A method for defining dispersed community territories

Kenneth B. Vernon[1] and Scott G. Ortman[1]

[1] *Center for Collaborative Synthesis in Archaeology, Institute of Behavioral Science, University of Colorado, Boulder, 1440 15th Street, Boulder, CO 80302*

Draft compiled on 2024-06-14

Abstract

The transition from dispersed to aggregated forms of settlement reflects a critical shift in the relative value of social and primary (food) modes of production. However, investigating trade-offs between these different forms of settlement requires estimates of the extent of community territories, including their nearby arable land. Here we demonstrate a simple algorithm to do that. Our algorithm is analogous to that used to define core-based statistical areas for the US census, though instead of central business districts, we rely on community centers (or areas of known and persistent interaction between unrelated individuals). We provide examples of our algorithm by applying it to archaeological sites in the central Mesa Verde, northern Rio Grande, and Cibola regions in the US Southwest. A sensitivity analysis is also conducted to demonstrate how each tuning parameter contributes to the algorithm.

Keywords: Aggregation, Core-Based Statistical Area, Community Detection, Complex Systems, Least-Cost Path Analysis, Population Density

Target Journal(s): Journal of Archaeological Science

Corresponding author email: *Kenneth.Vernon@colorado.edu*

1 Introduction

Urban geographers recognize density as a crucial property of human communities which influences many other properties, including productivity, resource needs, and infectious disease rates (Angel et al., 2021; Bettencourt, 2021; Duranton & Puga, 2020). The archaeological record of subsistence farming societies contains physical traces of community formations with a wide range of densities, often glossed as "dispersed" or "aggregated" (Birch, 2013; Drennan et al., 2015; Gyucha, 2019). The primary difference between the two has to do with the spatial distribution of residents relative to areas of primary (food) production. In dispersed formations, households are scattered across an area, interspersed with the land that they farmed, thus requiring them to commute to central places for various forms of social production. In aggregated formations, in contrast, households are clustered in a village or town, such that they reside in locations of social production, but must commute to fields for primary production.

From a complex systems perspective, the main way these two settlement morphologies vary is with respect to which costs are being minimized. In the dispersed pattern, costs associated with primary (food) production by the household are being minimized, whereas in the aggregated pattern, it is costs associated with various forms of social production, from government and ritual to economic exchange and warfare. Presumably, choices regarding which sorts of costs to minimize are driven by the relative productivity of primary and social production, which is to say, their relative importance, broadly construed, for human well-being in a given context.

There are many episodes in history where human settlements transitioned from dispersed to aggregated, and presumably this is a signal of a change in the relative value of social vs. primary production for the residents. This can occur for a variety of reasons, including an increase in the social cost of not being aggregated (e.g., due to warfare); a decrease in transport costs for staples (due to animal traction, wheeled vehicles, roads, etc.); an increase in land productivity which changes the balance of costs of transport for primary vs. social production; or an increase in the contribution of exchange to household incomes, which can derive from increases in community size and density. In addition, while one can conceptualize transitions between dispersed and aggregated forms of settlement in general systems terms, in real human communities a variety of political or ideological factors can keep people from adopting an energetically balanced form of settlement given prevailing conditions. So, there are opportunities for archaeologists to study deviations from general equilibrium conditions in addition to factors that shift the equilibrium in a given context.

For all this work, however, there is a fundamental issue: the identification of community territories. It is relatively straightforward to define the boundaries of aggregated settlements by finding the extent of built space and/or artifactual remains. However, to define the community boundary one needs to know the extent of agricultural land used by the residents, and this is often difficult to determine from the village remains themselves (Varien, 1999b). In contrast, it is often quite difficult to determine the boundaries of dispersed communities due to gaps in survey coverage or relatively consistent distributions of farmsteads, but once one has determined the community territory it is straightforward to define the associated agricultural land.

In this paper, we present a method we developed to solve such problems, which enables us to compare the properties and resilience of dispersed vs. aggregated communities on the same landscapes. We draw on three extensive compilations of archaeological data from different portions of the ancestral Pueblo region of the US Southwest, using the locations of what we call community centers as tethers for defining territories that we infer were used by social communities over extended periods of time (Glowacki & Ortman, 2012; Varien et al., 2007). In some regions, community centers are sites containing civic architecture, or more households than could plausibly be related by reckoned kinship. The key feature of these centers is that they represent locations against which travel time distances to other sites can be compared.

While not strictly required by our algorithm, we make the simplifying assumption that once established, community territories were relatively fixed in space for the duration of the occupations of their associated sites. This allows us to use the distribution of all recorded residences, regardless of their periods of occupation, in defining community territories. There are several reasons why we believe this is a reasonable simplification. First, there are strong cross-cultural regularities in the time individuals spend in daily travel from and to their residences, so dispersed farming communities tend to be dispersed across the distance associated with this typical travel time, regardless of the number of households involved (Marchetti, 1994). Second, the archaeological record of our study areas supports the idea that community territories did not shrink when their populations aggregated into villages. In some cases, farmsteads were converted to field houses, with some building materials being reused elsewhere and evidence for continued limited activity use (Varien, 1999b, 1999a, 2002). In other cases, grid gardens were constructed directly within older rubble mounds by residents of newer, aggregated villages nearby (Gauthier & Herhahn, 2005). Third, an analysis that allowed community territories to vary by time step would reduce variation in the density of the resulting units, thus washing out variation in one of the most important social properties one would ideally want to examine. Fourth, in areas where there is strong survey coverage, dispersed residences form clusters, often but not always centered on a few larger settlements, and this clustering is apparent even when the dispersed residences are plotted by time step (Schachner, 2012; Varien, 2002). Fifth, the effects of population growth for dispersed farming communities are not parallel to the effects of population expansion in aggregated settlements. In the latter, the area of the settlement must grow somewhat because the residential density is already relatively high. But in dispersed settlements, there is often unused land between existing farmsteads that can be settled. This, combined with the strong cross-cultural regularity in typical commute times, leads to much greater consistency in the spatial extent of residents in a dispersed community than in the case of an aggregated community. Finally, given the inherent incompleteness of the archaeological record, it seems most appropriate to use all remains to define what is effectively the maximum extent of community territories. While it was undoubtedly true that individual community territories expanded or contracted over time, the assumption of constant area seems more reasonable than the assumption of constant density.

2 Data

The datasets we apply our algorithm to come from three regions in the US Southwest: Central Mesa Verde (CMV), the Northern Rio Grande (NRG), and the Cibola or Zuni region (CIB) (see Figure 1). The CMV and NRG datasets derive from Phase II of the Village Ecodynamics Project (Ortman, 2016b; Schwindt et al., 2016). The CIB data derive from data syntheses produced by Schachner and Peeples (Peeples, 2018; Peeples & Schachner, 2012; Schachner, 2012). All of these data are now included in cyberSW (Mills et al., 2020). The distribution of farms and centers is shown in Figure 2. For summaries of the archaeological and environmental context in each region, see Table 1.

The CMV area encompasses the southwestern corner of Colorado, from Mesa Verde National Park in the southeast to the Dolores River in the north, and the Utah state line in the west. The elevation ranges from roughly 1,399.8 m to 3,032.8 m, with an average elevation of 1,999.0 m (σ = 276.5 m). SKOPE (Bocinsky et al., 2022) estimates suggest annual precipitation in the CMV study area has averaged around 39.5 cm (σ = 9.1 cm) over the last 2,000 years, and it puts summer growing degree days for maize around 2,390.1 °F

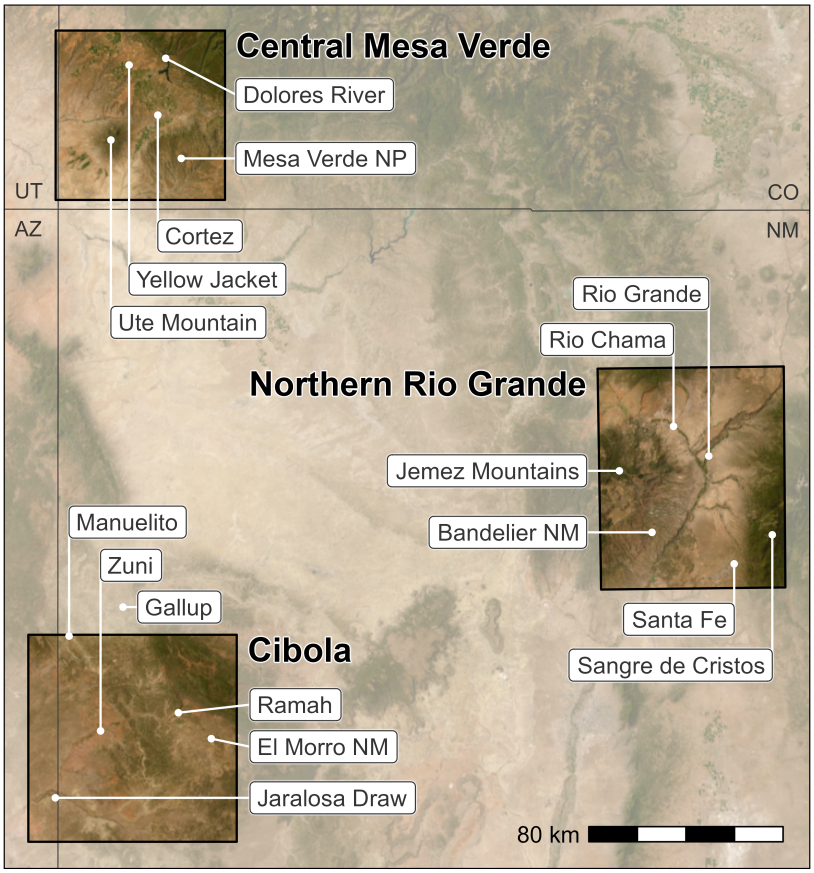

Figure 1. Overview map showing the locations of the three study areas: Central Mesa Verde, Northern Rio Grande, and Cibola. Prominent towns, administrative units, and landforms are also labeled to help orient the reader.

(σ = 97.9 °F). Across most areas in the region, it is believed that direct precipitation was sufficiently high to support dry farming in most years (Bocinsky et al., 2016; Bocinsky & Kohler, 2014).

The NRG study area in north central New Mexico includes basically all of the Rio Grande valley between the Jemez Mountains to the west and the southernmost mountains of the Sangre De Cristos to the east, with the Rio Grande itself running from the northeast down to the western edge of the valley at the foot of the Pajarito Plateau around Bandelier National Monument. The elevation ranges from roughly 1,587.9 m to 3,836.2 m, with an average elevation of 2,220.0 m (σ = 371.3 m). SKOPE estimates the average annual precipitation in the NRG over the last 2,000 years to be around 40.6 cm (σ = 8.2 cm), and it puts summer growing degree days for maize around 2,154.4 °F (σ = 102.1 °F). It is argued that direct rainfall at Bandelier NM made dry farming possible there, but in the valley farming required floodwater irrigation along watercourses (Bocinsky et al., 2016; Bocinsky & Kohler, 2014; Duwe & Anschuetz, 2013; Ortman, 2016a).

The CIB region straddles the southern slope of the Colorado Plateau along the border between Arizona and New Mexico, from El Morro National Monument in the east to the centrally located Zuni Pueblo to the confluence of Jaralosa Draw and the Zuni River in the southwest. Although this is a rugged terrain by any reasonable standard, it is notably less so than the other study areas. The elevation ranges from roughly 1,771.3 m to 2,777.1 m, with an average elevation of 2,145.9 m (σ = 160.0 m). SKOPE estimates suggest annual precipitation in the CIB area over the last 2,000 years has averaged around 34.3 cm (σ = 7.6 cm), and it puts summer growing degree days for maize around 2,317.3 °F (σ = 93.8 °F). However, no area in the region receives sufficient direct rainfall to support dry farming, so some form of water management was required (Kintigh, 1985; Muenchrath et al., 2002).

Table 1. Regional Summaries

	Area [km2]	**Dates [CE]**	**Archaeology**	**Environment**
CMV	4,566.4	725-1280	Rooms: 68,940 Farms: 5,272 Centers: 170	Precipitation: 39.5 cm (σ: 9.1) Maize GDD: 2,390.1 °F (σ: 97.9) Elevation: 1,999.0 m (σ: 276.5)
NRG	6,958.0	900-1550	Rooms: 46,226 Farms: 2,316 Centers: 181	Precipitation: 40.6 cm (σ: 8.2) Maize GDD: 2,154.4 °F (σ: 102.1) Elevation: 2,222.0 m (σ: 371.3)
CIB	7,424.1	700-1540	Rooms: 26,818 Farms: 713 Centers: 77	Precipitation: 34.3 cm (σ: 7.6) Maize GDD: 2,317.3 °F (σ: 93.8) Elevation: 2,145.9 m (σ: 160.0)

3 Methods

Our clustering algorithm draws inspiration from the procedure used by the United States Office of Management and Budget to define Core Based Statistical Areas (CBSA) (2020 Standards for Delineating Core Based Statistical Areas, 2021). We start by defining a core area, in this case a community center location. We then associate outlying areas with their nearest core area and merge core areas into larger agglomerations based on the proportion of their populations that they share. Our method may be described as a guided density-based clustering algorithm in that it does not rely on a random set of points when initialized. Instead, the selection of core areas is guided by archaeological data and regional expertise. We think this is a key argument in its favor.

From an archaeological perspective, one can think of our algorithm as combining insights from Varien (1999b) and Reese et al. (2019), specifically Varien's idea of a community catchment, which is an isochrone defined by a uniform commute time in all directions from a community center, and Reese's suggestion that a spatial or geographic community can be identified by grouping individuals in terms of the commute distances between them. Our way of combining these ideas is to associate farm sites with specific community centers in terms of their commute distances and then draw the community catchment around those associated farms. We also share with them and others (Lipe & Hegmon, 1989; Murdock, 1949; Peterson & Drennan, 2005) a focus on "geographic" as opposed to "purely social" communities; that is, communities constrained in space and comprised of individuals that can be expected to interact on a regular, even daily, basis.

An important difference for our algorithm is that it incorporates information about the total mass or population size around each center, so it is at least conceptually similar to the classic Xtent model proposed by (Renfrew & Level, 1979; see also Ducke & Kroefges, 2008), as well as the more widely known gravity model (Isard, 1954). In broad outline, it has the following steps:

1. Identify community centers
2. Join farms to their nearest community center
3. Exclude farms beyond a commute time threshold from all centers (*D-max*)
4. Join centers based on their overlapping populations (P, *D-join*)
5. Draw smallest concave hull encompassing all farms, centers, and paths

After step 4, we also apply a filter to remove communities that have less than a minimal number of dispersed farmsteads, specifically four. Although this is a somewhat arbitrary threshold, we note that three vertices are the bare minimum required to define a polygon and measure its area. By removing small communities, we also minimize potential underestimates of total arable land area. In addition, our larger ambition is to understand processes occurring within dispersed farming communities, not necessarily to define all the dispersed communities that once existed in an area. For community centers with few or no farms, there is no meaningful dispersed farming community to examine, regardless of whether this is due to limited survey coverage or actual past behavior, hence our desire to exclude them.

The primary tuning parameters in our algorithm are P, the proportion of a community population used to join communities; *D-join*, the commute time required to calculate P; and

D-max, the commute time used to exclude distant farms. While values for these variables were carefully chosen based on theory and empirical research, we recognize that they are, at the end of the day, arbitrary selections, so we include some limited sensitivity analysis to show how changes to these parameters affect the algorithm (see Figure 7). This involves changing the value of a focal parameter while holding the other parameters at their default values, which we describe below.

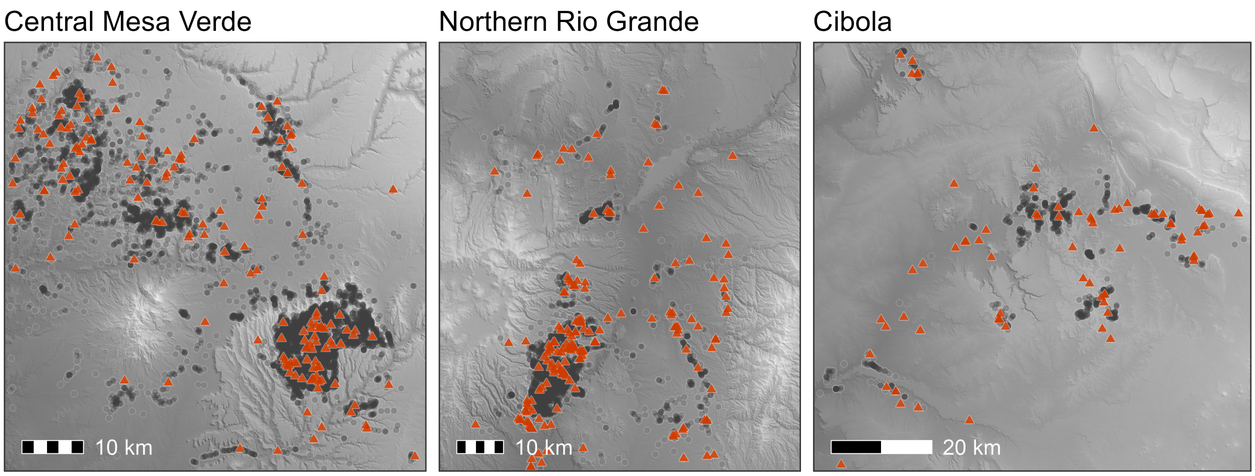

Figure 2. Overview map showing the locations of farmsteads (black dots) and community centers (red triangles) in Central Mesa Verde, Northern Rio Grande, and Cibola.

3.1 Identify community centers

Rule: Any site, whether residential or otherwise, with known and persistent interaction between unrelated individuals should count as a community center.

Why persistent interactions? Intuitively, a community center is a place where individuals from different households cross paths with the intention to interact (Berrey et al., 2021; Glowacki & Ortman, 2012; Varien, 1999b). A simple way to demonstrate that intention is to show that individuals would still be willing to pay some non-negligible cost to visit those centers under various counterfactual changes to local conditions. This serves to distinguish them from merely accidental exchanges or chance encounters. Of course, there can be some path-dependence to this process, with locations of ephemeral or one-off encounters coming to be locations of persistent interaction. The point is that the location will come to be a community center so long as individuals continue to visit that location despite the underlying variability in their circumstances.

Why unrelated individuals? The simple answer is that a community is not a family. As a unit of social organization, it does share something in common with a family, namely, a set of shared interests; but unlike in a family, those shared interests form only a small part of a much larger set of potentially conflicting goals and desires. As a consequence, communities are more susceptible to problems of coordination or collective action (Smith, 2003; Smith, 2010). On the other hand, the mere existence of a community center implies some level of coordination - even if that coordination is just about the location of the community center itself! So, community members share less in common than families, but more in common than a purely random sample of the larger population.

We can summarize these ideas by noting that persistent interactions at a location make that location a *center*, and persistent interactions between unrelated individuals at that location make it a *community* center (Gilpin, 2003; Peterson & Drennan, 2005; Stone, 2016; Varien & Potter, 2008). To make these considerations more useful for defining community boundaries, however, we need to operationalize them, to make them applicable to specific archaeological contexts, and that requires a deeper understanding of the archaeology in those contexts. For given differences in population density and settlement patterning, the frequency and intensity of interaction between individuals at centers is likely to vary by region, so it is reasonable to expect that community centers themselves (at least in so far as they play the role of centers) should vary in their size and composition across regions, too.

For the two regions represented by the VEP II dataset, we rely on two different but related sets of criteria to identify community centers. In the Central Mesa Verde area, a site is considered a community center if it includes civic architecture, specifically a great kiva or great house, or eight or more pit structures (small kivas) representing eight or more households. In the Northern Rio Grande, a site counts as a community center if it is classified as a village (a site with 50 or more rooms) or town (a site with 500 or more rooms). In the Cibola region, a community center includes at least 50 rooms or civic architecture. The large number of households in these regions are thought to be too large to be plausibly related by reckoned kinship.

3.2 Join farms to their nearest community center

Rule: For each farm, assign it to the community of its nearest community center.

Following Reese et al. (2019), we define nearness or proximity in terms of commute time rather than linear geographic distance. This is done for the obvious reason that the primary opportunity cost associated with daily pedestrian movement is the time spent walking, during which other productive activities cannot take place. There is also a much stronger cross-cultural regularity in the length of time individuals devote to commuting than the distance covered, as the latter is a function of the speed of movement, which is influenced by transport technology (Marchetti, 1994). As a result, both the frequency and intensity of interaction at community centers should decay with commute time. The longer it takes to get to a community center, the less frequent and less intense the interactions at that center should be (Peterson & Drennan, 2005; Varien, 1999b).

While different constraints will naturally arise for different modes of transportation, leading to differences in both the routes taken and the speed and distance covered, for the populations considered here the only available mode of transportation was pedestrian. The biggest obstacle to interaction was thus the landscape itself, in particular its topography - intuitively, the steeper the uphill slope, the slower the hiking speed, the longer the commute time. So, all else being equal, individuals at a farm that is equidistant (in terms of geographic distance) between two community centers should prefer traveling to the one separated by less rugged terrain.

Commute time is a critical variable in many places in our algorithm, but as a first approximation, we incorporate the idea by simply joining farms to the community of their nearest center, meaning the center with the shortest commute time from the farm. For details of how we estimate commute times over costly terrains, see "Least-cost path analysis," below.

3.3 Exclude distant farms

Rule: Exclude all farms farther than commute time, *D-max*, from their nearest center.

The commute time threshold, *D-max*, defines an isochrone around each community center. All farm sites that fall outside that isochrone are considered outliers and dropped from the analysis. Doing this helps to minimize potential overestimates of total arable land area for each community. This is owing to the fact that distant farms on average add more area to the community polygon than nearer farms. Here, we have defined *D-max* as one hour. This is about twice the median daily commute time observed in cross-cultural studies and it includes roughly 95% of the farmsteads in each region (see Figure 5).

3.4 Join centers with overlapping neighbors

Rule: For any two centers c and d with populations $N_c < N_d$, if $P \cdot N_c$ is within distance *D-join* of d, then c is part of the same community as d.

The larger and more dense farming populations become, the harder it gets to tell them apart and, more importantly, the harder it gets to justify keeping them apart, so we need to articulate a rule that specifies when two communities should be merged into one. For CBSAs, that decision hinges on a distinction between central and outlying counties, or counties where some fraction of the population live in an urban core and counties where some fraction commute to and from a central county on a regular basis. Basically, if the central counties of one CBSA qualify as the outlying counties of another CBSA, then the two CBSAs are merged. The intuition here is that when rates of interaction and exchange between CBSAs are roughly equal to rates within a CBSA, the CBSAs should be merged (2020 Standards for Delineating Core Based Statistical Areas, 2021).

We carry that intuition over to our merger rule for dispersed maize farming communities, albeit with some modification to account for the fact that we are working with point locations rather than administrative boundaries. We assume that if some proportion, P, of the population related to one community center lives within a certain commute time, *D-join*, of another community center, that those two community centers are effectively parts of the same dispersed community. We do this in the direction of letting larger communities absorb smaller ones.

Here we operationalize the population size N of center c using the well-established rule of thumb in US Southwest archaeology that each surface room in a settlement represents a single person, leading to a one to one conversion from rooms to residents (Duwe et al., 2016; Kintigh, 1985; Lekson, 1989; Lipe, 1989; Ortman, 2016b). In this case, we use the catchment room count, meaning the number of rooms within *D-join* of c, defined as:

$$N_c = \sum_{i=1}^{S} R_i \cdot I(t_{ic} \leq D\text{-}join)$$

for all sites S (including both farms and centers), with R_i being the room count at site i, t_{ic} the travel time from i to c, *D-join* the threshold travel time to a center, and I the indicator function that is 1 if $t \leq D\text{-}join$ and 0 otherwise. Note that S includes c and that $t_{cc} = 0$ (the travel time from a center to the same center is precisely 0), so R_c is always included in N_c.

The catchment room count within *D-join* of two centers c and d is then

$$N_{cd} = \sum_{i=1}^{S} R_i \cdot I(t_{ic} \leq D\text{-}join) \cdot I(t_{id} \leq D\text{-}join)$$

and

$$P_{cd} = N_{cd}/N_c$$

so we could also state our join rule as: if $P_{cd} \geq \alpha$ for some critical threshold α, then c is part of the same community as d. This promotes joins when c and d are themselves in close proximity to each other, as the farther apart they are, the smaller the number of sites that overlap within *D-join*. It also tends to encourage joining two communities when a substantial proportion of the population lives between the two centers. In either case, the result is what Varien (1999b) sometimes refers to as a "macro-community" or "multi-community cluster," meaning a community organized around multiple community centers.

The rule was partly inspired by the density-based clustering algorithm known as DBSCAN (Ester et al., 1996), as we rely on a moving window (an isochrone, in this case) set around each center, and look for some measure of density to guide join decisions. As noted above, however, our method does not rely on a random initialization of points, but rather targets community centers. It is also hierarchical in the sense that it exploits additional structure in our data, namely the distinction between community centers and farm sites.

An important obstacle to estimating population is that room counts cannot always be directly measured. For many sites in the CMV, pit structures or small kivas are more reliably identified than surface rooms. Fortunately, each pit structure is typically the central roofed space of a unit pueblo residence that includes an average of six surface rooms (Adler, 1990), so we simply multiply the number of pit structures by six to obtain room estimates for each site. CIB and NRG sites all have room estimates, so no additional assumptions are made about them.

For this analysis, we specify *D-join* to be one half hour and P to be 0.8. We acknowledge that these values are somewhat arbitrary. In their defense, they set an extremely high bar, so joining centers that satisfy the rule seems reasonable.

3.5 Draw community boundary

Rule: For each community, draw a polygon that encompasses all farms, community centers, and commute paths using a concave hull.

So far, we have only grouped site points into clusters or communities. The goal, however, is to define their spatial extent. That means, for each set of site points, we need some way of drawing a polygon that encompasses all of them. This can be done in a number of ways. The simplest strategy would be to find the centroid of the set of points and draw a circle with the smallest radius that still includes them all. This would be utterly arbitrary, however, and frankly unrealistic, not representing any sort of meaningful boundary given the landscape. An alternative would be to use the convex hull of the set of points, but this tends to exaggerate the total area of the community, especially when the distribution of points is suggestive of a concave shape, like the letter ‘C’. A concave hull would handle peculiar shapes, but it could also introduce the opposite problem, restricting the area of the community to an unreasonable degree, depending on how strict the convexity ratio is. Our somewhat brute force way of balancing these factors is to identify paths with shortest commute times between the outer most farms in a community, extract the vertices from those paths, and then incorporate those into the set of points used to define the concave hull (see Figure 3). The result is a polygon that is concave as well in the dimension of time, meaning the paths with shortest commute times between any two farms never leave the community.

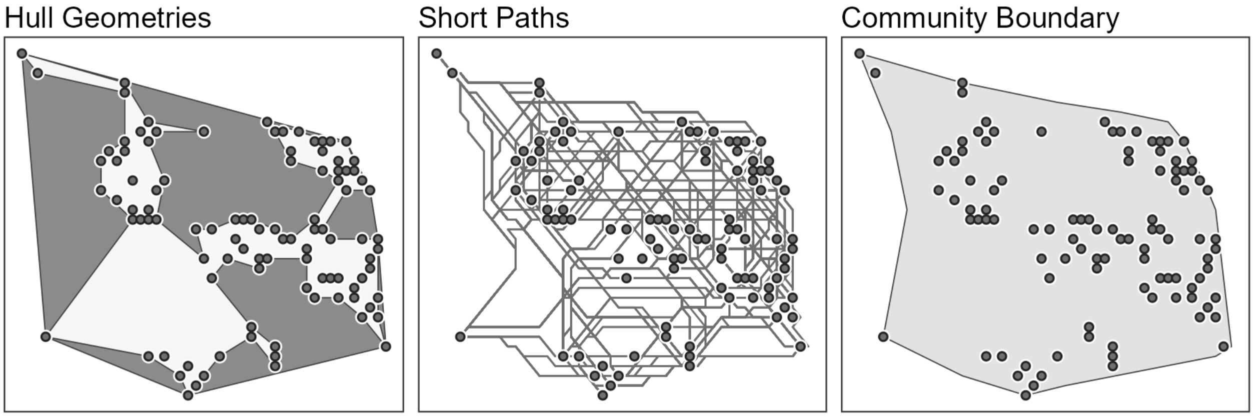

Figure 3. Example of the method used to define community boundaries. Left panel shows two hull geometries: (i) convex hull with dark gray fill and (ii) concave hull (ratio = 0.1) with light gray fill. Middle panel shows every pairwise short path between the points along the outer edge of the convex hull. Right panel shows community boundary derived by applying the concave hull (ratio = 0.5) to the combined set of site points and vertices of short paths. The most notable differences are along the left boundary.

In some cases, this may lead to overlapping community boundaries, though the amount of overlap should be minimal given the way that we cluster points, particularly when there are many densely distributed points. To some extent, of course, this sort of overlap is both unavoidable and expected as even in the modern world there are disagreements about the exact locations of legal and political boundaries.

As a final gloss on community boundaries, we also dilate each concave hull, briefly expanding it with a positive buffer, then shrinking it with a negative buffer having a slightly smaller size. This should have only a minimal effect on the total area of each polygon and also serves to smooth out noise along the edges.

3.6 Least-cost path analysis

The key to our path analysis is to take information about the terrain in our study regions and convert it first into meaningful estimates of hiking speed and then into meaningful estimates of travel time. Fortunately, attempts to model hiking speed across a range of slopes have advanced considerably in recent years, most notably in (Campbell et al., 2022; see also

Campbell et al., 2019). Our application of Campbell's hiking function relies on coefficients from the median range of individuals in their sample and assumes that everyone in the areas of our analysis was largely equal in their walking speed and endurance. This makes their walking speeds comparable to, but not the same as, what would be estimated by Tobler's hiking function (Tobler, 1993).

To calculate travel or commute times, we first download a 1 Arc-second digital elevation model (DEM) for each study area from the US Geological Survey (2023) 3D Elevation Program and convert it into a graph with each node representing a grid cell and each edge representing a straight line between adjacent grid cells in the Moore neighborhood (the eight adjacent grid cells). The degree slope of each edge is calculated, and edges with slope estimates greater than or equal to 45 degrees are removed. This prevents the path analysis from assuming that a hypothetical hiker would, for instance, scale steep slopes or walk off a cliff. The remaining slope estimates are then fed to Campbell's hiking function to derive estimates of hiking speed along each edge. The inverse of the hiking speed (the pace) is then multiplied by the distance along that slope to estimate travel time.

Next, we associate each farm and community center with the grid cell that contains its centroid. In effect, we treat all the sites falling into a grid cell as a single site. We then apply Dijsktra's algorithm to the entire set of nodes with associated sites. This gives us two datasets: (i) a dense distance matrix with least cost travel-times from all farms and community centers to all farms and community centers and (ii) a set of linestring geometries whose vertices are the grid cells traversed on shortest paths. For simplicity, we average travel times to and from each origin and destination point, thus making the distance matrix symmetric. All of the clustering steps in our algorithm are implemented as operations on this matrix.

Importantly, the original resolution of the gridded data is ~30 m, but we aggregate grid cells by a factor of 3, taking the mean in all cases, and making them closer to ~90 m in resolution. Aggregation is a familiar speed-up technique for least-cost analysis on large grids containing tens of millions of cells, though it is not without its risks (Cushman & Landguth, 2010; Doyle et al., 2012; Etherington, 2016), so we also conduct some limited sensitivity analysis with varying grid resolutions, which are reported below (see also Figure 7).

All analyses, including the cost-distance modeling and our implementation of the algorithm, are conducted in the R programming language and environment (R Core Team, 2023). For more details, please see Supplementary Materials, which include all code necessary to reproduce these analyses and the figures in the text.

4 Results

The goals of this analysis are (i) to identify dispersed farming communities based on the distribution of farmsteads and their proximity to community centers and then (ii) to regionalize those communities by giving them meaningful boundaries. This was done by developing a simple algorithm inspired by the concept of a CBSA combined with previous work by Varien (1999b) and Reese et al. (2019). Using parameter values *D-join* = 30 minutes, *D-max* = 1 hour, and $P = 0.8$, our algorithm identified 103 communities in CMV, 88 communities in NRG, and 45 communities in CIB (see Figure 4). The number of communities removed because they included an insufficient number of sites was 5 for CMV, 33 for NRG, and 23 for CIB, leaving 98, 55, and 22 communities with definable spatial extents.

The mean areas of these communities are 9.51 km^2 in CMV, 6.5 km^2 in NRG, and 7.74 km^2 in CIB; the mean room counts (for all community centers and farms) are 663 in CMV, 598.65 in NRG, and 768.05 in CIB; and the mean room densities are 83.83 rooms per km^2 in CMV, 183.93 rooms per km^2 in NRG, and 207.18 rooms per km^2 in CIB. We emphasize that the means in several cases deviate to a considerable degree from median values (denoted in Table 2 with the Greek letter η), indicating skew in the data and the potential presence of large outliers. This is illustrated in the boxplots in Figure 6, which show the distributions of these and other variables of interest that can be derived from our analysis and also used to evaluate its merits (see also Table 2). Note that these summaries are time-insensitive. They represent the total archaeological sample in each region.

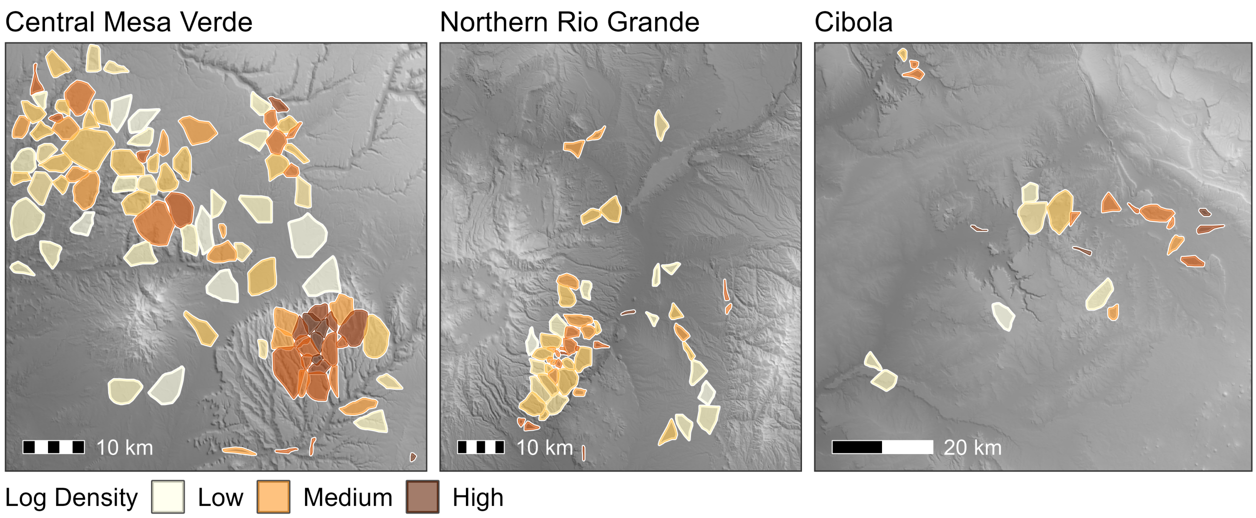

Figure 4. Map showing the size, shape, and spatial distribution of community polygons. Color represents the log density of rooms (the number of rooms per square kilometer), with lighter colors indicating lower values and darker colors indicating higher values. Note that densities in each region were re-scaled to have the same range, so this figure cannot be used to evaluate absolute differences between regions!

In addition to showing the community polygons, Figure 4 also offers a simple demographic profile for each region based on the spatial distribution of all rooms (their count per square kilometer) across communities irrespective of time period. For visualization purposes, the log of the density is used, as it helps to normalize the distribution and, in particular, to rein in large outliers that swamp the variance. It is important to note, too, that values in each region are re-scaled to be in the same range, so the figure cannot be used to estimate absolute differences between regions. Still, the relative differences within each region can be enlightening, though even they must be interpreted with caution.

Table 2. Community Summaries

		Area (sq.km)	Rooms (N)	Farms (N)	Centers (N)	Commute (mins)	Room Density (N/sq.km)
CMV	η	: 7.61	: 399	: 25.5	: 1	: 21.51	: 55.08
	μ	: 9.51	: 663	: 45.61	: 1.67	: 22.72	: 83.83
	σ	: 7.2	: 753.62	: 59.11	: 1.39	: 8.3	: 71.93
NRG	η	: 5.45	: 475	: 18	: 2	: 18.14	: 98.69
	μ	: 6.5	: 598.65	: 36.91	: 2.53	: 18.28	: 183.93
	σ	: 5.07	: 468.12	: 37.55	: 2.17	: 8.65	: 269.84
CIB	η	: 4.59	: 492	: 13.5	: 2	: 20.85	: 157.01
	μ	: 7.74	: 768.05	: 24.23	: 2	: 24.36	: 207.18
	σ	: 7.96	: 626.91	: 25.39	: 1.11	: 12.14	: 218.53

η: median, μ: mean, σ: standard deviation

Not surprisingly, virtually all of the Mesa Verde NP communities show high densities. The community that includes Goodman Point and Sand Canyon Pueblo just north of Ute Mountain also has a relatively high density, as do several of the communities west of Yellow Jacket and along the Dolores River. In the NRG, the highest density community is near the center of the region. It juts out from the Rio Grande at its confluence with the Pojoaque River and includes the San Ildefonso and Pojoaque pueblos. Another high density community lies in a shallow north-south canyon at the base of the Sangre de Cristos. Additional high density communities are observed on the Pajarito Plateau in Bandelier NM. In the CIB area, the three densest communities can be found in the far eastern edge of the study area near El Morro NM, with additional dense communities near Ramah just to the west of El Morro, southeast of Zuni near the center of Cibola, and south of Manuelito in the northwest.

Results of the sensitivity analysis are as expected given the role that the main tuning parameters play in the algorithm (see Figure 7). Increasing *D-join* makes it easier to join two communities, thus leading to growth in the total area of communities and declines in the total number of resulting communities. Increasing P makes it harder to join communities, thus leading to more smaller communities. Increasing *D-max* increases the total area of communities and the average commute time to centers as it serves to filter out distant farmsteads. Importantly, finer-grained grid resolutions had no meaningful impact on the results of this analysis, so the level of aggregation we chose appears acceptable. One puzzling result is that the CIB region shows some volatility in its room densities with respect to the main tuning parameters. The number of CIB communities is also largely invariant. We comment on this in the discussion.

5 Discussion

Our approach offers several advantages over previous efforts to define dispersed community territories. First, it handles variation in survey coverage by seeking to define communities we have evidence for, rather than all communities that may once have existed. Second, it handles variation in the degree to which surveyors have lumped or split architectural remains in defining sites (Kintigh, 2003). This is a special problem in the CMV. The tradition of site survey recording within Mesa Verde National Park has generally involved treating each room block as a separate site, even when there is a continuous artifact scatter between them. Outside of the park, however, the tradition has been to include all room blocks within a single artifact scatter as part of a single site. Because our method incorporates room count estimates, it is robust to these differences in site definition, enabling properties of the resulting communities to be compared more directly across these areas. Finally, our method focuses on travel time, a strong regularity in human affairs, and can be applied to landscapes with dramatically different topographies. The resulting community territories take obvious natural topographic barriers to interaction into account, but also allow territories to cross such boundaries when this is suggested by the data themselves.

Because our algorithm requires that we calculate the commute time between each farm and community center, we can also evaluate potential spatial - strictly, commute time -

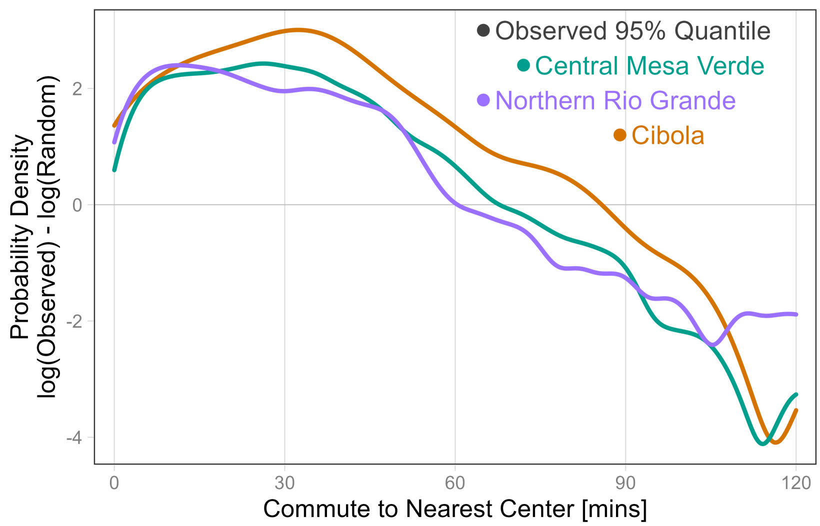

Figure 5. Figure shows the difference between the log density of commute times (in minutes) from observed farms and random locations to the nearest community center for each region, along with the location of their 95% quantile for observed farms (the location of the point next to the respective region label, all being between 60 and 90 minutes).

interactions between them. Figure 5 illustrates this idea by showing the difference between (i) the log probability density of commute times from observed farms to their nearest centers and (ii) the log probability density of commute times from random locations to their nearest centers, along with the 95% quantiles of the observed distributions. This is analogous to what Reese et al. (2019) refer to as the "null difference cost distance." One can interpret it as showing where commute times differ from what one might expect if farm sites were distributed randomly across the landscape. Values larger than zero indicate that there are more farms located at those commute times from centers than you would expect by chance. Conversely, values less than zero indicate that there are less farms located at those locations than one might expect by chance. In either case, these differences tell us that there is positive covariance between farmsteads and nearby centers at commute times of less than approximately one hour (the value we assigned to *D-max*).

The shapes of the density curves are not, however, identical. We attribute this to three factors: sampling intensity, local topography, and regional settlement patterning. For instance, the spike in the CIB region at or around the 30 minute point along with the longer tail (and the larger 95% quantile) is at least partially owing to the fact that block survey in that area is limited to a few disconnected locations. The area is also much flatter than in the CMV and NRG regions, making the background commute distribution much more uniform, at least within the surveyed areas.

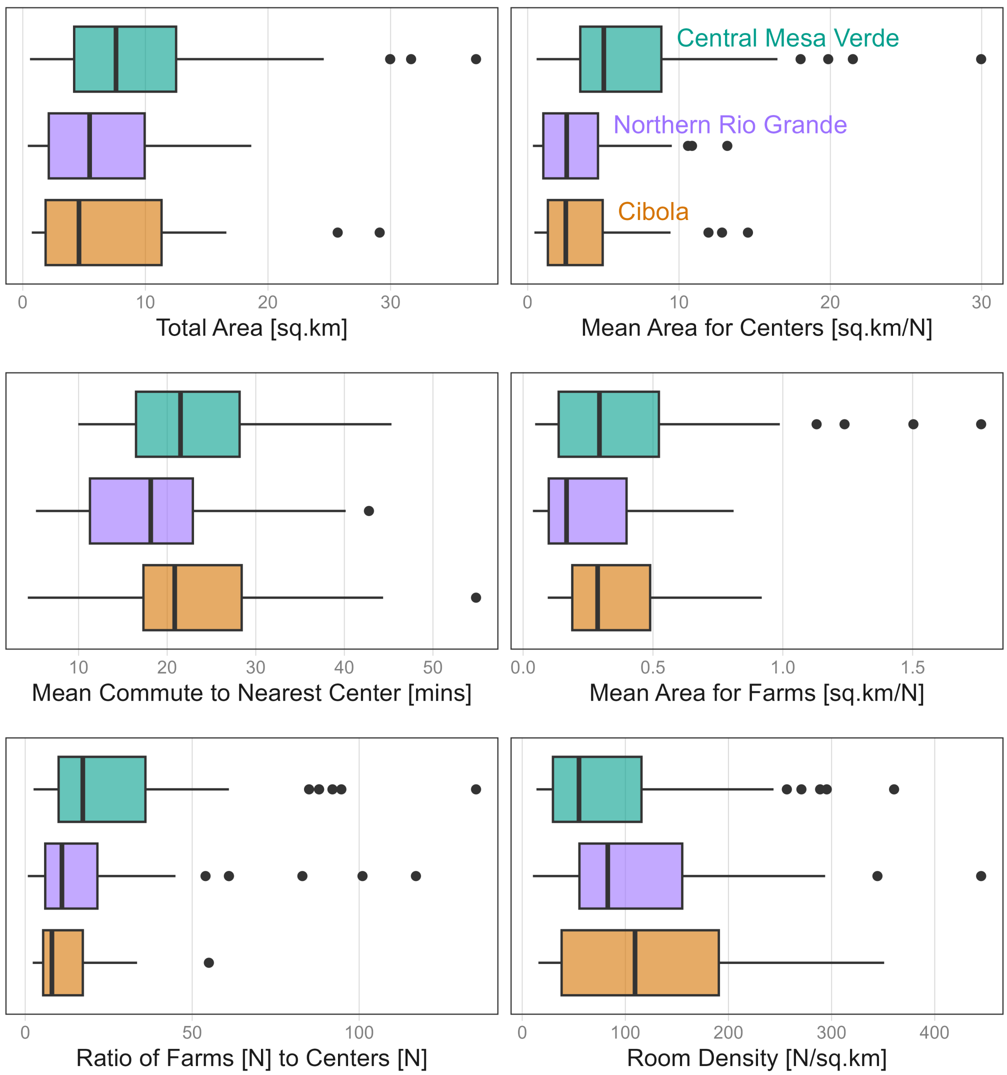

Figure 6. Figure shows for each region the distribution of relevant variables across communities. Note that these are time-insensitive distributions of the samples in each region.

Perhaps more interesting though is the fact that these regions represent different settlement patterns over different periods of time. In the CMV area, the pattern of settlement was largely one of growth and dispersal, with local populations reaching an asymptote after which they spread out into neighboring areas (Schwindt et al., 2016). Nearly all centers in the CMV area arose within farming clusters, but they were relatively small compared to those in the other study areas, and occurred only to a limited degree in the period leading up to the Great Drought in 1280 CE (and to a lesser degree the period leading up to and including much of the 10th century), with no large aggregations occurring afterwards (Glowacki & Ortman, 2012; Ortman, 2016a; Varien, 1999b). In the NRG region, communities had a few dispersed farms early on, more in the period immediately following 1280 (mostly concentrated in the northern Pajarito Plateau, the area of Bandelier NM), and then very large aggregations after that (Ortman, 2016b, 2016a). As suggested by the relatively small ratio of farms to centers (see Figure 6), communities in the CIB region had almost no dispersed farming for most of the sequence, with only an abbreviated phase of dispersed settlement that quickly morphed into large aggregations as populations surged in the period after 1280 (Kintigh, 1985; Peeples, 2018; Schachner, 2012). As in the CMV area, aggregations in the NRG and CIB regions sometimes developed within dispersed farming communities, but unlike in the CMV, many also seem to have been established after 1280 CE in areas away from where dispersed farms had been previously located (Ortman, 2016a; Schachner, 2015).

These settlement patterns explain why our algorithm dropped 23 communities in the CIB region, 35 in the NRG, but only 4 in the CMV. Many of the CIB community centers are simply too isolated from farmsteads for our algorithm to define a meaningful extent. Settlement patterning probably also explains the volatile response of the CIB communities to changes in the tuning parameters, as the small number of farms means small changes are just enough for communities to add or drop community centers that are excluded by our main analysis.

This is perhaps the biggest limitation of our algorithm. It cannot define community boundaries when the relationship between community centers and dispersed farms is not obvious, either because centers are geographically isolated from farms or because there is an insufficient number of farms to serve as a bridge between centers. We note, however, that our concept of a community center does not technically require the presence of an archaeological site. Community center sites are just a clear signal of the locations of persistent interaction in the archaeological record.

Archaeologists and anthropologists have investigated various dimensions of community organization and the types of communities they foster. Examples include biological communities organized around genetic relatedness or physical similarities (Becker & Juengst, 2017; Blom, 2005, 2017; Kakaliouras, 2017), imagined or ideational communities organized around perceived membership and metaphors (sometimes called *styles of imaginings;* Anderson, 1983; Isbell, 2000; Ortman, 2011), reproductive communities organized around mating opportunities (Kolb & Snead, 1997; Mahoney et al., 2000), religious communities organized around a shared sense of the sacred (Bernardini, 2004; Malville & Malville, 2001), linguistic communities organized around a shared grammar and vocabulary (Gumperz, 1968; Silverstein, 1998), even sight communities organized around a shared set of visually striking landmarks (Bernardini & Peeples, 2015). In fact, there are probably as many types of communities as there are types of things for people to care about.

Because the type of community we aim to delineate is a dispersed *farming* community, we structure our algorithm around the dimensions of food production, travel time, and social interaction, as these are major constraints on settlement for subsistence farmers (Bocinsky et al., 2016; Bocinsky & Kohler, 2014; Vernon et al., 2022, in press; Yaworsky et al., 2023). To give just one example, our algorithm suggests that 50% of CMV communities offer anywhere from 15-50 hectares per farm. For the CIB and NRG regions, the corresponding figures are

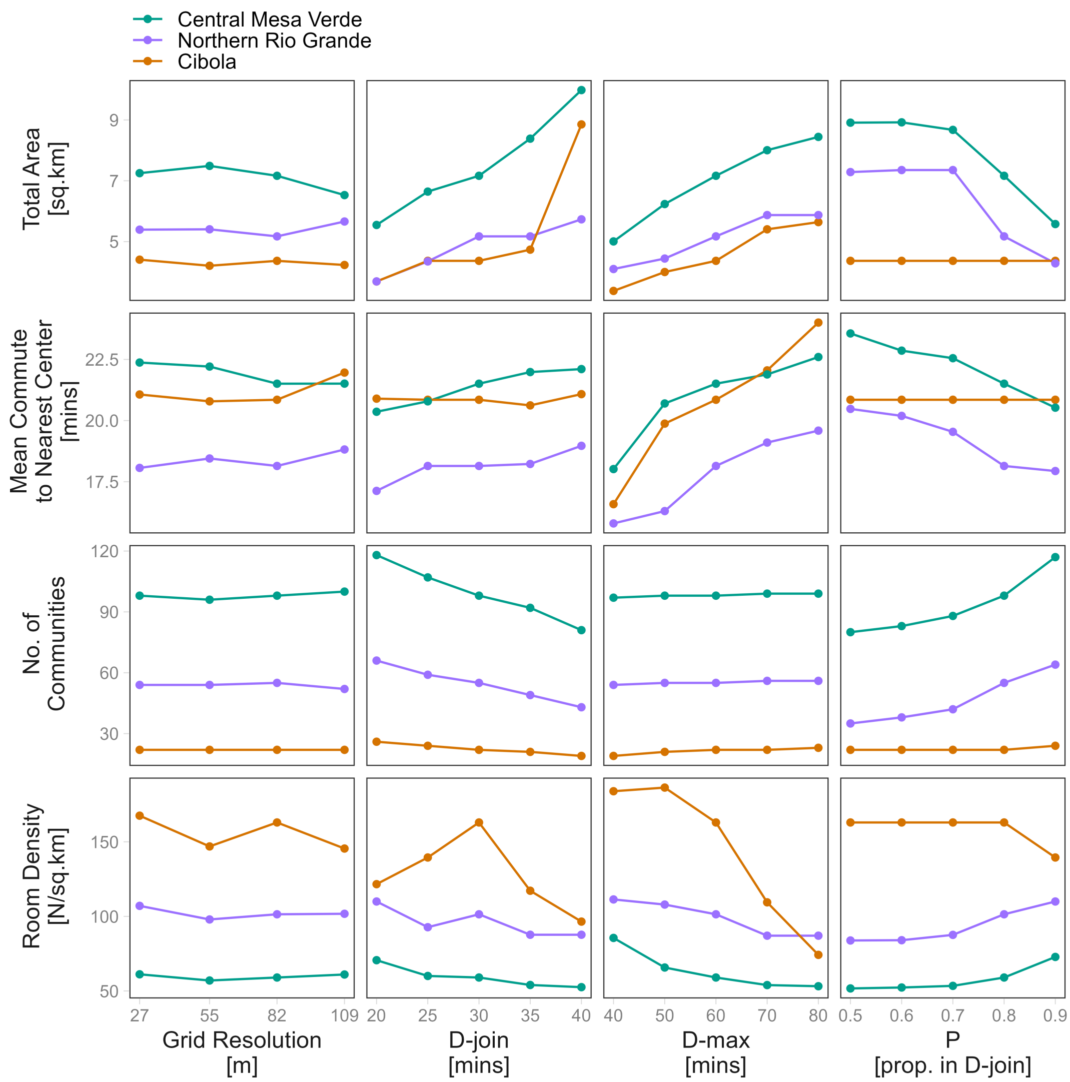

Figure 7. Figure shows results of sensitivity analysis, focusing on three main parameters (*D-join*, *D-max*, and P) and approximate grid resolution. Measures of some key outcomes are also shown, including the number of communities, their total area, room density, and average commute time from farms to the nearest community center.

25-45 and 15-35 hectares per farm, respectively. This means that farms in all three areas meet the 14 hectare threshold required for a household to subsist on maize in these regions (Benson, 2011; Bocinsky & Varien, 2017; Reese et al., 2019). And that is before taking into account that the area of the community is held fixed (pace Reese et al., 2019) and all sites are used regardless of time period, so the derived ratios of land to farms represent minimum possible levels, not the levels actually experienced by farmers in these communities at any given time.

6 Conclusion

In this paper, we develop an algorithm that leverages the difference between farmsteads and community centers to cluster sites into dispersed farming communities and define their boundaries. This was inspired by the concept of a CBSA, along with important contributions from Varien (1999b) and Reese et al. (2019). This algorithm allows us to capture and compare the changing properties of both dispersed and aggregated communities using consistent measurements. For instance, we can use uniform probability density analysis of pottery assemblages from each site (Ortman, 2016b) to apportion rooms to time slices and thus estimate changes in population density through time. The ability to compare these properties for communities of varying scales and densities is crucial for evaluating their relative costs and benefits for human welfare.

7 Acknowledgements

Portions of this research were supported by a grant from the National Science Foundation (#2213921, to Ortman).

8 References

Adler, M. A. (1990). *Communities of soil and stone: An archaeological investigation of population aggregation among the Mesa Verde region Anasazi, AD 900–1300 (Order No. 9116107)* [PhD thesis, University of Michigan]. https://www.proquest.com/dissertations-theses/communities-soil-stone-archaeological/docview/303851225/se-2

Anderson, B. (1983). *Imagined communities: Reflections on the origin and spread of nationalism*. Verso.

Angel, S., Lamson-Hall, P., Blei, A., Shingade, S., & Kumar, S. (2021). Densify and Expand: A Global Analysis of Recent Urban Growth. *Sustainability*, *13*(7). https://doi.org/10.3390/su13073835

Becker, S. K., & Juengst, S. L. (2017). Establishing a bioarchaeology of community. *Archaeological Papers of the American Anthropological Association*, *28*(1), 6–12. https://doi.org/10.1111/apaa.12084

Benson, L. V. (2011). Factors Controlling Pre-Columbian and Early Historic Maize Productivity in the American Southwest, Part 1: The Southern Colorado Plateau and Rio Grande Regions. *Journal of Archaeological Method and Theory*, *18*(1), 1–60. https://doi.org/10.1007/s10816-010-9082-z

Bernardini, W. (2004). Hopewell geometric earthworks: A case study in the referential and experiential meaning of monuments. *Journal of Anthropological Archaeology*, *23*(3), 331–356. https://doi.org/10.1016/j.jaa.2004.06.001

Bernardini, W., & Peeples, M. A. (2015). Sight communities: The social significance of shared visual landmarks. *American Antiquity*, *80*(2), 215–235. https://doi.org/10.7183/0002-7316.80.2.215

Berrey, C. A., Drennan, R. D., & Peterson, C. E. (2021). Local economies and household spacing in early chiefdom communities. *PLOS ONE*, *16*(5), e0252532. https://doi.org/10.1371/journal.pone.0252532

Bettencourt, L. M. A. (2021). *Introduction to Urban Science: Evidence and Theory for Cities as Complex Systems*. The MIT Press.

Birch, J. (2013). *From Prehistoric Villages to Cities: Settlement Aggregation and Community Transformation*. Routledge.

Blom, D. E. (2005). A bioarchaeological approach to tiwanaku group dynamics. In R. M. Reycraft (Ed.), *Us and them: Archaeology and ethnicity in the andes* (pp. 153–182). UCLA Cotsen Institute of Archaeology.

Blom, D. E. (2017). A bioarchaeological perspective on community and the tension between individual and population. *Archaeological Papers of the American Anthropological Association*, *28*(1), 104–111. https://doi.org/10.1111/apaa.12092

Bocinsky, R. K., Gillreath-Brown, A., Kintigh, K., Kinzig, A., Kohler, T., Lee, A., Ludascher, B., McPhillips, T., Nguyen, C., & Pritchard, C. (2022). *Synthesizing Knowledge of Past Environments (SKOPE)* [Web Application]. SKOPE: Synthesizing Knowledge of Past Environments. https://doi.org/10.5281/zenodo.6522974

Bocinsky, R. K., & Kohler, T. A. (2014). A 2,000-year reconstruction of the rain-fed maize agricultural niche in the US Southwest. *Nature Communications*, *5*, 5618.

Bocinsky, R. K., Rush, J., Kintigh, K. W., & Kohler, T. A. (2016). Exploration and exploitation in the macrohistory of the pre-Hispanic Pueblo Southwest. *Science Advances*, *2*(4), e1501532. https://doi.org/10.1126/sciadv.1501532

Bocinsky, R. K., & Varien, M. D. (2017). Comparing maize paleoproduction models with experimental data. *Journal of Ethnobiology*, *37*(2), 282–307. https://doi.org/10.2993/0278-0771-37.2.282

Campbell, M. J., Dennison, P. E., Butler, B. W., & Page, W. G. (2019). Using crowdsourced fitness tracker data to model the relationship between slope and travel rates. *Applied Geography*, *106*, 93–107. https://doi.org/10.1016/j.apgeog.2019.03.008

Campbell, M. J., Dennison, P. E., & Thompson, M. P. (2022). Predicting the variability in pedestrian travel rates and times using crowdsourced GPS data. *Computers, Environment and Urban Systems*, *97*, 101866. https://doi.org/10.1016/j.compenvurbsys.2022.101866

Cushman, S. A., & Landguth, E. L. (2010). Scale dependent inference in landscape genetics. *Landscape Ecology*, *25*(6), 967–979. https://doi.org/10.1007/s10980-010-9467-0

Doyle, J. A., Garrison, T. G., & Houston, S. D. (2012). Watchful realms: Integrating GIS analysis and political history in the southern maya lowlands. *Antiquity*, *86*(333), 792–807. https://doi.org/10.1017/S0003598X0004792X

Drennan, R. D., Berrey, C. A., & Peterson, C. E. (2015). *Regional Settlement Demography in Archaeology*. Eliot Werner Publications.

Ducke, B., & Kroefges, P. C. (2008). From points to areas: Constructing territories from archaeological site patterns using an enhanced xtent model. In A. Posluschny, K. Lambers, & I. Herzog (Eds.), *Layers of perception* (pp. 245–251).

Duranton, G., & Puga, D. (2020). The Economics of Urban Density. *Journal of Economic Perspectives*, *34*(3), 3–26. https://doi.org/10.1257/jep.34.3.3

Duwe, S., & Anschuetz, K. F. (2013). Ecological Uncertainty and Organizational Flexibility on the Prehispanic Tewa Landscape: Notes from the Northern Frontier. In B. J. Vierra (Ed.), *From Mountain Top to Valley Bottom: Understanding Past Land Use in the Northern Rio Grande Valley, New Mexico*. University of Utah Press.

Duwe, S., Eiselt, B. S., Darling, J. A., Willis, M. D., & Walker, C. (2016). The pueblo decomposition model: A method for quantifying architectural rubble to estimate population size. *Journal of Archaeological Science*, *65*, 20–31. https://doi.org/10.1016/j.jas.2015.10.011

Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. *Proceedings of the 2nd ACM International Conference on Knowledge Discovery and Data Mining (KDD)*, 226–231. https://cdn.aaai.org/KDD/1996/KDD96-037.pdf

Etherington, T. R. (2016). Least-cost modelling and landscape ecology: Concepts, applications, and opportunities. *Current Landscape Ecology Reports*, *1*(1), 40–53. https://doi.org/10.1007/s40823-016-0006-9

Gauthier, R. P., & Herhahn, C. (2005). Why Would Anyone Want to Farm Here? In R. P. Powers (Ed.), *The Peopling of Bandelier: New Insights from the Archaeology of the Pajarito Plateau* (pp. 27–34). School of American Research Press.

Gilpin, D. (2003). Chaco-era site clustering and the concept of communities. *KIVA*, *69*(2), 171–205. https://doi.org/10.1080/00231940.2003.11758490

Glowacki, D. M., & Ortman, S. G. (2012). Characterizing Community-Center (Village) Formation in the VEP Study Area. In T. A. Kohler & M. D. Varien (Eds.), *Emergence and Collapse of Early Villages: Models of Central Mesa Verde Archaeology*. Univeristy of California Press.

Gumperz, J. J. (1968). The speech community. In *International encyclopedia of the social sciences* (pp. 381–386).

Gyucha, A. (2019). *Coming Together: Comparative Approaches to Population Aggregation and Early Urbanization*. State University of New York Press.

Isard, W. (1954). Location theory and trade theory: Short-run analysis. *The Quarterly Journal of Economics*, *68*(2), 305–320. https://doi.org/10.2307/1884452

Isbell, W. H. (2000). What we should be studying: The "imagined community" and the "natural community." In M. A. Canuto & J. Yaeger (Eds.), *The archaeology of communities: A new world perspective* (pp. 243–266). Routledge.

Kakaliouras, A. M. (2017). Theory for a bioarchaeology of community: Potentials, practices, and pitfalls. *Archaeological Papers of the American Anthropological Association*, *28*(1), 13–23. https://doi.org/10.1111/apaa.12085

Kintigh, K. W. (1985). *Settlement, Subsistence, and Society in Late Zuni Prehistory*. University of Arizona Press.

Kintigh, K. W. (2003). Coming to Terms with the Chaco World. *KIVA*, *69*(2), 93–116. https://doi.org/10.1080/00231940.2003.11758487

Kolb, M. J., & Snead, J. E. (1997). It's a small world after all: Comparative analyses of community organization in archaeology. *American Antiquity*, *62*(4), 609–628. https://doi.org/10.2307/281881

Lekson, S. H. (1989). Kivas? In W. D. Lipe & M. Hegmon (Eds.), *The Architecture of Social Integration in Prehistoric Pueblos* (pp. 161–167). Crow Canyon Archaeological Center.

Lipe, W. D. (1989). Social Scale of Mesa Verde Anasazi Kivas. In W. D. Lipe & M. Hegmon (Eds.), *The Architecture of Social Integration in Prehistoric Pueblos* (pp. 53–71). Crow Canyon Archaeological Center.

Lipe, W. D., & Hegmon, M. (1989). Historical Perspectives on Architecture and Social Integration in the Prehistoric Pueblos. In W. D. Lipe & M. Hegmon (Eds.), *The Architecture of Social Integration in Prehistoric Pueblos* (pp. 15–34). Crow Canyon Archaeological Center.

Mahoney, N. M., Adler, M. A., & Kendrick, J. W. (2000). The Changing Scale and Configuration of Mesa Verde Communities. *KIVA*, *66*(1), 67–90. https://doi.org/10.1080/00231940.2000.11758422

Malville, J. M., & Malville, N. J. (2001). Pilgrimage and periodic festivals as processes of social integration in chaco canyon. *KIVA*, *66*(3), 327–344. https://doi.org/10.1080/00231940.2001.11758436

Marchetti, C. (1994). Anthropological invariants in travel behavior. *Technological Forecasting and Social Change*, *47*(1), 75–88. https://doi.org/10.1016/0040-1625(94)90041-8

Mills, B. J., Ram, S., Clark, J. J., Ortman, S., & Peeples, M. A. (2020). cyberSW Version 1.0. *Archaeology Southwest*.

Muenchrath, D. A., Kuratomi, M., Sandor, J. A., & Homburg, J. A. (2002). Observation Study of Maize Production Systems of Zuni Farmers in Semiarid New Mexico. *Journal of Ethnobiology*, *22*(1), 1–33.

Murdock, G. P. (1949). *Social structure*. Macmillan.

Ortman, S. G. (2011). Bowls to gardens: A history of tewa community metaphors. In D. M. Glowacki & S. Van Keuren (Eds.), *Religious transformations in the late pre-hispanic pueblo world* (pp. 84–108). The University of Arizona Press.

Ortman, S. G. (2016a). 5 Discourse and Human Securities in Tewa Origins. *Archaeological Papers of the American Anthropological Association*, *27*(1), 74–94. https://doi.org/10.1111/apaa.12075

Ortman, S. G. (2016b). Uniform Probability Density Analysis and Population History in the Northern Rio Grande. *Journal of Archaeological Method and Theory*, *23*(1), 95–126. https://doi.org/10.1007/s10816-014-9227-6

Peeples, M. A. (2018). *Connected Communities: Networks, Identity and Social Change in the Ancient Cibola World*. University of Arizona Press.

Peeples, M. A., & Schachner, G. (2012). Refining correspondence analysis-based ceramic seriation of regional data sets. *Journal of Archaeological Science*, *39*(8), 2818–2827. https://doi.org/10.1016/j.jas.2012.04.040

Peterson, C. E., & Drennan, R. D. (2005). Communities, Settlements, Sites, and Surveys: Regional-Scale Analysis of Prehistoric Human Interaction. *American Antiquity*, *70*(1), 5–30. https://doi.org/10.2307/40035266

R Core Team. (2023). *R: A language and environment for statistical computing* [Manual]. R Foundation for Statistical Computing. https://www.R-project.org/

Reese, K. M., Glowacki, D. M., & Kohler, T. A. (2019). Dynamic Communities on the Mesa Verde Cuesta. *American Antiquity*, *84*(4), 728–747. https://www.jstor.org/stable/26818404

Renfrew, C., & Level, E. V. (1979). Exploring dominance: Predicting polities from centers. In C. Renfrew & K. L. Cooke (Eds.), *Transformations* (pp. 145–167). Academic Press. https://doi.org/10.1016/B978-0-12-586050-5.50016-6

Schachner, G. (2012). *Population Circulation and the Transformation of Ancient Zuni Communities.* University of Arizona Press.

Schachner, G. (2015). Ancestral Pueblo archaeology: The value of synthesis. *Journal of Archaeological Research*, *23*, 49–113.

Schwindt, D. M., Bocinsky, R. K., Ortman, S. G., Glowacki, D. M., Varien, M. D., & Kohler, T. A. (2016). The Social Consequences of Climate Change in the Central Mesa Verde Region. *American Antiquity*, *81*(1), 74–96. https://doi.org/10.7183/0002-7316.81.1.74

Silverstein, M. (1998). Contemporary transformations of local linguistic communities. *Annual Review of Anthropology*, *27*(1), 401–426. https://doi.org/10.1146/annurev.anthro.27.1.401

Smith, E. A. (2003). Human cooperation: Perspectives from behavioral ecology. In P. Hammerstein (Ed.), *Genetic and cultural evolution of cooperation* (pp. 401–427). The MIT Press.

Smith, E. A. (2010). Communication and collective action: Language and the evolution of human cooperation. *Evolution and Human Behavior*, *31*(4), 231–245. https://doi.org/10.1016/j.evolhumbehav.2010.03.001

Stone, T. (2016). Organizational variability in early aggregated communities in middle-range societies: An example from the kayenta region of the american southwest. *American Antiquity*, *81*(1), 58–73. https://doi.org/10.7183/0002-7316.81.1.58

Tobler, W. R. (1993). *Three Presentations on Geographical Analysis and Modeling: Non-isotropic Geographic Modeling, Speculations on the Geometry of Geography And, Global Spatial Analysis* (93 (1)). National Center for Geographic Information and Analysis.

US Geological Survey. (2023). *3D elevation program 1 arc-second digital elevation model*. https://prd-tnm.s3.amazonaws.com/StagedProducts/Elevation/1/TIFF/USGS_Seamless_DEM_1.vrt

2020 Standards for Delineating Core Based Statistical Areas, 86 Federal Register Vol 86 pp 37770-37778 37770 (2021). https://www.federalregister.gov/documents/2021/07/16/2021-15159/2020-standards-for-delineating-core-based-statistical-areas

Varien, M. D. (1999a). Regional context: Architecture, settlement patterns, and abandonment. In M. D. Varien (Ed.), *The Sand Canyon archaeological project: Site testing*. Crow Canyon Archaeological Center.

Varien, M. D. (1999b). *Sedentism and Mobility in a Social Landscape: Mesa Verde \& Beyond*. University of Arizona Press.

Varien, M. D. (2002). Persistent communities and mobile households: Population movement in the central Mesa Verde region, A.D. 950 to 1290. In M. D. Varien & R. H. Wilshusen (Eds.), *Seeking the center place: Archaeology and ancient communities in the Mesa Verde region* (pp. 163–184). University of Utah Press.

Varien, M. D., Ortman, S. G., Kohler, T. A., Glowacki, D. M., & Johnson, C. D. (2007). Historical Ecology in the Mesa Verde Region: Results from the Village Ecodynamics Project. *American Antiquity*, *72*(2), 273–299. https://doi.org/10.2307/40035814

Varien, M. D., & Potter, J. M. (2008). The social production of communities: Structure, agency, and identity. In M. D. Varien & J. M. Potter (Eds.), *The social construction of communities: Agency, structure, and identity in the prehispanic southwest*. Altamira Press.

Vernon, K. B., Yaworsky, P. M., McCool, W. C., Spangler, J. D., Brewer, S., & Codding, B. F. (in press). The fremont frontier: Living at the margins of maize farming. *American Antiquity*. https://doi.org/10.1017/aaq.2024.22

Vernon, K. B., Yaworsky, P. M., Spangler, J., Brewer, S., & Codding, B. F. (2022). Decomposing habitat suitability across the forager to farmer transition. *Environmental Archaeology*, *27*(4), 420–433. https://doi.org/10.1080/14614103.2020.1746880

Yaworsky, P. M., Vernon, K. B., McCool, W. C., Hart, I. A., Spangler, J., & Codding, B. F. (2023). The goldilocks zone for maize agriculture and the settlement and abandonment of the west tavaputs plateau. *Quaternary International*. https://doi.org/10.1016/j.quaint.2023.12.003