeferencing and Digitization

Petrie et al. (2019), Garcia et al. (2019) and Green et al. (2019) developed methods for georeferencing and interpreting 1 inch to 1 mile Survey of India maps to glean data related to ancient settlement in northwest India, in the form of anthropogenic mounds that surveyors in the late 19th and early 20th centuries recorded. This study draws on their methods since the map was made in the style of British Survey of India maps. The first step was to georeference the 1907 Zanzibar map to features on a modern basemap of the island. This can be achieved using the Georeferencer in QGIS, and by uploading base maps from the QuickMapServices plugin.

No datum or coordinate system is specified on the map itself, but it was likely created using the Everest 1830, Clarke 1866, or Clarke 1880 datum (Mugnier 2021; Petrie et al. 2019). The map has longitude and latitude graticules with specified coordinate points, which were converted into a point vector file. Though the unreferenced map aligns nearly perfectly with these points when projected in the Clarke 1880 datum, the features on the map can be up to a kilometer or more off from their actual locations when the map is georeferenced in this way. This inaccuracy may be due to mapmaker error, or changes to the map when it was printed in 1907 or digitized in recent years. Other examples of Survey of India maps have also been found to be internally inconsistent in this way (Petrie et al. 2019). Given this problem, the solution following Petrie et al. (2019) was to georeference features on the map by hand using the WGS 1984 datum, and to allow the graticules to distort. Petrie et al. (2019) suggest that the "Adjust" transformation in ESRI ArcGIS may give the best possible results. 111 control points were used to georeference this map with the Adjust transformation in ArcGIS; this was done prior to conceiving of an all-open-access study. However, georeferencing in QGIS is equally advanced and while the Adjust transform is not available in QGIS 3.28, the Thin Plate Spline transform may give comparable results since it similarly allows for rubber-sheeting with a large number of control points. After georeferencing, map features were digitized by hand by creating point, line, and polygon features. The following sections describe these features.

2.2.2 Settlements and roads

The map depicts settlements and a road network, digitized in Figure 3. There are 489 settlements, with squares drawn in varying numbers to represent village size. Imam Sherif Khan Bahadur's survey methodology is not apparent, but
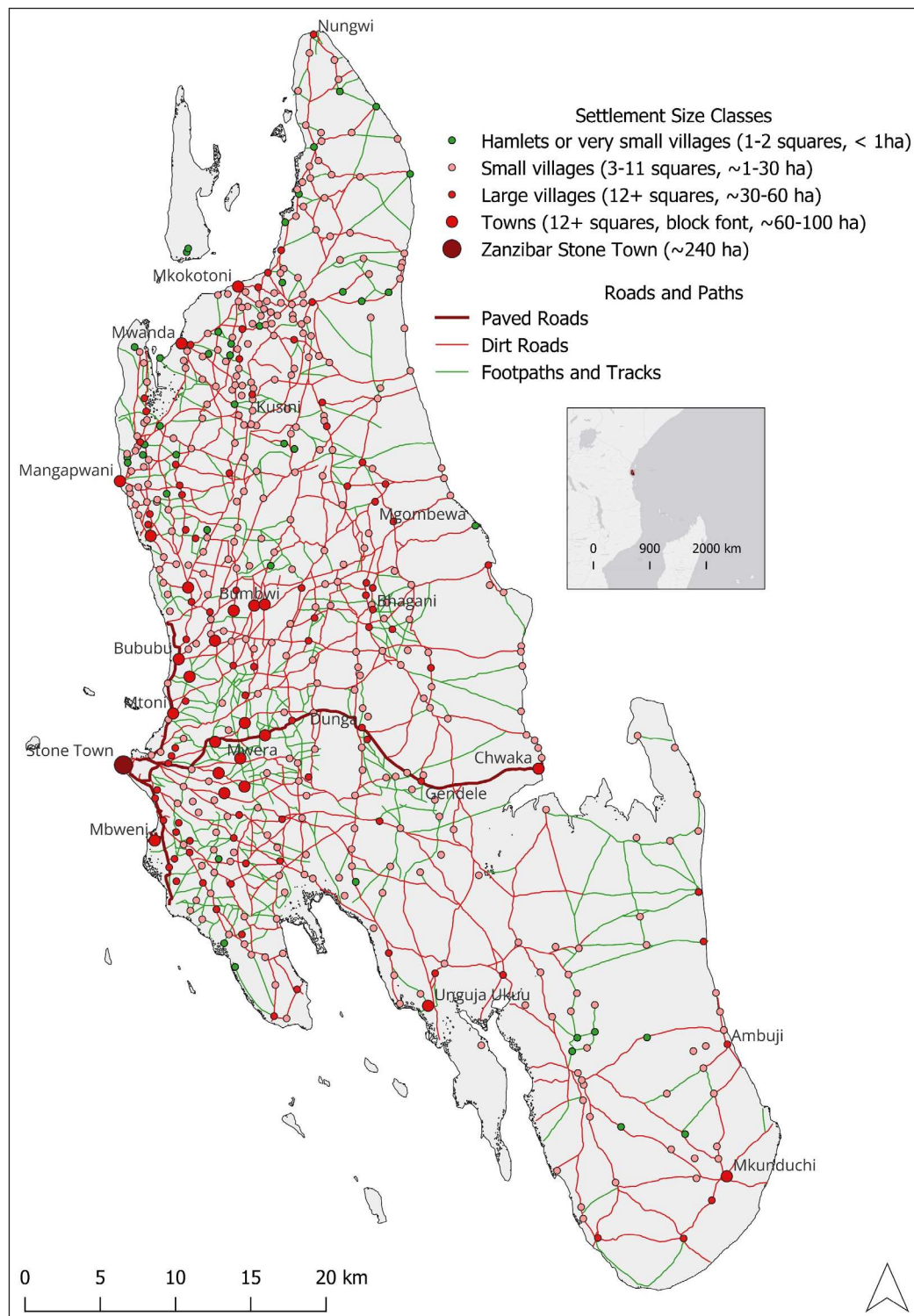
from the map it is likely that two methods were used to indicate settlement size and importance. The first method was through the illustration of squares, which represent settlement areas. Squares are not drawn to the scale of a Zanzibari house—by the scale of the map, the average square ranges from 600 to 1000 square meters. A house this size would be a mansion. Rather than indicating the actual sizes of houses, it is likely that overall settlement size was estimated by a surveyor and then indicated by the number of squares drawn. Officially, the legend of the map depicts a scatter of squares and describes this as a village, further suggesting that surveyors were drawing settlement squares with the aim of capturing settlement size, rather than individual households. The legend also states: "N.B.—Very few of the *Miji* (settlements) in Zanzibar Island are compact villages, the houses are rather scattered over each district". This is the case today as well in many areas.

The second method for distinguishing settlement types is typographical—there are Latin letters in italics with capitalized first letters and additional lowercase letters to name most settlements, as well as to indicate place names. Bolded, non-italic letters are used to name larger towns, also with capitalized first letters and additional letters in lower-case. Finally, bolded, non-italic, and all-capitalized letters are used for Zanzibar Stone Town, the largest settlement on the map. The inset in Figure 2 shows an example of all three types. Zanzibar Stone Town is a city, Mtoni is a town, and Gulioni, Mianzini, and Miwaleni are three of many villages.

While the map legend describes the squares as representing villages, they can be interpreted to represent a settlement hierarchy from their placement and count in conjunction with typographical differences. Square counts for each village appear to be significant and relate to different sized settlements: 1) hamlets or very small villages, 2) small villages, and 3) large villages. Settlements were divided into these 5 size classes based on their typography and number of settlement squares depicted. A sample of each size class was measured in area, to convert square counts into estimated average settlement size, in hectares. Table 3 shows these estimated size classes.

This method of distinguishing settlements by size is an imposition for the sake of regional analysis. A contrasting perspective is the view derived from mid-20th century ethnographies of the Swahili, which divided permanent settlements into "stone-towns" and "country-towns", irrespective of size. These towns, though differentiated by the degree of political and economic specialization, functioned similarly as places that were the basis of social rights for their residents. They were also characterized by different forms of production and trade, with stone-towns emphasizing mercantile activity and country-towns emphasizing agricultural production (Horton and Middleton 2000: 55–58). Since the legend of the 1907

Figure 3 Settlements and road network of the 1907 map.

SIZE CLASS	NUMBER OF SQUARES	FONT EXAMPLE	ESTIMATED SIZE	COUNT
Hamlet/Very Small Village	1–2	*Kongeni*	≤1 ha	42
Small Village	3–11	*Gamba*	1–30 ha	337
Large Village	12+	*Maangani*	30–60 ha	86
Town	12+	**Mkokotoni**	60–100 ha	23
City	n/a	**ZANZIBAR STONE TOWN**	~240 ha	1

Table 3 Size classes for settlement in the 1907 Zanzibar map.

map specifically refers to villages as *miji*, it is likely that the surveyors had some familiarity with an idealized Swahili system of land tenure. Nevertheless, the size-based and typographical differences that are visible on the map attest to the material differences in settlements that the map makers encountered and adapted to as they produced their survey.
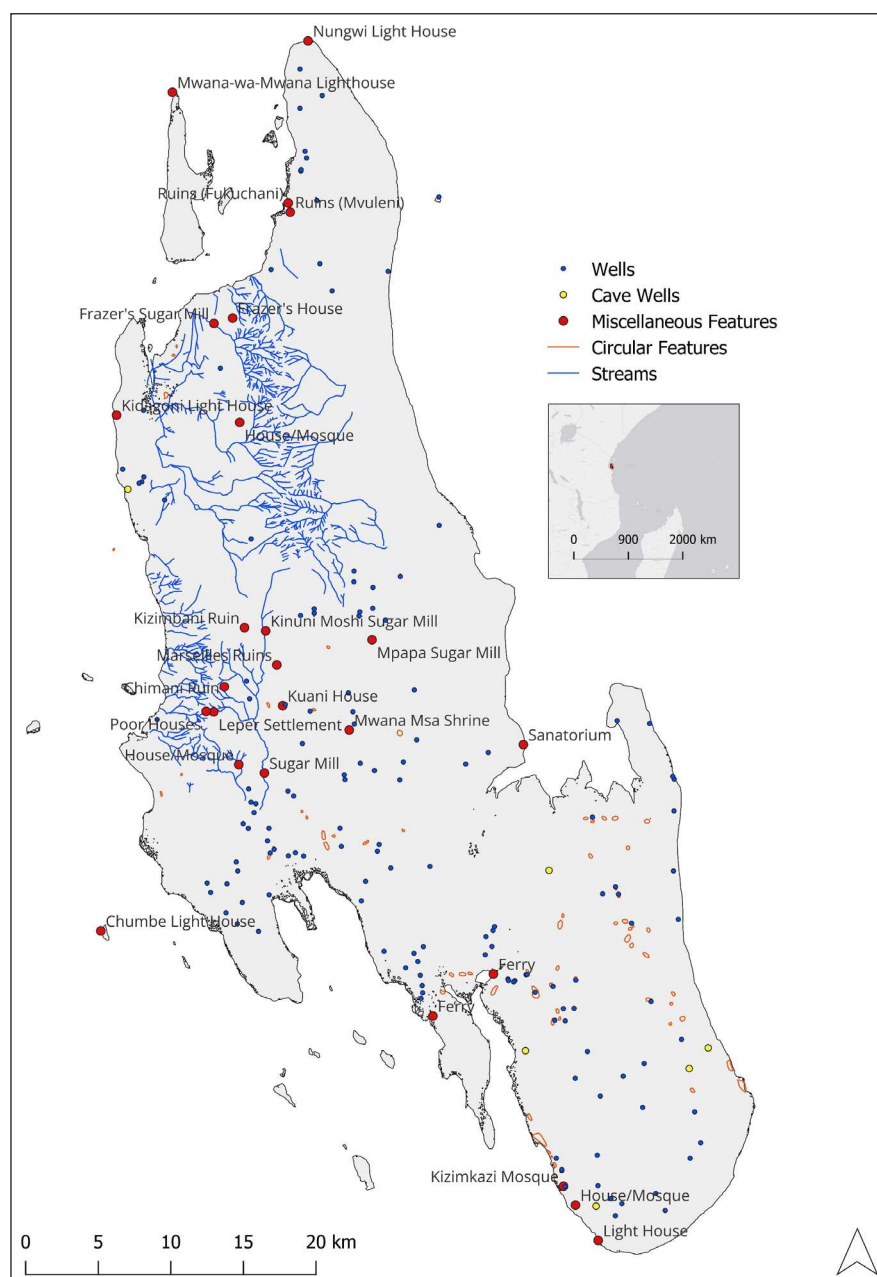
In addition to the settlements, the map also depicts a network of roads, paths, tracks, and other ambiguous dotted lines on the map. The longest paved or "metalled" road during this time ran from Zanzibar Stone Town to Chwaka, connecting the east and west of the island. This road was under construction in the early 1890s (Owens 2007), and its presence on the map may indicate that it was just finished when the map was completed. Surveyors recorded two other paved roads on the map as

well. One went north from Stone Town to Bububu, and the other went south to Stone Town to Mbweni.

Other roads are those which the legend calls "village roads"; these are likely dirt paths for foot traffic, and may have also been accessible to mules, camels, horses, and carts. A third category is called "Other Roads and Tracks". In practice, these are likely not qualitatively distinguishable from village roads in that they were also dirt footpaths, though possibly smaller and less frequently used.

2.2.3 Streams, Wells, and Other Miscellaneous features

Figure 4 shows a map of streams, wells, and other miscellaneous features on the 1907 map. Streams are drawn as black lines on the 1907 map and labeled in italic

Figure 4 Streams, wells, and other miscellaneous features.

letters by their Swahili names. This stream map likely represents the oldest model of hydrology on Zanzibar, existing prior to many landscape transformations which occurred during the latter half of the 20th century. The map shows that major streams did not flow in the south, east and far north of the island, where porous limestone bedrock draws water underground. It also shows that the courses of the larger streams of the early 20th century were slightly different than comparable streams today. This may be due to variations in local geology, urban development, or changes in landforms that have altered the courses of streams since the early 20th century.

Wells are indicated on the map by circles. Their preponderance in karstic limestone areas far from the streams of the northwest region aligns inversely with the stream network—the wells are most common in places where streams are not shown above ground. The lack of wells in places with above-ground streams suggests that people in the early 20th century relied considerably on above-ground stream water for daily use where it could be found, and dug wells in places where stream water was not available. Six wells in the south are marked as either "*Cave Wells*" or "*C.W.*"*,* which likely also stands for cave well. One of these cave wells is Kuumbi Cave, a well-known site from the late Pleistocene to late Holocene (Shipton et al. 2016). The data showing other cave wells contained in this map may point to other cave site locations.

Other miscellaneous features on the 1907 Zanzibar map include lighthouses, "poor houses", a leper colony, a sanatorium, ferries, sugar mills, ruins, and a depiction of buildings with steepled roofs that may represent mosques or large houses. Two other areas are places called the Mwana Msa Shrine and the Kuani House. Finally, dotted lines which form small circles are not described in the legend but appear to correspond to some labeled settlements in the south and east of the island. It is unclear what these circles represent, but further research might investigate whether they correspond to abandoned settlements, as boundary markers or raised areas. Comparisons of these circles with satellite imagery are inconclusive. Some circles fall over modern field plots or settlements, but others fall in areas that are today covered in brush. Horton and Middleton (2000: 56) describe historical areas of built-up soil in the south and east where village communities repeatedly constructed and demolished earth and thatch houses; it is possible that these circular features could represent the mounds created by this practice. Further research and ground-truthing in the southern region of Zanzibar might clarify this question. As is the case with Survey of India maps in India (Green et al. 2019; Petrie et al. 2019), surveyors may have unknowingly mapped archaeological sites by recording mounds as landscape features.

Other archaeologically significant features are the five places listed as ruins on the map. Two of these locations are the known Portuguese-period sites of Fukuchani and Mvuleni, which LaViolette and Norman (2023) have recently investigated. The three others are the Chimani Ruin, the Kizimbani Ruin, and the Marseilles Ruin, all of which are located just northeast of Stone Town near Mwera. The Marseilles Ruin may be the site of the Marseilles plantation, the site of a battle where Barghash bin Said surrendered to his brother the Sultan Sayyid Majid in 1865 after an abortive attempt to seize the throne of Zanzibar (Ruete 1888: 107). Further investigation is needed to determine whether any of these ruins still exist today.

2.3 STEPS 4–7: CONVERTING SETTLEMENT CLASSES TO TRAINING FEATURES, CALCULATING ZONAL STATISTICS, AND DETERMINING WEIGHT CLASSES

The digitization process outlined above produced five settlement classes from the 1907 map: 1) hamlets/very small villages, 2) small villages, 3) large villages, 4) towns, and 5) the main urban center, Zanzibar Stone Town. To analyze these settlements in relation to raster data, they were converted into polygons that reflect their area. Buffer polygons for each class were created, encompassing the average area of each class. To simplify analyses, these buffer polygons were then merged into two settlement class groups: small settlements (n = 379) and large settlements (n = 110). Small settlements were comprised of hamlets/very small villages and small villages. Large settlements were comprised of large villages, towns, and Zanzibar Stone Town. These two polygon vector files constituted the training data for two distinct models.

The Majority statistic was calculated across these two training feature classes for each zonal raster image in QGIS using the Zonal Statistics process. This statistic reports which unique pixel value (reflecting a zone) is most numerous within the space of a settlement polygon; summarizing this statistic using the Statistics by Categories process gives a count of settlement polygons per zone for each raster image. Table 4 through Table 11 display this data for each raster image and include statistics that were used to weight zones with the densest count of settlement polygons.

Density was calculated by dividing the count of settlement polygons per zone by the area of each zone, to derive a count of settlement polygons per square kilometer. The zone with the highest density of settlements per raster image was selected for weighting in the final predictive model. Calculating the zone with the highest density of settlement polygons is a better measure of zone favorability than a simple

count of settlement polygons per zone, since each zone within a raster image can vary significantly in size. To compare the evenness or unevenness of the distribution of settlements across zones, another statistic calculated was the coefficient of variation (CV), found by dividing the standard deviation of settlement polygons per zone by the mean of settlement polygons per zone.

The following sections show settlement class distributions across the eight zonal raster images. Each table is divided into two groups, showing the distribution of small settlements (n = 379) and large settlement (n = 110) across each zonal raster image. Each table also lists the zone with the highest density of settlements for each settlement group, and the coefficient of variation for each settlement group. Maps of each zonal raster image are available as supplementary materials.

2.3.1 Zonal statistics across aspect zones
A zonal raster image for aspect derived from 30 m free SRTM imagery from USGS shows hillslope orientation. Table 4 shows settlement classes from the 1907 map across these aspect zones. East and west-facing slopes have the highest site density as well as the highest site counts for both classes. No other patterns are apparent. See Figure A in the Supplementary Materials section for a map of this raster image.

2.3.2 Zonal statistics across elevation zones
Table 5 shows the 2019 sites distributed across five elevation zones, categorized using a Natural Breaks (Jenks) algorithm on a digital elevation model from SRTM 30 m imagery. Higher elevation zones are favored. The 69–135 m zone is most densely settled for small settlements, while the 48–69 m zone was most densely settled for large settlements. See Figure B in the Supplementary Materials section for a map of this raster image.

2.3.3 Zonal statistics across geology zones
Seven geology zones exist on Zanzibar (Colbert et al. 1987, also see Hardy et al. 2015). These are catenas of M3 sandy clay marl, Q2 coralline limestone, M1 Miocene limestone, a Q2/M1 mixture, Q1 recent deposits, a Q2/Q3/M1 mixture, and mangrove zones with no data. Table 6 shows settlement classes from the 1907 map across these zones. Settlements of both classes are most numerous and most dense in areas of M3 Sandy Clay Marl. See Figure C in the Supplementary Materials section for this raster image.

2.3.4 Zonal statistics across rainfall zones
Rainfall zones on Zanzibar (Colbert et al. 1987) are divided into three zones: 1000–1500 mm, 1500–2000 mm, and 2000–2500 mm of rainfall per year. Table 7 shows the zonal statistics for settlement classes across these zones. Large settlements are most dense in the

	NORTH	NE	EAST	SE	SOUTH	SW	WEST	NW	TOTAL	HDW	CV
area (km²)	190	195	223	193	177	199	223	195	1595		
Small Settlements (count)	35	46	71	31	28	38	72	58	379		
Small Settlements (percent)	9%	12%	19%	8%	7%	10%	19%	15%	100%		0.37
Small Settlements per km²	0.18	0.24	0.32	0.16	0.16	0.19	0.32	0.30	0.24	West	
Large Settlements (count)	9	7	27	7	7	15	22	16	110		
Large Settlements (percent)	8%	6%	25%	6%	7%	13%	20%	15%	100%		0.56
Large Settlements per km²	0.05	0.04	0.12	0.04	0.05	0.07	0.10	0.08	0.07	East	

Table 4 Zonal statistics for map settlement classes and aspect zones.

	0–17 m	17–31 m	31–48 m	48–69 m	69–135 m	TOTAL	HIGHEST DENSITY	CV
area (km²)	545	542	310	144	72	1613		
Small Settlements (count)	93	90	103	59	34	379		
Small Settlements (percent)	25%	24%	27%	16%	9%	100%		0.38
Small Settlements per km²	0.17	0.17	0.33	0.41	0.47	0.23	69–135 m	
Large Settlements (count)	26	32	27	18	7	110		
Large Settlements (percent)	24%	29%	25%	16%	6%	100%		0.44
Large Settlements per km²	0.05	0.06	0.09	0.13	0.10	0.07	48–69 m	

Table 5 Zonal statistics for map settlement classes and elevation.

	M3 SANDY CLAY MARL	MANGROVE (NO DATA)	Q2 CORALLINE LIMESTONE	M1 MIOCENE LIMESTONE	Q2/M1 MIXTURE	Q1 RECENT DEPOSITS	Q2/ Q3/M1 MIXTURE	TOTAL	HIGHEST DENSITY WITHIN	CV
area (km²)	218	47	612	188	409	102	59	1635		
Small Settlements (count)	104	2	113	69	50	32	9	379		
Small Settlements (percent)	27%	1%	30%	18%	13%	8%	2%	100%		0.81
Small Settlements per km²	0.48	0.04	0.18	0.37	0.12	0.31	0.15	0.23	M3 Sandy Clay Marl	
Large Settlements (count)	32	1	34	23	7	12	1	110		
Large Settlements (percent)	29%	1%	31%	21%	6%	11%	1%	100%		0.89
Large Settlements per km²	0.15	0.02	0.06	0.12	0.02	0.12	0.02	0.07	M3 Sandy Clay Marl	

Table 6 Zonal statistics for map settlement classes across the geology zones.

	1000–1500 mm	1500–2000 mm	2000–2500 mm	TOTAL	HIGHEST DENSITY WITHIN	CV
area (km²)	666	843	104	1613		
Small Settlements (count)	94	224	61	379		
Small Settlements (percent)	25%	59%	16%	100%		0.68
Small Settlements per km²	0.14	0.27	0.59	0.23	2000–2500 mm	
Large Settlements (count)	15	86	9	110		
Large Settlements (percent)	14%	78%	8%	100%		1.17
Large Settlements per km²	0.02	0.10	0.09	0.07	1500–2000 mm	

Table 7 Zonal statistics for settlement classes across rainfall zones.

1500–2000 mm zone, while smaller settlements are most dense in the 2000–2500 mm zone. See Figure D in the Supplementary Materials section for a map of this raster image.

2.3.5 Zonal statistics across reef distance buffer zones
Reefs (Khamis et al. 2017: 120) were buffered by distance to create a zonal raster image for the island. Table 8 shows the zonal statistics for settlement classes across these zones. Small settlements were densest within 500 meters of reefs, while large settlements were densest within 3 kilometers. See Figure E in the Supplementary Materials section for a map this raster image.

2.3.6 Zonal statistics across slope degree zones
Slope degree zones, also derived from 30 m SRTM imagery from USGS, fall into three categories: 0-3-degree slope, 3-10-degree slope, and areas with >10-degree slope. Table 9 shows settlement classes across these zones. No site is found on a slope of 10 degrees or more, and both settlement classes are most dense in 0-3-degree slope areas. See Figure F in the Supplementary Materials section for a map this raster image.

2.3.7 Zonal statistics across soil type zones
Table 10 shows settlement distributions across soil type zones, derived from Khamis et al. (2017: 120). Swahili speakers distinguish five soil types on Zanzibar based on color, organic composition, and depth: *kinamo* (deep dark sandy clays), *kinongo* (deep red laterites of variable clay and organic content), *uwanda* (shallow red laterites with moderate organic content, over coralline limestone bedrock), *mchanga* (deep sandy soil built on recent alluvial sediments) and *maweni* (shallow dark laterites with high organic content over coralline limestone bedrock) (Khamis et al. 2017: 120). Small settlement classes are associated with *kinongo* soils while large settlement classes are associated with *mchanga* soils. See Figure 1 above or Figure G in the Supplementary Materials section for a map of this raster image.

2.3.8 Zonal statistics across stream buffer zones
Table 11 shows the settlement classes across a zonal raster image of buffered distances from streams, which were digitized from the 1907 map of Zanzibar. Both settlement classes are densest within 500 meters of streams. See Figure H in the Supplementary Materials section for a map of sites across this raster image.

	100 m FROM REEFS	500 m FROM REEFS	1 km FROM REEFS	3 km FROM REEFS	5 km FROM REEFS	10 km FROM REEFS	15 km FROM REEFS	TOTAL	HDW	CV
area (km²)	1	6	29	287	318	671	300	1612		
Small Settlements (count)	0	6	11	61	66	167	68	379		
Small Settlements (percent)	0%	2%	3%	16%	17%	44%	18%	100%		1.07
Small Settlements per km²	0.00	1.00	0.38	0.21	0.21	0.25	0.23	0.24	500 m	
Large Settlements (count)	0	0	0	23	24	51	12	110		
Large Settlements (percent)	0%	0%	0%	21%	22%	46%	11%	100%		1.20
Large Settlements per km²	0.00	0.00	0.00	0.08	0.08	0.08	0.04	0.07	3 km	

Table 8 Zonal statistics for settlement classes across reef distance buffer zones.

	0–3-DEGREE SLOPE	3–10-DEGREE SLOPE	>10–DEGREE SLOPE	TOTAL	HIGHEST DENSITY	CV
area (km²)	1208	378	8	1594		
Small Settlements (count)	313	66	0	379		
Small Settlements (percent)	83%	17%	0%	100%		1.31
Small Settlements per km²	0.26	0.17	0.00	0.24	0–3-degree slope	
Large Settlements (count)	100	10	0	110		
Large Settlements (percent)	91%	9%	0%	100%		1.50
Large Settlements per km²	0.08	0.03	0.00	0.07	0–3-degree slope	

Table 9 Zonal statistics for settlement classes across slope degree zones.

	MCHANGA	KINAMO	KINONGO	MAWENI	UWANDA	TOTAL	HIGHEST DENSITY WITHIN	CV
area (km²)	145	206	338	761	164	1614		
Small Settlements (count)	49	67	139	96	28	379		
Small Settlements (percent)	13%	18%	37%	25%	7%	100%		0.57
Small Settlements per km²	0.34	0.33	0.41	0.13	0.17	0.23	*Kinongo*	
Large Settlements (count)	29	17	56	4	4	110		
Large Settlements (percent)	26%	15%	51%	4%	4%	100%		0.98
Large Settlements per km²	0.20	0.08	0.17	0.01	0.02	0.07	*Mchanga*	

Table 10 Zonal statistics for settlement classes and soil types.

	WITHIN 100 m OF STREAMS	WITHIN 500 m OF STREAMS	WITHIN 1 km OF STREAMS	WITHIN 3 km OF STREAMS	BEYOND 3 km FROM STREAMS	TOTAL	HIGHEST DENSITY WITHIN	CV
area (km²)	103	197	105	186	1021	1612		
Small Settlements (count)	24	92	38	61	164	379		
Small Settlements (percent)	6%	24%	10%	16%	43%	100%		0.73
Small Settlements per km²	0.23	0.47	0.36	0.33	0.16	0.24	Within 500 m	
Large Settlements (count)	3	36	14	28	29	110		
Large Settlements (percent)	3%	33%	13%	25%	26%	100%		0.60
Large Settlements per km²	0.03	0.18	0.13	0.15	0.03	0.07	Within 500 m	

Table 11 Zonal statistics for settlement classes across the 1907 stream buffer raster image.

2.3.9 Weighting Zonal Raster Images

The next step was to weight zonal raster images based on the training features. Zones in which settlement class training data is most dense were selected as factors, and then weighted based on the coefficient of variation (CV). For each settlement class, the standard of deviation was calculated for the range of CV values, and then a CV range was created by adding and subtracting the standard of deviation from the mean CV. This range was then divided up by equal intervals, and each interval threshold was assigned a weight class. This created an internally consistent range of weight class thresholds for each training feature dataset. The CV for each zonal raster image was assigned a score based on these weight class thresholds.

One knowledge-based adjustment was made: the CV for large settlements across aspect zones was reduced from 3 to 2, for three reasons. First, the CV value is 0.56, only 0.01 points into weight class 3, producing an edge effect. Secondly, visual inspection of both settlement classes across aspect zones confirms that the distribution mostly reflects the fact that east and west facing slopes are the most numerous slopes on the island, due to the fact that Zanzibar's hill system runs like a spine from south to north. It was likely that settlement choices in the past were not strongly influenced by aspect zones, but rather fell in a relatively random distribution with regard to hillslope orientation. This is because hill slopes on Zanzibar are mild, and the equatorial sun means that hill orientation is less of a factor for agricultural production than in climates closer to the Earth's poles. Third, the 30 m resolution of the aspect raster image means that nearly all large settlement classes contain multiple aspect zones within them, so the zonal statistics for each training feature can vary considerably due to very slight adjustments in settlement placement by the manual digitizer. This third problem was not the case with any other class and was less the case with the small class of training features since the areas of these settlements tended to encompass far fewer aspect zones. To compensate for these factors and to minimize the impact of the aspect raster image on

weighting the large settlement zones, a knowledge-based adjustment reduced the aspect weight class for large settlements from 3 to 2, while still maintaining the most favored zone chosen by the Majority statistic.

Tables 12 and 13 depict the weight class thresholds for each model. Tables 14 and 15 depict the weighted scores for each zonal raster image for both settlement classes, based on the zonal statistics above. This table also summarizes the most favored zones for both settlement classes across all zonal raster images. For small settlement classes, this produced Model A; for large settlement classes, this produced Model B.

The eight zonal raster images were reclassified to reflect weight classes, using the Reclassify by Table process. Favored zones were classified with a unique pixel value of each weight, and all other pixels were assigned a value of zero. This was done twice, once for each model.

WEIGHT CLASSES	CV THRESHOLDS FOR SMALL SETTLEMENT CLASSES (MEAN CV = 0.74 + OR – STDDEV OF 0.32)
1	0–0.09
2	0.09–0.41
3	0.41–0.74
4	0.74–1.06
5	1.06–1.38

Table 12 CV thresholds and associated weight class values, based on mean CV and CV std. dev. for small settlement classes.

WEIGHT CLASSES	CV THRESHOLDS FOR LARGE SETTLEMENT CLASSES (MEAN CV = 0.92 + OR – STDDEV OF 0.37)
1	0–0.18
2	0.18–0.55
3	0.55–0.92
4	0.92–1.29
5	1.29–1.65

Table 13 CV thresholds and associated weight class values, based on mean CV and CV std. dev. for large settlement classes.

RASTER IMAGES	MODEL A: MOST FAVORED CLASS FOR SMALL SETTLEMENTS	CV	ZONE WEIGHT
Aspect	West	0.37	2
Elevation	69–135 m	0.38	2
Geology	M3 Sandy Clay Marl	0.81	4
Stream Buffer	Within 500 m	0.73	3
Rainfall	2000–2500 mm	0.68	3
Reefs	500 m from reefs	1.07	5
Soil Types	*Kinongo*	0.57	3
Slope Degree	0–3-degree slope	1.31	5

Table 14 Model A, favored zones, CVs, and weight classes for small settlement class training features across eight zonal raster images.

RASTER IMAGES	MODEL B: MOST FAVORED CLASS FOR LARGE SETTLEMENTS	CV	ZONE WEIGHT
Aspect	East	0.56	2 (knowledge-based adjustment, down from 3)
Elevation	48–69 m	0.44	2
Geology	M3 Sandy Clay Marl	0.89	3
Stream Buffer	Within 500 m	0.60	3
Rainfall	1500–2000 mm	1.17	4
Reefs	3 km from reefs	1.20	4
Soil Types	*Mchanga*	0.98	4
Slope Degree	0-3-degree slope	1.50	5

Table 15 Model B, favored zones, CVs, and weight classes for large settlement class training features across eight zonal raster images.

2.4 STEPS 8–9: PRODUCING THE PREDICTIVE MODEL

The two sets of weighted raster images from these models were summed using the Raster Calculator process in QGIS to produce two archaeological predictive models, Model A (Figure 5) and Model B (Figure 6). The summed raster images contained values from 1–22 in the case of the small settlement model and 1–25 in the case of the large settlement model, but both were reclassified into five zones of site detection probability using a Natural Breaks (Jenks) algorithm: Very Low, Low, Medium, High, and Very High (see Diwan 2020: 152). Since QGIS 3.28 does not yet support this reclassification algorithm for raster images, a workaround was to convert the raster images into polygons, re-symbolize the polygon files with graduated symbols using Natural Breaks (Jenks), and then manually reclassify the original raster images using the Reclassify by Table tool, with the values generated from the polygon symbology.

3. RESULTS

Summing weighted zonal raster images from two different settlement classes produced two raster images that reflect site detection probability zones, Model A (Figure 5, from small settlement classes) and Model B (Figure 6, from large settlement classes). The models predict site locations based on the density of training features within specific environmental zones. Both training feature classes were strongly associated with flat, level ground (0-3-degree slope), areas within 500 meters of above ground streams, and M3 Sandy Clay Marl geology zones. Both were also associated with east and west aspect zones, but as discussed in the prior section, this was likely not a strong factor influencing settlement location. However, smaller

settlement classes were more strongly associated with zones related to agricultural production and subsistence: higher elevation zones (preferential for clove plantations, see Sheriff et al. 2016), high rainfall areas (2000–2500 mm), *kinongo* soils (deep soils favorable for subsistence agriculture and earth and thatch house construction), and areas within 500 meters of offshore reefs. In contrast, larger settlement classes were associated with lower rainfall areas, lower elevation zones, sandy *mchanga* soils, and areas further from offshore reefs. The comparison between reef distance is noteworthy—Swahili people in small rural settlements may have preferred to live near reefs because they provided opportunities for subsistence fishing, while people in larger settlements may have preferred open seas without reefs. Reefs would have impeded the movement of larger ships and boats that brought trade to port towns like Mkokotoni, Tumbatu and Zanzibar Stone Town, with their deep-water ports and good anchorage (Fitton 2018). The lack of easily accessible reefs for small-scale fishing around Mkokotoni, Tumbatu and Zanzibar Stone Town may have not only facilitated the arrival of larger boats bringing trade, but also may have stimulated communities in these places to develop larger-scale fishing operations, necessitating greater social cooperation and coordination.

Having considered the environmental factors that structure Models A and B, the next section ground truths the models by comparing them to the locations of major precolonial sites of the late first and early second millennium on Zanzibar (Fitton 2018; Horton and Clark 1985), and to a dataset of sites recovered during a systematic survey across multiple environmental zones in northern Zanzibar in 2019 (Alders 2023). These latter surveys identified and recorded 44 new archaeological sites, with 31 found in the course of a systematic random stratified sample in a region of 32 km².
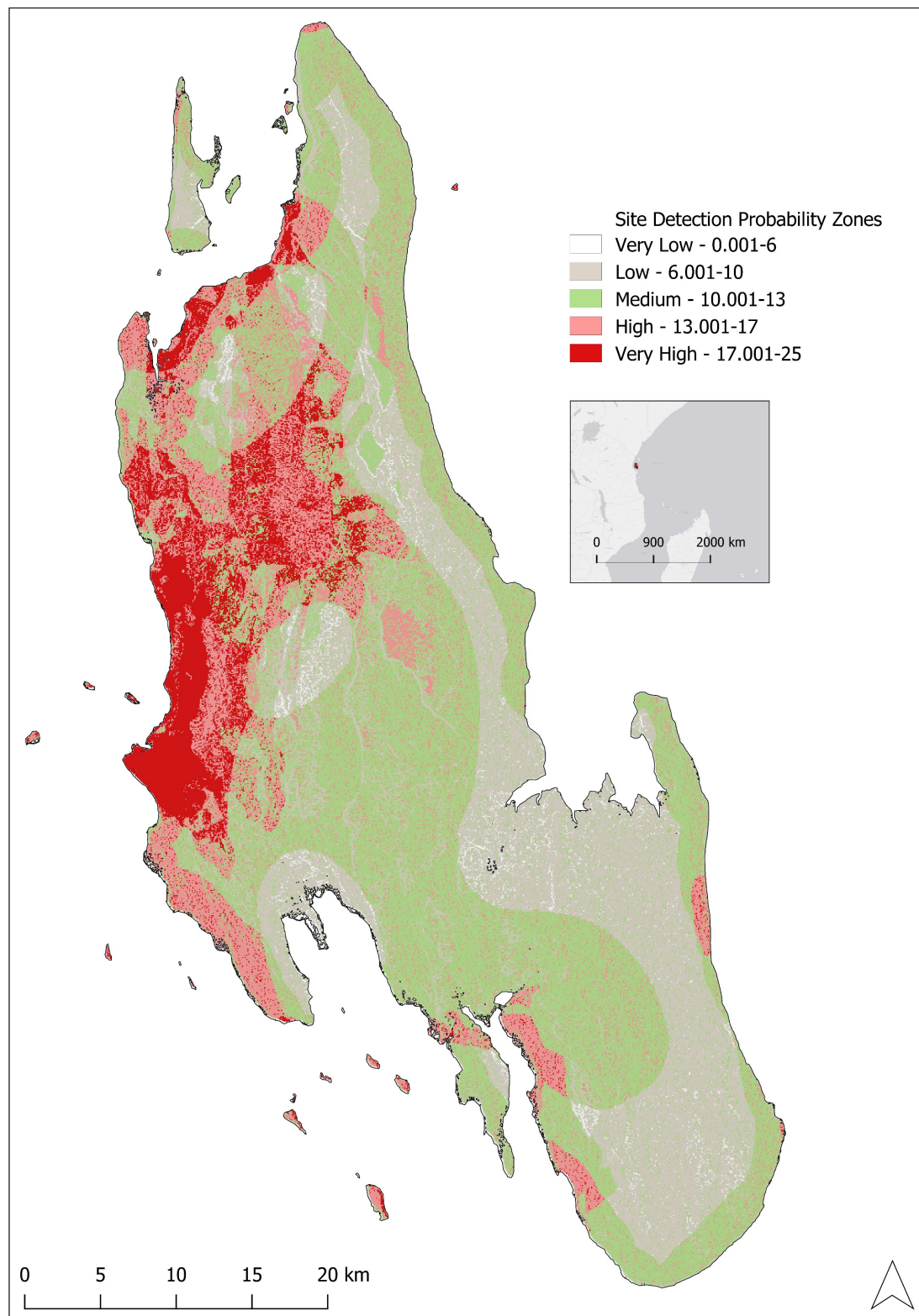
Figure 5 Model A. Site detection probability zones derived from small settlements on the 1907 map, in relation to environmental datasets.

3.1 COMPARING THE MODELS TO MAJOR KNOWN PRECOLONIAL SWAHILI SITES

Tables 16 and 17 show zonal statistics for known major precolonial sites in relation to the site detection probability zones of the two predictive models, and Figures 7 and 8 show the spatial distribution of these sites in relation to the two models.

Six out of eight precolonial sites fall within the High or Very High site detection probability zones from Model B. The fact that Model B predicts these precolonial site locations well attests to similarities in environmental favorability between precolonial site locations and the larger settlements which persisted and grew during the

19th century on Zanzibar, which became the training data for this model. The two outliers are Kuumbi Cave and Tumbatu, which lie in Low and Medium probability zones respectively. Kuumbi Cave is a famous precolonial site on the island, but it was not significantly inhabited by Swahili people. Rather, it is most well-known as one of a handful of late Pleistocene hunter-gatherer settlements on the East African coast (Shipton et al. 2016). Occupied by hunter-gatherers, agricultural suitability was not a concern, and the site was likely favored for the naturally occurring shelter that the cave provided. Similarly, the 11th–15th century town of Tumbatu was not founded as an agricultural center but was rather established by nascent

Figure 6 Model B. Site detection probability zones derived from large settlements on the 1907 map, in relation to environmental datasets.
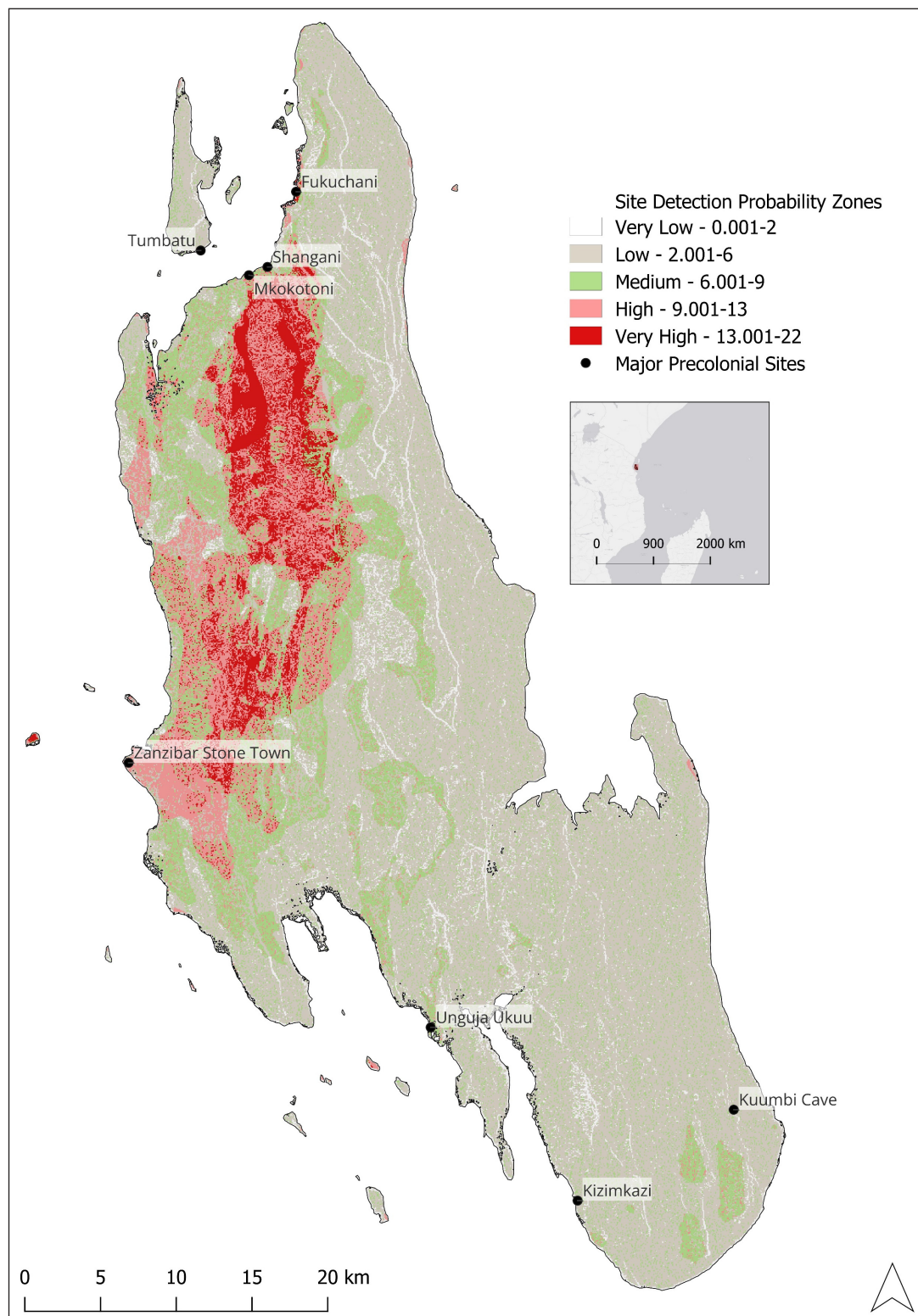
	VERY LOW	LOW	MEDIUM	HIGH	VERY HIGH	TOTAL	HIGHEST DENSITY WITHIN	CV
area (km²)	104	924	308	189	87	1612		
Known Sites (count)	1	3	1	2	1	8		
Known Sites (percent)	13%	38%	13%	25%	13%	100%		0.56
Known Sites per km²	0.0096	0.0032	0.0032	0.0106	0.0115	0.0050	Very High	

Table 16 Known precolonial sites in relation to Model A, based on small settlement classes from the 1907 map.

	VERY LOW	LOW	MEDIUM	HIGH	VERY HIGH	TOTAL	HIGHEST DENSITY WITHIN	CV
area (km²)	48	438	682	292	152	1612		
Known Sites (count)	0	1	1	3	3	8		
Known Sites (percent)	0%	13%	13%	38%	38%	100%		0.84
Known Sites per km²	0.0000	0.0023	0.0015	0.0103	0.0197	0.0050	Very High	

Table 17 Known precolonial sites in relation to Model B, based on large settlement classes from the 1907 map.

Figure 7 Known major precolonial sites on Zanzibar in relation to Model A, which is based on small settlement size classes in the 1907 map.

Figure 8 Known major precolonial sites on Zanzibar in relation to Model B, which is based on large settlement size classes in the 1907 map.

Swahili elites around the 11ᵗʰ century, who may have sought seclusion, security, and access to shipping routes rather than agricultural suitability. The town's residents relied on support from the residents of Mkokotoni across the channel, who may have continually ferried over food and water to the site (Rødland 2021: 254). Given this interdependence, Rødland (2021) has argued that Tumbatu and Mkokotoni formed a single urban landscape. That Model B does not predict Tumbatu's site location well was to be expected, given the specific history of the town.

Model A does not predict the locations of major precolonial as well, with four sites falling into Very Low or Low zones; however, the Very High zone still has the highest density of sites because the zone is the smallest relative to others. The poorer performance of Model A attests to the fact that in comparison to major precolonial sites, slightly different environmental affordances structured the small-scale settlements during the 19ᵗʰ century that were used as training data for this model.

3.2 COMPARING THE MODELS TO 2019 FIELD SURVEY DATA

Tables 18 and 19 show zonal statistics for the sites recorded during field surveys (see Alders 2022; 2023) in relation to the two predictive models. The tables are stratified by site type, considering artifact scatters in fields in the top rows and permanent, village-sized sites in the bottom rows. Figures 9 and 10 show the spatial distribution of these sites across both models.

Model A predicted the site locations of larger, permanently occupied sites in rural inland Zanzibar, with seven out of nine sites falling within High and Very High site detection probability zones. This reflects similarities between environmental affordances which structured the small settlement classes on the 1907 map of Zanzibar, and the larger, village-sized sites recovered archaeologically during survey. For 19th-century sites recovered this was expected; however, the model's ability to predict the location of a precolonial village site also suggests that the model reflects environmental factors that conditioned small-scale settlement for many centuries in rural Zanzibar.

Model A failed to predict the site locations of smaller artifact scatters in fields. These smaller sites represent ephemeral camps or field houses that were occupied during seasonal agricultural labor (see also Walshaw 2015), especially in the eastern region where stony landscapes and a lack of fresh water on the surface prohibit larger settlements in many areas. In these regions, farmers today bring food and water to swidden field plots and camp for several days during clearing and planting. Ceramic scatters and shell piles in these same fields dating to the 11th century at the earliest likely attest to similar land use patterns in the past (Alders 2022: 118–126). In contrast, larger sites in other parts of the survey region likely reflect more permanent occupations, ranging from small hamlets to plantation estates to the large, dispersed village or town of Chaani, which spanned at least 60 hectares by the 19th century (Alders 2023). The failure of this model to predict these smaller ephemeral camp sites reflects the fact the model was trained with permanent settlement classes, the smallest of which (Hamlets/Very Small Villages on the 1907 map) was still larger than the ephemeral camp sites that surveys recorded.
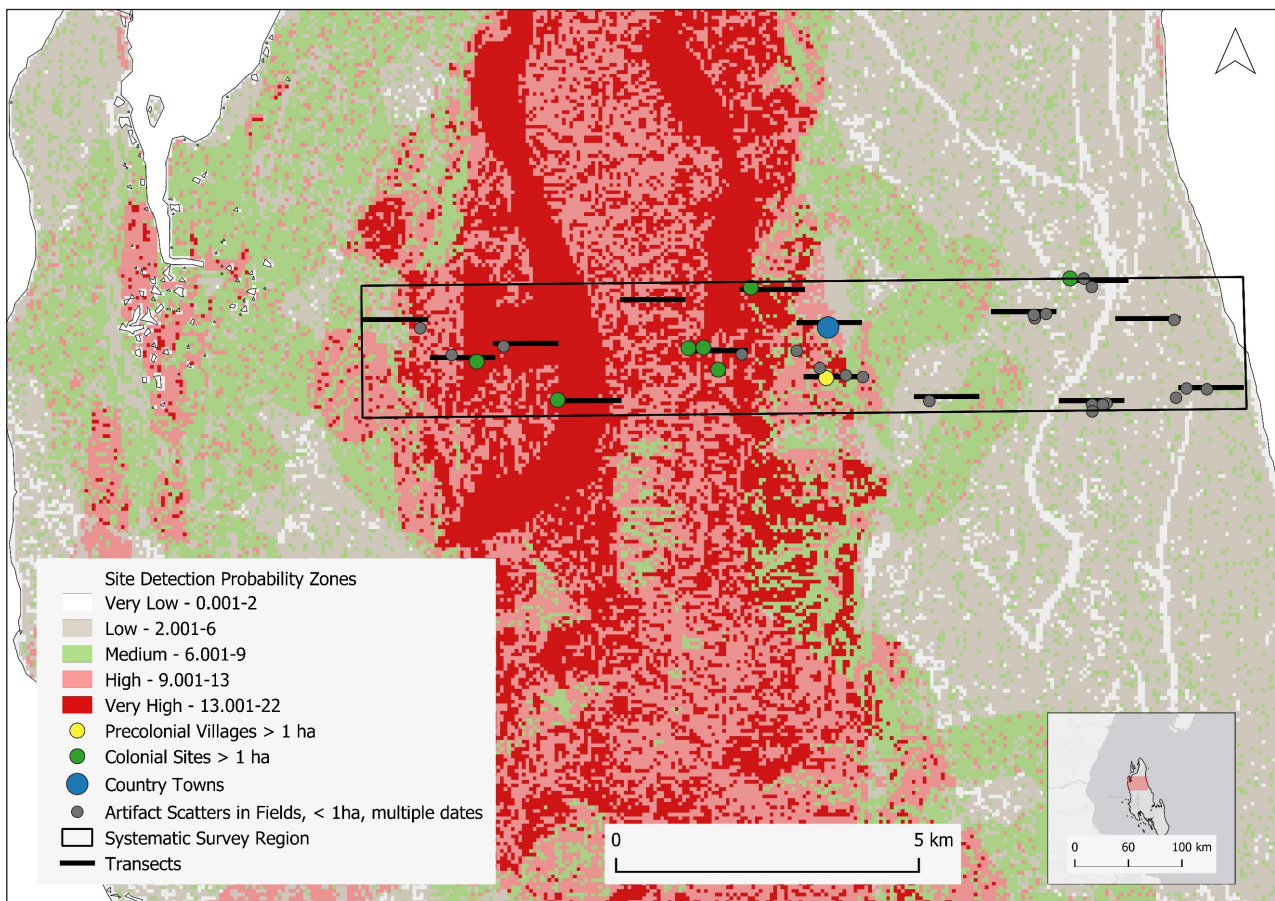
Model B was less successful at predicting the locations of small-scale sites recovered through field survey, with Low and Medium site detection zones having the highest density of sites for small artifact scatters and larger permanent sites, respectively. This result suggests that Model B, which was trained using large settlement classes from the 1907 map, reflects slightly different environmental affordances that did not apply to small-scale settlement in rural inland Zanzibar. Environmental conditions that influenced the locations of larger settlements were less constraining to small-scale communities in rural areas.

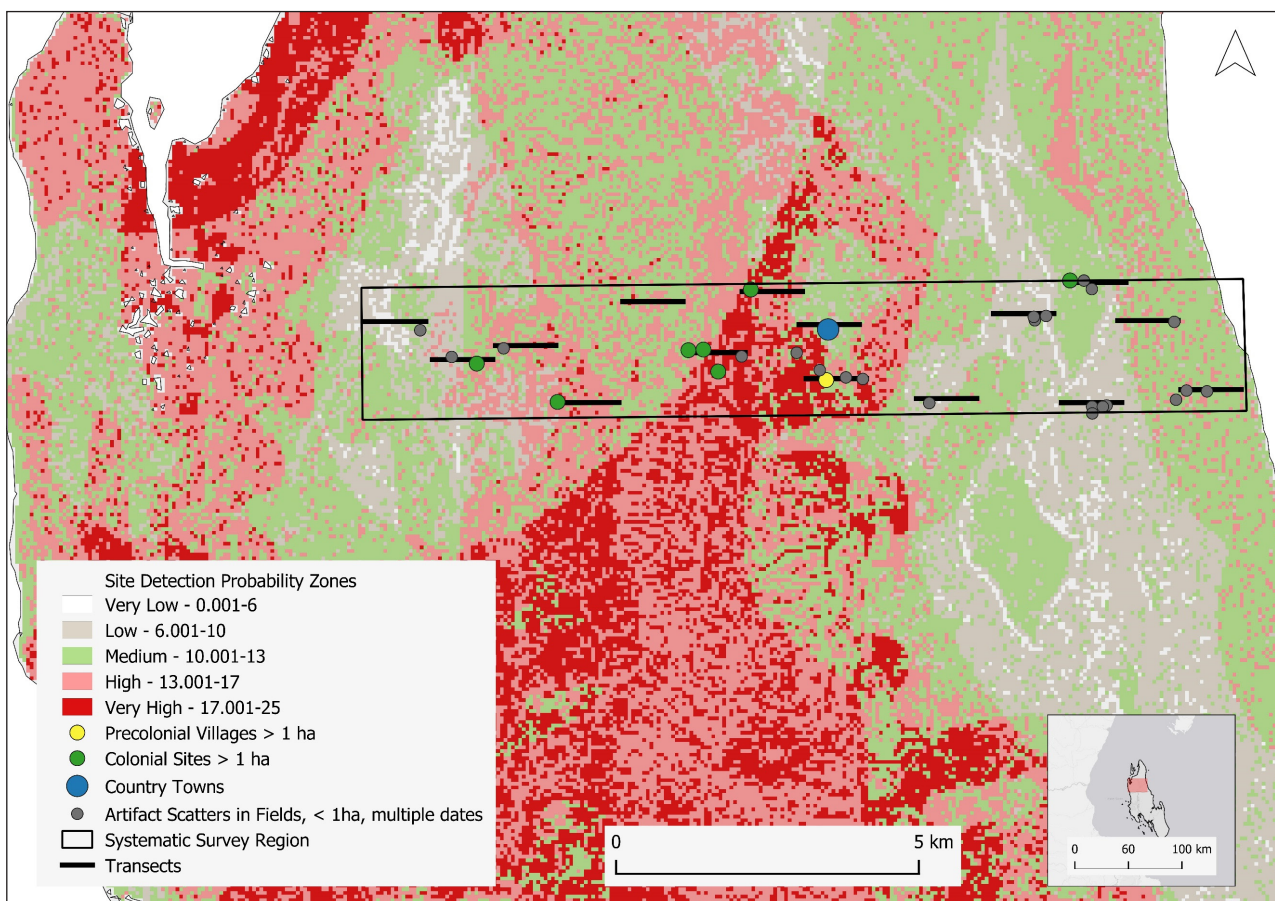|	VERY LOW	LOW	MEDIUM	HIGH	VERY HIGH	TOTAL	HIGHEST DENSITY WITHIN	CV
area (km^2)	1	9	5	7	10	32		
Artifact Scatters in Fields (count)	0	14	1	4	3	22		
Artifact Scatters in Fields (percent)	0%	64%	5%	18%	14%	100%		1.27
Artifact Scatters in Fields per km^2	0.0000	1.5556	0.2000	0.5714	0.3000	0.6875	Low	
Permanent Sites (count)	0	1	1	3	4	9		
Permanent Sites (percent)	0%	11%	11%	33%	44%	100%		0.91
Permanent Sites per km^2	0.0000	0.1111	0.2000	0.4286	0.4000	0.2813	High	

Table 18 2019 field survey sites in relation to Model A, which is based on small settlement classes from the 1907 map.

	VERY LOW	LOW	MEDIUM	HIGH	VERY HIGH	TOTAL	HIGHEST DENSITY WITHIN	CV
area (km^2)	1	7	13	8	3	32		
Artifact Scatters in Fields (count)	1	9	7	3	2	22		
Artifact Scatters in Fields (percent)	5%	41%	32%	14%	9%	100%		0.78
Artifact Scatters in Fields per km^2	1.0000	1.2857	0.5385	0.3750	0.6667	0.6875	Low	
Permanent Sites (count)	0	0	5	3	1	9		
Permanent Sites (percent)	0%	0%	56%	33%	11%	100%		1.20
Permanent Sites per km^2	0.0000	0.0000	0.3846	0.3750	0.3333	0.2813	Medium	

Table 19 2019 field survey sites in relation to Model B, which is based on large settlement classes from the 1907 map.

Figure 9 Sites from systematic survey in 2019 in relation to Model A, which is based on small settlement size classes in the 1907 map.

Figure 10 Sites from systematic survey in 2019 in relation to Model B, which is based on large settlement size classes in the 1907 map.

3.3 ASSESSING FALSE POSITIVES

A consideration for both models is the extent to which zones of High and Very High probability for site detection return false positives. For known major precolonial sites this was not possible to assess, since these site locations come from disparate sources and were not the result of a survey sample. For the 2019 field survey sites, Figures 9 and 10 show survey transects across and within site detection zones for Model A and B, which give an indication of false positive results for each model. Though Model A's Very High site zone was a better predictor of permanent sites recovered during survey than Model B, it also created High and Very High site detection zones across three transects which did not produce any permanent sites, one of which did not produce any sites at all. Model B's High and Very High site detection zones were less successful at predicting all site locations, but the model also has fewer false positives in Very High site detection zones. Model B's Very High site detection zone was smaller, but still included the precolonial village site, and the largest site in the survey region, the ~60 ha dispersed village of Chaani. The false positives in Model A especially show the limitations of this model at the scale of transect survey, but do not detract from the larger regional implications of the study.

4. DISCUSSION

Model A was successful at predicting the locations of smaller, permanent village sites in rural inland Zanzibar, while Model B was more successful at predicting the locations of major precolonial Swahili sites, especially in coastal areas. This section considers the utility of these models for future site detection and reflects on the environmental affordances that might have conditioned Swahili settlement over time at different scales.

Model A can be used to predict other permanent precolonial and colonial period village sites in rural inland areas. Another precolonial village site, Mwanakombo, was also discovered in 2019 during field surveys but was not included in this analysis because it was not recorded during systematic surveys (Alders 2023); nevertheless, this site also falls within the High and Very High site detection zones of Model A. Precolonial village sites like these took advantage of *kinongo* soils for farming and making earth and thatch houses, proximity to streams, and high rainfall. On the other hand, the small ephemeral sites outside of the predictive zones in Model A attest to the creative forms of land use that Swahili communities have employed for centuries in environmentally marginal landscapes. Though Swahili people did not settlement permanently in these zones they nevertheless transformed and occupied these landscapes through seasonal incremental processes, digging in coralline limestone bedrock to plant and crafting field walls out of limestone cobbles (Alders 2022: 123–125). Although Swahili communities favored specific environmental zones for permanent settlement, they were not constrained

from using and moving through less favorable zones on the island. Ecological affordances structured, but did not determine, long-term land use in rural inland areas.

Model B is a better fit than A for the data for known locations of major precolonial sites. In addition to predicting the locations of Stone Town, Mkokotoni, Shangani and Fukuchani in the northwest, the model identifies small strips of coastline in the south of the island as areas of High and Very High probability for site detection, and these locations line up well with the large precolonial port of Unguja Ukuu and the precolonial town of Kizimkazi, which hosts the oldest mosque in East Africa (Kleppe 2001). Further surveys in the High and Very High zones of Model B would likely reveal other important precolonial sites on the island. Areas for future surveys might include the southeastern coast near Paje, the southwest coast across the bay from Unguja Ukuu, the western peninsula south of Zanzibar Stone Town, the northwest coast, and many inland areas north of Zanzibar Stone Town. The inland region north of this urban center in particular likely contains a number of precolonial village sites that would help clarify urban-rural interactions. These sites may be under threat of destruction from growing agricultural and urban development.

The predictive models produced here are useful tools for archaeological prospection, but they also inform a long-term understanding of urban and rural settlement development on the East African Swahili Coast over the last millennium. Research on the Swahili Coast has definitively revealed the scale, complexity, and interconnectedness of non-elite, rural settlement (Kusimba et al. 2013; LaViolette and Fleisher 2018). Increasingly, archaeologists have sought to investigate the environmental dynamics of Swahili settlement landscapes (Faulkner et al. 2022; Fitton et al. 2023; Kotarba-Morley et al. 2022; Pawlowicz et al. 2014; Prendergast et al. 2017; Quintana Morales et al. 2022; Walshaw and Stoetzel 2018). This paper contributes to this growing body of research by modeling and testing the environmental affordances that influenced regional settlement trends. The fact that, out of the large area of southern Zanzibar, Unguja Ukuu and Kizimkazi developed in small regions identified by Model B attests to how local environmental conditions were significant factors for the development of Swahili settlements. The settlement locations of even the largest and wealthiest Swahili towns on Zanzibar developed in places where early Swahili communities capitalized on environmentally suitable zones for farming, fishing, house building, and procuring water. These zones continued to influence settlement trends into the colonial era when Omani planters settled rural inland landscapes with enslaved retinues and sought to produce cloves and other products for international markets.

Environmental factors were important, but Swahili people on Zanzibar were not constrained by them. In the case of Tumbatu, they settled on a rocky, agriculturally marginal offshore island with little water, and may

have relied on social networks to provision the town from the more agriculturally suitable territories around Mkokotoni. Also, as demonstrated by field surveys, small-scale Swahili communities farmed and camped in environmentally marginal zones in the rocky eastern region, though they did not settle there permanently. Nevertheless, ecological factors certainly influenced Swahili settlement trends over time, and the predictive models produced here help contextualize the material affordances that Swahili people dealt with, mobilized, and capitalized on over the last millennium.

5. CONCLUSION

A comparison with ground-truthed archaeological sites shows the effectiveness of archaeological predictive modeling through zonal statistics on Zanzibar, Tanzania. The results may help plan future surveys and inform emergent understandings of human-environment dynamics on the Swahili Coast. The development of this methodology using open-access software increases geospatial accessibility and affordability, a consideration that will be especially impactful for researchers in the Global South where funding and licensing is limited. Like recent studies that emphasize low-cost open-access remote sensing methods for archaeological prospection in Africa, this method takes advantage of a growing suite of freely available geospatial datasets. The methodology described here can be applied across a wide variety of contexts in Africa and globally. This method does not rely on having high-resolution multispectral imagery, LiDAR, or paywalled software. The only prerequisite is having a representative and theoretically-justified way to weight zonal raster images—in this case, settlement classes from a digitized historical map were the basis for weighting. The quality and representativeness of training features across all zones under consideration is an important factor for producing a model that is useful for archaeological prospection and for understanding regional environmental and spatial factors.

ADDITIONAL FILE

The additional file for this article can be found as follows:

• **Supplementary Materials.** Supplementary files contain all zonal raster images used in the analysis, a digitized close-up of the 1907 map, and a detailed workflow in QGIS. DOI: https://doi.org/10.5334/jcaa.107.s1

ACKNOWLEDGEMENTS

FUNDING INFORMATION

COMPETING INTERESTS

The author has no competing interests to declare.

AUTHOR AFFILIATIONS

Wolfgang Alders orcid.org/0000-0002-8656-3765
University of Arkansas, US

REFERENCES

A map of Zanzibar Island from surveys made by Imam Sherif Khan Bahadur, Survey of India, under the direction of the late General Sir Lloyd William Mathews, KCMG. 1907. London: Stanford's Geographical Establishment. Retrieved from https://gallica.bnf.fr/ark:/12148/btv1b52505995f.r.

Adjustment of certain charges in connection with the deputation of Asstt: surveyor Imam Sharif, Khan Bahadur, to Zanzibar. 1894. Calcutta: National Archives of India. Retrieved from https://www.abhilekh-patal.in/jspui/handle/123456789/2894038

Alders, W. 2022. *Uneven Ground: The Archaeology of Social Transformation in Zanzibar, Tanzania.* Berkeley: Department

of Anthropology Dissertation, UC Berkeley. Retrieved from https://escholarship.org/uc/item/3n46k1w9.

Alders, W. 2023. The archaeology of social transformation in rural Zanzibar, Tanzania, eleventh to nineteenth centuries CE. *African Archaeological Review*. DOI: https://doi.org/10.1007/s10437-023-09523-y

Behr, D, Ozgul, A and **Cozzi, G.** 2017. Combining human acceptance and habitat suitability in a unified socio-ecological suitability model: a case study of the wolf in Switzerland. *Journal of Applied Ecology*, 54: 1919–1929. DOI: https://doi.org/10.1111/1365-2664.12880

Biagetti, S, Merlo, S, Adam, E, Lobo, A, Conesa, F, Knight, J, …Madella, M. 2017. High and medium resolution satellite imagery to evaluate late Holocene human–environment interactions in arid lands: A case study from the Central Sahara. *Remote Sensing*, 9(4). DOI: https://doi.org/10.3390/rs9040351

Casana, J. 2020. Global-scale archaeological prospection using CORONA satellite imagery: automated, crowd-sourced, and expert-led approaches. *Journal of Field Archaeology*, 45: 89–100. DOI: https://doi.org/10.1080/00934690.2020.1713285

Castiello, M. 2022. Predictive Modeling. In M. Castiello, *Computational and Machine Learning Tools for Archaeological Site Modeling* (pp. 33–63). Cham: Springer. DOI: https://doi.org/10.1007/978-3-030-88567-0_3

Castiello, M and **Tonini, M.** 2021. An Explorative Application of Random Forest Algorithm for Archaeological Predictive Modeling. A Swiss Case Study. *Journal of Computer Applications in Archaeology*, 4(1): 110–125. DOI: https://doi.org/10.5334/jcaa.71

Cerasoni, J, Hallett, E, Ben Arous, E, Beyer, R, Krapp, M, Manica, A and **Scerri, E.** 2022. Archaeological sites and palaeoenvironments of Pleistocene West Africa. *Journal of Maps*. DOI: https://doi.org/10.1080/17445647.2022.2052767

Colbert, G, Wagner, B and **Pinther, M.** 1987. *Hydrogeological map of Zanzibar*. New York: United Nations.

Comer, J, Comer, D, Dumitru, I, Priebe, C and **Patsolic, J.** 2023. Sampling methods for archaeological predictive modeling: spatial autocorrelation and model performance. *Journal of Archaeological Science: Reports*, 48. DOI: https://doi.org/10.1016/j.jasrep.2022.103824

Cooper, F. 1977. *Plantation slavery on the east coast of Africa*. New Haven: Yale University Press.

Crowther, A, Faulkner, P, Prendergast, M, Quintana Morales, E, Horton, M, Wilmsen, E, …Boivin, N. 2016. Coastal subsistence, maritime trade, and the colonization of small offshore islands in Eastern African prehistory. *Journal of Island and Coastal Archaeology*, 11(2): 211–237. DOI: https://doi.org/10.1080/15564894.2016.1188334

D'Andrea, A, Manzo, A, Harrower, M and **Hawkins, A.** 2008. The pre-Aksumite and Aksumite Settlement of NE Tigrai, Ethiopia. *Journal of Field Archaeology*, 33: 151–176. DOI: https://doi.org/10.1179/009346908791071268

Davis, D and **Douglass, K.** 2020. Aerial and spaceborne remote sensing in African archaeology: a review of current research and potential future avenues. *African Archaeological Review*, 37: 9–24. DOI: https://doi.org/10.1007/s10437-020-09373-y

Davis, D and **Douglass, K.** 2021. Remote sensing reveals lasting legacies of land use by small-scale foraging societies. *Frontiers in Ecology & Evolution*, 9. DOI: https://doi.org/10.3389/fevo.2021.689399

Davis, D and **Sanger, M.** 2021. Ethical challenges in the practice of remote sensing and geophysical archaeology. *Archaeological Prospection*, 28: 271–278. DOI: https://doi.org/10.1002/arp.1837

Diwan, G. 2020. GIS-based comparative archaeological predictive models: a first application to Iron Age sites in the Bekaa (Lebanon). *Mediterranean Archaeology and Archaeometry*, 20(2): 143–158.

Faulkner, P, Sarathi, A, Crowther, A, Smith, T, Harris, M, Ali, A, …Boivin, N. 2022. Human-ecodynamics and the intertidal zones of the Zanzibar archipelago. *Frontiers in Earth Science*. DOI: https://doi.org/10.3389/feart.2022.982694

Fisher, M, Fradley, M, Flohr, P, Rouhani, B and **Simi, F.** 2021. Ethical considerations for remote sensing and open data in relation to the endangered archaeology in the Middle East and North Africa project. *Archaeological Prospection*. DOI: https://doi.org/10.1002/arp.1816

Fitton, T. 2018. Zanzibar. In S. Wynne-Jones, & A. LaViolette, *The Swahili World* (pp. 239–244). London: Routledge. DOI: https://doi.org/10.4324/9781315691459-22

Fitton, T, Sulas, F, Lisowski, M, Alexander, M, Juma, A and **Wynne-Jones, S.** 2023. Integrating spatial and legacy data to understand archaeological sites in their landscape. A case study from Unguja Ukuu, Zanzibar. *Archaeological Prospection*. DOI: https://doi.org/10.1002/arp.1885

Fitzgerald, W. 1898. *Travels in the coastlands of British East Africa and the islands of Zanzibar and Pemba; their agricultural resources and general characteristics*. London: Chapman and Hall.

Fleisher, J and **LaViolette, A.** 1999. Elusive watte-and-daub: Finding the hidden majority in the archaeology of Swahili. *Azania: Archaeological Research in Africa*, 34: 87–108. DOI: https://doi.org/10.1080/00672709909511473

Garcia, A, Orengo, H, Conesa, F, Green, A and **Petrie, C.** 2019. Remote sensing and historical morphodynamics of alluvial plains: the 1909 Indus flood and the city of Dera Ghazi Khan (province of Punjab, Pakistan). *Geosciences*, 9(1). DOI: https://doi.org/10.3390/geosciences9010021

Green, A, Orengo, H, Alam, A, Garcia-Molsosa, A, Green, L, Conesa, F, …Petrie, C. 2019. Re-discovering ancient landscapes: Archaeological survey of mound features from historical maps in northwest India and implications for investigating the large-scale distribution of cultural heritage sites in South Asia. *Remote Sensing*, 11(18). DOI: https://doi.org/10.3390/rs11182089

Hardy, A, Mageni, Z, Dongus, S, Killeen, G, Macklin, M, Majambare, S, …Thomas, C. 2015. Mapping hotspots of malaria transmission from pre-existing hydrology, geology, and geomorphology data in the pre-elimination context of Zanzibar, United Republic of Tanzania. *Parasites and Vectors*, 8(41). DOI: https://doi.org/10.1186/s13071-015-0652-5

Harrower, M. 2010. Geographic Information Systems (GIS) hydrological modeling in archaeology: an example from the origins of irrigation in southwest Arabia (Yemen). *Journal of Archaeological Science*, 37(7): 1447–1452. DOI: https://doi.org/10.1016/j.jas.2010.01.004

Harrower, M, Nathan, S, Mazzariello, J, Zerue, K, Dumitru, I, Meresa, Y, … Anderson, M. 2020. Water, geography, and aksumite civilization: The southern red sea archaeological histories (SRSAH) project survey (2009–2016). *African Archaeological Review*, 37(1): 51–67. DOI: https://doi.org/10.1007/s10437-020-09369-8

Harrower, M, Oches, E and **McCorriston, J.** 2012. Hydro-geospatial analysis of ancient pastoral/agro-pastoral landscapes along Wadi Sana (Yemen). *Journal of Arid Environments*, 86: 131–138. DOI: https://doi.org/10.1016/j.jaridenv.2011.11.020

Horton, M and **Clark, C.** 1985. Zanzibar archaeological survey. *Azania*, 20: 161–171. DOI: https://doi.org/10.1080/00672708509511362

Horton, M and **Middleton, J.** 2000. *The Swahili: The Social Landscape of a Mercantile Society*. Wiley-Blackwell.

Hritz, C. 2010. Tracing settlement patterns and channel systems in southern Mesopotamia using remote sensing. *Journal of Field Archaeology*, 35: 184–203. DOI: https://doi.org/10.1179/009346910X12707321520477

Kelly, D, Clark, M, Palace, M and **Howey, M.** 2023. Expanding omnidirectional geospatial modeling for archaeology: A case study of dispersal in a "New England" colonial frontier (ca. 1600–1750). *Journal of Archaeological Science*, 150. DOI: https://doi.org/10.1016/j.jas.2022.105710

Khalaf, N and **Insoll, T.** 2019. Monitoring Islamic archaeological landscapes in Ethiopia using open source satellite imagery. *Journal of Field Archaeology*, 44: 401–419. DOI: https://doi.org/10.1080/00934690.2019.1629256

Khamis, Z, Kalliola, R and **Käyhkö, N.** 2017. Geographical characterization of the Zanzibar coastal zone and its management perspectives. *Ocean and Coastal Management*, 149: 116–134. DOI: https://doi.org/10.1016/j.ocecoaman.2017.10.003

Klehm, C, Barnes, A, Follett, F, Simon, K, Kiahtipes, C and **Mothulatshipi, S.** 2019. Toward archaeological predictive modeling in the Bosutswe region of Botswana: Utilizing multispectral satellite imagery to conceptualize ancient landscapes. *Journal of Anthropological Archaeology*, 54: 68–83. DOI: https://doi.org/10.1016/j.jaa.2019.02.002

Klehm, C and **Gokee, C.** 2020. Geospatial analysis in African archaeology: current theories, topics, and methods. *African Archaeological Review*, 37: 1–7. DOI: https://doi.org/10.1007/s10437-020-09376-9

Kleppe, EJ. 2001. Archaeological investigations at Kizimkazi Dimbani. In Scarcia Amoretti, B (ed.), *Islam in East Africa: new sources; (archives, manuscripts and written historical sources, oral history, archaeology)* (pp. 361–384). Herder.

Kotarba-Morley, A, Kourampas, N, Morley, M, MacAdams, C, Crowther, A, Faulkner, P, … Boivin, N. 2022. Coastal landscape changes at Unguja Ukuu, Zanzibar: Contextualizing the archaeology of an early Islamic port of trade. *Journal of Island and Coastal Archaeology*. DOI: https://doi.org/10.1080/15564894.2022.2030441

Kuria, D, Ngari, D and **Waithaka, E.** 2011. Using geographic information systems (GIS) to determine land suitability for rice crop growing in the Tana delta. *Journal of Geography and Regional Planning*, 4(9): 525–532.

Kusimba, C, Kusimba, SB and **Dussubieux, L.** 2013. Beyond the coastalscapes: preindustrial social and political networks in East Africa. *African Archaeological Review*, 30: 399–426. DOI: https://doi.org/10.1007/s10437-013-9133-4

LaViolette, A and **Fleisher, J.** 2018. Developments in rural life on the eastern African coast, A.D. 700–1500. *Journal of Field Archaeology*, 43(5): 380–398. DOI: https://doi.org/10.1080/00934690.2018.1489661

LaViolette, A, Fleisher, J and **Horton, M.** 2023. Assembling Islamic practice in a Swahili urban landscape, 11th–16th centuries. *Journal of Social Archaeology*, 23(3). DOI: https://doi.org/10.1177/14696053231151667

LaViolette, A and **Norman, N.** 2023. The archaeology of Portuguese agricultural outposts in the 17th century Zanzibar countryside. In Haines, J and Hauser, M (eds.), *The Archaeology of Modern Worlds in the Indian Ocean*. Gainesville: University Press of Florida.

Magnini, L and **Bettineschi, C.** 2021. Object-based predictive modeling (OBPM) for archaeology: finding control places in mountainous environments. *Remote Sensing*, 13. DOI: https://doi.org/10.3390/rs13061197

Mansab, M. 2021. Mariam Mansab, Director of Zanzibar's Museum and Antiquities, in conversation with Dr. Akinbowale Akintayo, University of York, Department of Archaeology. (A. Akintayo, Interviewer) Retrieved from https://maeasam.org/heritage-management-in-zanzibar/.

Mugnier, C. 2021. Grids & datums. *Photogrammetric Engineering & Remote Sensing*, 87: 399–403. DOI: https://doi.org/10.14358/PERS.87.6.399

Ochungo, P, Khalaf, N, Merlo, S, Beldados, A, M'Mbogori, F, Tiki, W and **Lane, P.** 2022. Remote Sensing for Biocultural Heritage Preservation in an African Semi-Arid Region: A Case Study of Indigenous Wells in Northern Kenya and Southern Ethiopia. *Remote Sensing*, 14. DOI: https://doi.org/10.3390/rs14020314

Owens, R. 2007. Exploring the Articulation of Governmentality and Sovereignty: The Chwaka Road and the Bombardment of Zanzibar, 1895–1896. *Journal of Colonialism and Colonial History*, 8(2). DOI: https://doi.org/10.1353/cch.2007.0036

Pawlowicz, M, Fleisher, J and **De Luna, K.** 2020. Capturing people on the move: spatial analysis and remote sensing in the Bantu Mobility Project, Basanga, Zambia. *African Archaeological Review*, 37: 1–25. DOI: https://doi.org/10.1007/s10437-020-09363-0

Pawlowicz, M, Stoetzel, J and **Macko, S.** 2014. Environmental archaeology at Mikindani, Tanzania: towards a historical ecology of the southern Swahili coast. *Journal of African Archaeology*, 12(2): 119–139. DOI: https://doi.org/10.3213/2191-5784-10260

Petrie, C, Orengo, H, Green, A, Walker, J, Garcia, A, Conesa, F, … Singh, R. 2019. Mapping archaeology while mapping

an empire: Using historical maps to reconstruct ancient settlement landscapes in modern India and Pakistan. *Geosciences*, 9(1). DOI: https://doi.org/10.3390/geosciences9010011

Prendergast, M, Quintana Morales, E, Crowther, A, Horton, M and **Boivin, N.** 2017. Dietary diversity on the Swahili coast: the fauna from two Zanzibar trading locales. *International Journal of Osteoarchaeology*, 27: 621–637. DOI: https://doi.org/10.1002/oa.2585

Quintana Morales, E, Craig, O, Prendergast, M, Walshaw, S, Cartaciano, C, Mwebi, O … Wynne-Jones, S. 2022. Diet, economy, and culinary practices at the height of precolonial Swahili urbanism. *Journal of Anthropological Archaeology*, 66: 101406. DOI: https://doi.org/10.1016/j.jaa.2022.101406

Rayne, L, Carmela Gatto, M, Abdulaati, L, Al-Haddad, M, Sterry, M, Sheldrick, N and **Mattingly, D.** 2020. Detecting change at archaeological sites in North Africa using open-source satellite imagery. *Remote Sensing*, 12(22). DOI: https://doi.org/10.3390/rs12223694

Reid, S. 2016. Satellite remote sensing of archaeological vegetation signatures in coastal West Africa. *The African Archaeological Review*, 33: 163–182. DOI: https://doi.org/10.1007/s10437-016-9222-2

Reid, S. 2020. Forests of history: satellite remote sensing and archaeological survey in southern Ghana. *African Archaeological Review*, 37: 597–614. DOI: https://doi.org/10.1007/s10437-020-09393-8

Rhodes, D, Breen, C and **Forsythe, W.** 2015. Zanzibar: a nineteenth-century landscape of the Omani elite. *International Journal of Historical Archaeology*, 19(2): 334–355. DOI: https://doi.org/10.1007/s10761-015-0291-8

Rødland, H. 2021. *Swahili Social Landscapes: Material Expressions of Identity, Agency, and Labour in Zanzibar, 1000–1400 CE*. Uppsala: Dissertation, Department of Archaeology and Ancient History, University of Uppsala.

Ruete, E. 1888. *Memoirs of an Arabian Princess from Zanzibar*. New York: Appleton.

Sadr, K. 2016. A comparison of accuracy and precision in remote sensing stone-walled structures with Google earth, high resolution aerial photography and LiDAR; a case study from the South African iron age. *Archaeological Prospection*, 23: 95–104. DOI: https://doi.org/10.1002/arp.1532

Sheriff, A, Teelock, V, Omar Wahab, S and **Peerthum, S.** 2016. *Transition from Slavery in Zanzibar and Mauritius: A Comparative History*. Dakar, Senegal: CODESRIA. DOI: https://doi.org/10.2307/j.ctvh8qxjz

Shipton, C, Crowther, A, Kourampas, N, Prendergast, M, Horton, M, Douka, K, … Petraglia, M. 2016. Reinvestigation of Kuumbi Cave, Zanzibar, reveals Later Stone Age coastal habitation, early Holocene abandonment and Iron Age reoccupation. *Azania: Archaeological Research in Africa*, 51(2): 197–233. DOI: https://doi.org/10.1080/0067270X.2016.1173308

Thabeng, O, Merlo, S and **Adam, E.** 2020. From the bottom up: assessing the spectral ability of common multispectral sensors to detect surface archaeological deposits using field spectrometry and advanced classifiers in the Shashe-Limpopo confluence area. *African Archaeological Review*, 37: 25–49. DOI: https://doi.org/10.1007/s10437-020-09372-z

Verhagen, P and **Whitley, T.** 2020. Predictive Spatial Modeling. In Gillings, M, Hacıgüzeller, P and Lock, G (eds.), *Archaeological Spatial Analysis*. London: Routledge. DOI: https://doi.org/10.4324/9781351243858-13

Walshaw, S. 2015. *Swahili Trade, Urbanization, and Food Production: Botanical Perspectives from Pemba Island, Tanzania*. Oxford: Cambridge Monographs in African Archaeology 90 (British Archaeological Reports series 2755, Archaeopress). DOI: https://doi.org/10.30861/9781407314167

Walshaw, S and **Stoetzel, J.** 2018. Plant use and the creation of anthropogenic landscapes: coastal forestry and farming. In Wynne-Jones, S and LaViolette, A (eds.), *The Swahili World* (pp. 350–362). New York: Routledge. DOI: https://doi.org/10.4324/9781315691459-33

Yaworsky, P, Vernon, K, Spangler, J, Brewer, S and **Codding, B.** 2020. Advancing predictive modeling in archaeology: An evaluation of regression and machine learning methods on the Grand Staircase-Escalante National Monument. *Plos One*. DOI: https://doi.org/10.1371/journal.pone.0239424