ARTICLE

Check for updates

Time-dependent taphonomic site loss leads to spatial averaging: implications for archaeological cultures

Emily Coco [1,3] & Radu Iovita [1,2,3 ✉]

Archaeologists typically define cultural areas on the basis of similarities between the types of material culture present in sites. The similarity is assessed in order of discovery, with newer sites being evaluated against older ones. Despite evidence for time-dependent site loss due to taphonomy, little attention has been paid to how this impacts archaeological interpretations about the spatial extents of material culture similarity. This paper tests the hypothesis that spatially incomplete data sets result in detection of larger regions of similarity. To avoid assumptions of cultural processes, we apply subsampling algorithms to a naturally occurring, spatially distributed dataset of soil types. We show that there is a negative relationship between the percentage of points used to evaluate similarity across space and the absolute distances to the first minimum in similarity for soil classifications at multiple spatial scales. This negative relationship indicates that incomplete spatial data sets lead to an over-estimation of the area over which things are similar. Moreover, the location of the point from which the calculation begins can determine the size of the region of similarity. This has important implications for how we interpret the spatial extent of similarity in material culture over large distances in prehistory.

[1] Center for the Study of Human Origins, Department of Anthropology, New York University, New York, NY, USA. [2] Department of Early Prehistory and Quaternary Ecology, Eberhard Karls University of Tübingen, Tübingen, Germany. [3] These authors contributed equally: Emily Coco, Radu Iovita. ✉email: iovita@nyu.edu

Introduction

Archaeologists have a unique interest among historical scientists to understand the spatial structure of past material culture. This often manifests itself in the definition of large regional cultural-technological areas on the basis of similarity in material culture between sites. However, the inherent incompleteness of the archaeological record due to taphonomic and discovery biases can inhibit the search for cultural areas. In recent years, the effects of this incompleteness have been studied in relation to temporal trends (e.g., Surovell and Brantingham, 2007; Surovell et al., 2009; Contreras and Meadows, 2014; Miller-Atkins and Premo, 2018; Perreault, 2018, 2019). Yet, comparatively little research has been done to understand how the decreased spatial resolution of the archaeological record due to time-dependent site loss (Surovell and Brantingham, 2007; Surovell et al., 2009) impacts the spatial patterns we are attempting to find.

Cultural classification or the definition of cultural taxonomies is fundamental to anthropological and archaeological research (Barth, 1981; Reynolds and Riede, 2019). This is because the past is often seen as made up of bounded, homogeneous groupings with different ways of life and therefore different material cultures (Shennan, 1994; Lucy, 2005). These cultural units are used to structure research, organise the archaeological record, analyse past cultural dynamics, and communicate results (Reynolds and Riede, 2019; Riede et al., 2019). Cultural classification based on material remains in archaeology has been critiqued from a number of directions. Chief among those is the issue with attempting to draw discrete boundaries in material culture variation as a reference for discrete boundaries among cultural or ethnic groups (Barth, 1981; Shennan, 1994; Furholt, 2008; Riede et al., 2019) or, indeed, subspecies (Villa and Roebroeks, 2014). What has become clear to archaeologists is that variations in material culture cannot be directly mapped onto certain groups and that considering different subsets of the archaeological record can reveal different patterns (Barth, 1981; Shennan, 1994; Furholt, 2008; Hodder, 2012). For this reason, many researchers have been advocating for reformulation of named cultural units either through using attributes that track shared transmission histories of cultural traits (Riede et al., 2019), or by considering a more polythetic version of culture to account for cross-cutting patterns of different aspects of material culture (Lucy, 2005; Furholt, 2008).

Despite these critiques, cultural classification remains a mainstay of archaeological practice. Archaeologists routinely compare their excavated materials with those they know from nearby areas and either accept or reject a degree of similarity. These comparisons are sometimes done formally, based on attribute analysis (sometimes supported by multivariate statistics), or, quite often, based on the inspection of published drawings or photographs. Although some contemporary archaeologists use sophisticated mathematical methods to compare artefact shapes (e.g., Eerkens et al., 2006; Lycett, 2016), or indeed, entire assemblages (e.g., Grove and Blinkhorn, 2020), the general classificatory framework is often based on categories that were established earlier using the informal method described above. In addition to the everyday practice of informal assessments of similarity, some studies have turned to quantitative methods to explore the relationship between geographic distance and similarity (i.e., density of particular artefact types, presence/absence of particular traits, etc.) for detecting structure of archaeological cultures (Renfrew, 1977; Kimes et al., 1982; Shennan et al., 2015; Lycett, 2019). Here, the premise is that groups in closer proximity to each other tend to interact and share knowledge more frequently than those farther apart (Johnson et al., 2006; Ross and Atkinson, 2016; Derungs et al., 2018). Similarity based on proximity is also known as Tobler's First Law of Geography, which states that everything is related to everything else, but near things are more related than distant things (Tobler, 1970, 2004; Sui, 2004). Incidentally, Tobler's First Law does not only apply to cultural entities, but to all geographic patterns. Examples of this in non-human geography literature most frequently look at the distance decay of similarity in species composition of ecological communities (Nekola and White, 1999; Bjorholm et al., 2008; e.g., Astorga et al., 2012; Wetzel et al., 2012).

Given that archaeologists are still generally interested in finding meaningful spatial structure of material culture despite the low spatial resolution of the archaeological record, we must ask: to what extent are our perceived spatial entities influenced by the incompleteness of the archaeological record? Cultural entities are often defined over vast distances. Although human interaction surely plays a role in the creation of spatially detectable similarity in material culture, the appearance of similar assemblages in distant areas of the world might require a different explanation. Typically, long-distance similarity in material culture is explained by the migration of groups, the vertical transmission of cultural-technological traditions through interaction, or convergent evolution through independent innovation (Crema et al., 2014; Ross and Atkinson, 2016; O'Brien et al., 2018). However, even if these explanations have validity on certain scales, cultural regions defined especially in the earlier periods of prehistory, such as the Palaeolithic, are massive. Some, such as the Aurignacian or the Gravettian, extend across the European continent and beyond (Reynolds and Riede, 2019). The Acheulian stretches over three continents and lasts over a million years. The ubiquity of such continental-sized cultural regions prompts us to doubt that material culture similarities on such scales were in fact the result of sociocultural processes.

Recent studies have shown that time- and taphonomy-dependent factors can distort the spatial component of behavioural signals. For example, Miller-Atkins and Premo (2018) demonstrate that time averaging of assemblages can increase the apparent spatial spread of a cultural signature beyond the area actually occupied by that population at any given moment in time. We hypothesise that a similar problem exists for spatial averaging, such that as fewer and fewer points are used to examine similarity across space, the area included in a region of similarity increases. This would mean that cultural regions are systematically overestimated, given the inherent incomplete nature of the archaeological record. To test this hypothesis, we chose to measure the effects of spatial averaging on spatially autocorrelated point data. As any archaeological data set is already incomplete in an uncontrollable way, we chose a contemporary data source that is both spatially autocorrelated and culturally independent: soil classifications. Soil types represent a spatially complete data set with a certain amount of temporal stability that demonstrates spatial autocorrelation but is not the result of cultural processes. For these reasons, it is possible to spatially subsample this data to determine how that affects analysis of similarity without invoking cultural explanations. This allows us to test the effects of spatial averaging without arguing for a similarity between the initial distribution of traits in our toy data set and that of any particular archaeological case.

Methods

Data. The European Soil Database v2.0 (European Soil Data Centre, 2004) is a soil database for all of Europe and Russia, which contains a geographical database of soil polygons with classifications attributes of different soil properties. For our analysis, we converted the polygons to point data using the

centroid of the polygons for point placement in ArcGIS 10.4.1 (ESRI, 2016). The analysis was run using two subsets of the data, one which included all of Europe (~25,000 points) and another which contained only the points for a single country (Germany, with ~2500 points) given there are often inconsistencies in large-scale maps due to different national mapping systems (Sprafke, 2016).

The categorical attributes used for analysis were classifications of the topsoil and subsoil, slope, accumulate temperature, and depth to rock, as well as the full soil code from the World Reference Base for Soil Resources (see Table 1). These were selected from over 70 total soil attributes for the raw data set to give a wide range of comparisons for calculating the similarity between soil points. Additionally, this range of variables was used to best approximate how informal archaeological comparisons are made: through comparison of absence/presence of particular trait-based variables and/or proportions of different categories of artefacts between sites.

Calculating geographic distance to minimum of similarity. Our spatial averaging methodology consists of repeatedly constructing decay curves for similarity from one random point to a set of other points and then comparing the absolute geographic distances to the first minimum in similarity. All analyses were carried out in R 3.6.1 (R Core Team, 2019) and all scripts can be found on GitHub (https://github.com/cocoemily/spatial-averaging). Geographic distances between the points are calculated in kilometres. Similarity between points is calculated through simple matching of categorical attributes as a percentage of all categories. This means that points that share all of the same values for each of the 13 attributes would have a similarity score of 1, whereas points that share the same values for 12 of the 13 attributes would have a similarity score of 0.92. We modelled similarity in this way because some of the attributes have similar values (as seen in Table 1), and therefore needed to be treated as an ordered set, which standard similarity indices do not account for. Geographic distance is then used as a predictor variable for similarity in fitting a smoothed function to the data using a generalised additive model. The function of this model is then analysed to find the first local minimum (see Fig. 1 for example curves). We chose to analyse the first local minimum of the similarity curve (hereafter, first minimum of similarity) because it best approximates how archaeologists construct cultural zones in practice.

This process is repeated for progressively smaller proportions of the total points. Specifically, we used randomly sampled points at 10% proportion intervals from 90 and 10%, and at 1% intervals from 10 to 1%. These intervals were chosen because the total number of soil data points is ~25,000, so therefore these small percentages better approximate the number of sites we might expect in any given archaeological time period.

Additionally, we calculated what proportion of points could theoretically remain at three time intervals, 10 ka BP, 50 ka BP, and 100 ka BP, by fitting an exponential function to radiometric dates for European Palaeolithic sites dating from 7000 to 900,000 BP (Vermeersch, 2019). The database contains radiocarbon, TL, OSL, ESR, Th/U, and AAR dates from the European Lower, Middle, and Upper Palaeolithic from >10,000 sites (Vermeersch, 2019). For this study, we used over 14,000 radiometric dates of the nearly 18,000 dates in the total data set; we did not include dates coded as "unreliable." The best fit function (1) below explains ~72% of the variance in date frequency, where n_t is the number of radiocarbon dates surviving at time t.

$$n_t = 412.5e^{-4.073 \times 10^{-5}t} \quad (1)$$

Using the above equation, the number of sites at the three aforementioned times is calculated and divided by the total

number of radiometric dates. These proportions were then used as sampling proportions for the methodology described above. This analysis allowed us to model how much data *could* potentially be lost due to taphonomy using a curve based on empirical archaeological data.

As the points are randomly sampled, we repeated the methodology 100 times for all proportions to capture variation. For each subset of the total points, the distance corresponding with the first minimum in similarity for the fitted function was recorded.

Results
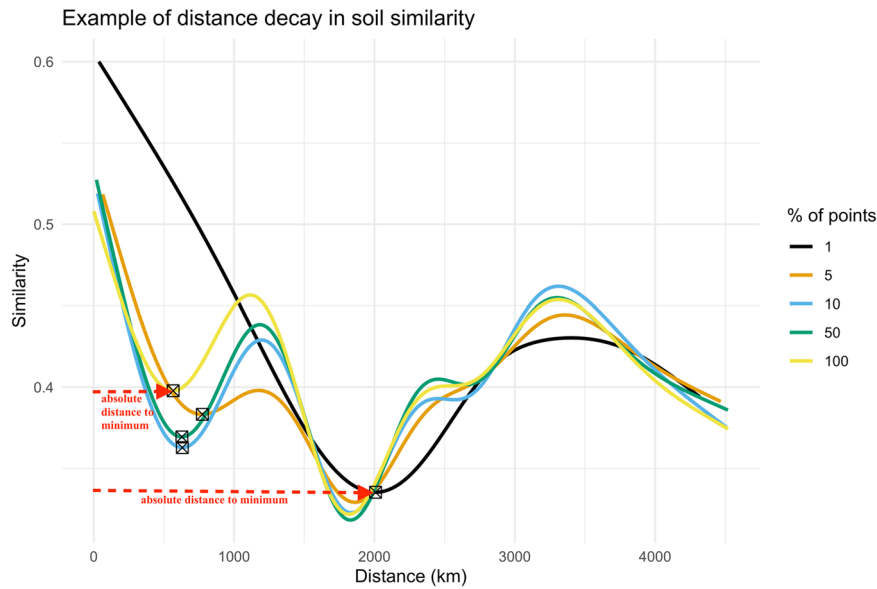**Absolute distance to first minimum for similarity.** A negative power function of the form: $d = \alpha P^\beta$, where d is the distance (in kilometres) to the first similarity minimum, P is the percentage of total points, and α and β are estimated parameters, was fit to the values of absolute distance to the first minimum in similarity by percentage of total points via nonlinear least squares regression (Fig. 2a). This function was determined to be a better fit than a linear regression model and an exponential function using the AIC. The estimates of both α and β were significant to $p < 0.001$; α was estimated to be 1322 with a 95% confidence interval of [81.34, 183.06], meanwhile β was estimated to be -0.04148 with a 95% confidence interval of $[-0.052, -0.031]$. The residual standard error of the model is relatively high at 513.1, which makes sense considering the wide range of variation in the distance to the first minimum at each percentage that results from each sample having a different starting point for comparison. That being said, the negative power relationship is still significant, meaning that as fewer points are considered, the distance over which things are similar increases for this particular dataset. This negative power relationship between distance to first minimum of similarity and percentage of points is confirmed by linear regressions on the logged distances at percentages from 100 to 10% and 10 to 1% separately (Fig. 2b, red line and blue line respectively). The negative slope for lower percentages is nearly twice as great as that of the higher percentages (see Table 2), further demonstrating how the effects of spatial averaging are amplified at the smallest amounts of total points on a continental scale.

Results within national boundaries for soil similarity. Because of inconsistencies at the international level among soil classification systems, we applied our model to a data set at the national level as well, in this case Germany. This allows for examination of how spatial averaging affects analysis of similarity at a smaller scale. The results for soil points within Germany demonstrate similar negative relationships between the percentage of total points and the absolute distances to the first minimums of similarities. However, for this smaller data set, a log-linear model better fits the data then the negative power model that was fit to the distance decay results from the total data set (see Fig. 3). The regression results are presented below in Table 3.

Comparison of "aged" samples. We wanted to examine how the areas of similarity compared when the soil types were sampled at rates corresponding to specific ages of deposits. Surovell and colleagues have demonstrated that the expected number of sites surviving in the archaeological record depends on the time since their deposition (Surovell and Brantingham, 2007; Surovell et al., 2009). Using Surovell and colleagues (2007, 2009) methodology, we fit an exponential function to radiometric dates for European Palaeolithic sites dating from 7000 to 900,000 BP, in order to determine how many points to use for looking at similarity decay corresponding to modelled ages of deposition. We then compared the difference in distance to the first minimum of similarity for

Table 1 Soil attributes used for analysis.

Attribute code	Attribute Name	Values
WRBFU	Full soil code	Four-letter soil code from the World Reference Base for Soil Resources
SLOPEDO	Dominant slope class	0 = No information
		1 = Level (dominant slope ranging from 0 to 8 %)
		2 = Sloping (dominant slope ranging from 8 to 15 %)
		3 = Moderately steep (dominant slope ranging from 15 to 25 %)
		4 = Steep (dominant slope over 25 %)
DR	Depth to rock	S = Shallow (<40 cm)
		M = Moderate (40–80 cm)
		D = Deep (80–120 cm)
		V = Very deep (>120 cm)
ATC	Accumulated mean annual temperature class	H = High (>3000 deg C)
		M = Medium (1800–3000 deg C)
		L = Low (<1800 deg C)
TXSUBDO	Dominant sub-sufrace textural calss	0 = No information
		9 = No mineral texture (Peat soils)
		1 = Coarse (18% < clay and > 65% sand)
		2 = Medium (18% <clay < 35% and > = 15% sand, or 18% < clay and 15% < sand < 65%)
		3 = Medium fine (<35% clay and <15% sand)
		4 = Fine (35% < clay < 60%)
		5 = Very fine (clay > 60%)
PD_SUB	Subsoil packing density	L = Low
		M = Medium
		H = High
STR_SUB	Subsoil structure	G = Good
		N = Normal
		P = Poor
		O = Peaty subsoil
MIN_SUB	Subsoil mineralogy	KQ = 1/1 Minerals + Quartz
		KX = 1/1 min + Oxy. & Hydroxy.
		MK = 2/1 & 1/1 Minerals
		M = 2/1 & 2/1/1 non swel. Min.
		MS = Swel. & non swel. 2/1 min
		S = Swelling 2/1 Minerals
		TV = Vitric Minerals
		TO = Andic Minerals
		NA = Not applicable
TXSRFDO	Dominant surface textural class	0 = No information
		9 = No mineral texture (Peat soils)
		1 = Coarse (18% < clay and > 65% sand)
		2 = Medium (18% < clay < 35% and > = 15% sand, or 18% < clay and 15% < sand < 65%)
		3 = Medium fine (<35% clay and < 15% sand)
		4 = Fine (35% < clay < 60%)
PD_TOP	Topsoil packing density	L = Low
		M = Medium
		H = High
STR_TOP	Topsoil structure	G = Good
		N = Normal
		P = Poor
		H = Humic or Peaty topsoil
MIN_TOP	Topsoil mineralogy	KQ = 1/1 Minerals + Quartz
		KX = 1/1 min + Oxy. & Hydroxy.
		MK = 2/1 & 1/1 Minerals
		M = 2/1 & 2/1/1 non swel. Min.
		MS = Swel. & non swel. 2/1 min
		S = Swelling 2/1 Minerals
		TV = Vitric Minerals
		TO = Andic Minerals
		NA = Not applicable
PMH	Parent material hydrogeological type	R = Porous - Stor. ~ Perm. +
		C = Porous 2 Stor. ~ Perm. +
		S = Porous 1 Stor. + Perm. +
		L = Stor. - Perm. -
		H = Hard. Stor. – Perm. –
		M = Soft. Stor. – Perm. –
		# = No information

Fig. 1 Example of distance decay curves at five different proportion percentages with first minimum of similarity indicated on each curve by a boxed X. Absolute distance to first minimum on the curve constructed with 100% of points and that with 1% of points indicated by the red arrows.
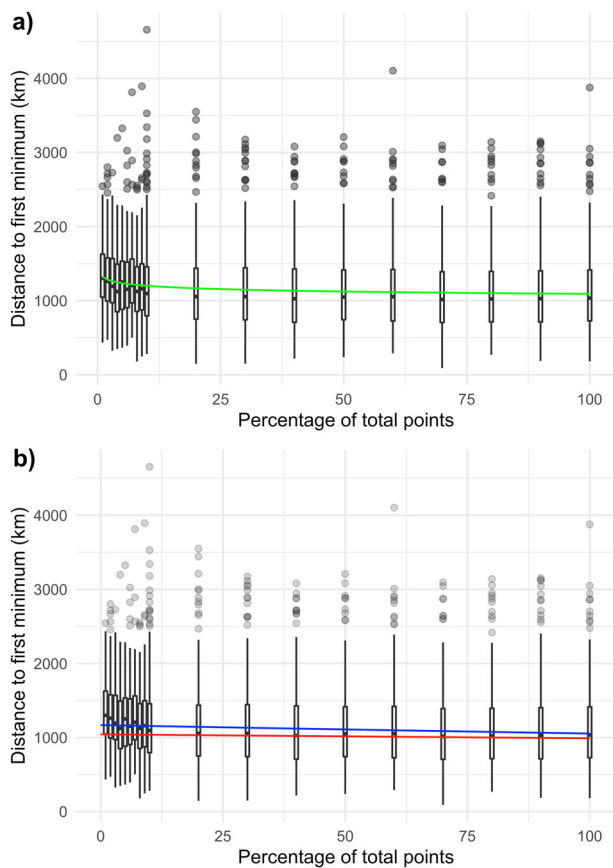
Fig. 2 Regressions fitted to values of absolute distance to first minimum of similarity by the percentage of total points sampled for the total dataset. In **a**, negative power function (green line) fitted to all data for the absolute distance to first minimum. In **b**, linear regressions fitted to subsets of the data; the red line is fitted to samplings up to 10%, whereas the blue line is fitted to samplings from 10 to 1%.

	Table 2 Linear regression results for percentage subsets (soil data).	
	Dependent variable:	
	Log (absolute distance)	
	100–10%	**10–1%**
Percentage of points	−0.0005**	−0.001**
	0.0002	0.0005
Constant	6.950***	7.063***
	0.014	0.016
Observations	4,939	843
R^2	0.001	0.006
Adjusted R^2	0.001	0.004
Residual std. error	0.465 (df = 4937)	0.419 (df = 841)
F-Statistic	5.109** (df = 1; 4937)	4.764** (df = 1; 841)

*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

deposits aged at 10,000, 50,000, and 100,000 years ago based on 491, 96, and 12 points, respectively.

We found again that there were fairly wide ranges of variation for the relative differences in distance to the first minimum at each of the three modelled ages of deposition. However, when comparing these groups using a Mann–Whitney U-test there were significant ($p < 0.01$) differences between the 10 and 100 ka and between the 10 and 50 ka distance to first similarity minimum. Additionally, as can be seen in Fig. 4, there is an increase in the absolute distances to the first minimum as the age of deposits increases, meaning that as the modelled age of our deposits increases, so does the distance to the first minimum in similarity. That being said, there appears to be a certain proportion of the total data set at which the relationship between the increase in distance to the first minimum of similarity and number of points becomes negligible, as demonstrated by the lack of significant difference between the 50 ka sampling and the 100 ka sampling, which represent ~0.004% and ~0.0005% of the total soil points, respectively.
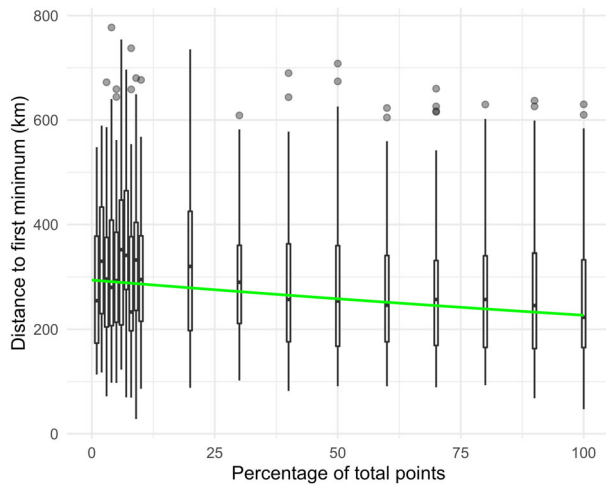
Fig. 3 Linear regression for soil types within Germany. Linear regression models fitted to data for the absolute distances to first minimum in similarity using soil points from within Germany.

**Table 3 Linear regression results for points within a national boundary (Germany).**

	Dependent variable:
	Log (absolute distance)
Percentage of points	−0.003***
	0.0004
Constant	5.684***
	0.021
Observations	1354
R^2	0.036
Adjusted R^2	0.036
Residual std. error	0.481 (df = 1352)
F-Statistic	50.825*** (df = 1; 1352)
*p < 0.1; **p < 0.05; ***p < 0.01.	

Location of reference point. Our results showed that the location of the starting point does not affect the overall trend of increasing distance to similarity with more site loss. However, the identified region of similarity will have different sizes depending on where the initial comparison point is located, as is evident in Fig. 5. The area of similarity at the lowest proportion of points in Fig. 5b is nearly twice times as large as in Fig. 5a, when the reference point is in a different location. This finding is consistent with the large ranges of variation in distance to first minimum of similarity noted above. This finding is of particular importance to archaeological research because the definition of cultural areas is typically based on similarity to 'type assemblages,' which are defined by the order of discovery. Our analysis shows that attempting to define regions of similarity based on comparison to a single starting point not only makes those regions subject to the effects of spatial averaging, but also this means the effects of spatial averaging will differ depending on where that starting point is. This quantitatively demonstrates the well-known limitations of applying the definitions of cultural areas in one area of the world to the archaeological record of another; depending on where the initial definition occurs, the region of similarity will be of a different size, no matter how spatially incomplete the dataset is.

Discussion

The results presented above indicate that as we consider fewer points for determining regions of similarity, the distance to the
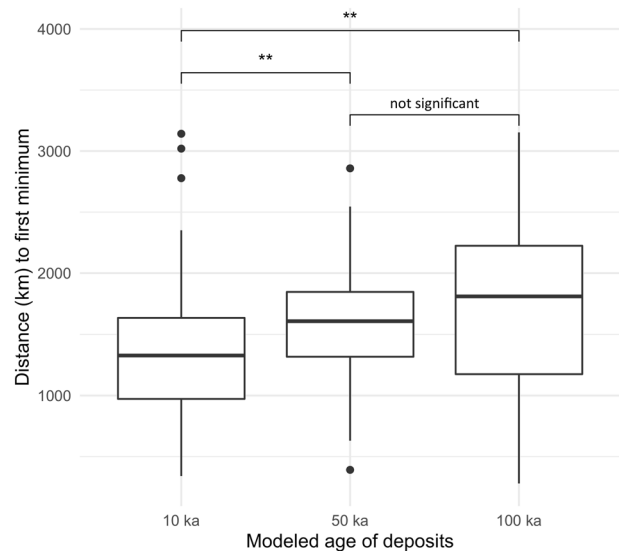

Fig. 4 Comparison of distance to first minimum of similarity between "aged" deposits. Comparison via Mann–Whitney U-tests in absolute difference in distance to first minimum in similarity for expected number of surviving archaeological sites at 10,000, 50,000, and 100,000 years ago. ** = significant to $p < 0.01$.

first minimum in similarity increases. This conclusion should apply to all spatially autocorrelated datasets, whether they are anthropogenic or not, but it is particularly important for anthropogenic ones, where the behavioural signal can be all but erased by spatial averaging. All of the analyses showed that the distance to first minimum in similarity increases by 0.5–3% with each percentage point of reduction in the number of points considered. When we consider that our analysis is on the scale of thousands of kilometres, these small percentage increases in distance to first minimum of similarity result in large over-estimates of similarity areas. It should be noted that the R-squared values on the regressions are small, explaining at most 4% of the variation. This is likely due to the large variation in how similar points are to each other at a given distance (see the large error bars in Figs. 1–3). As stated above, this variation results from different reference points and different comparison points used for each subsampling. Soil types do not exhibit perfect gradients in terms of regions of similarity; instead they are distributed in a mosaic, with similarity decreasing at one distance only to increase again at a farther distance. Despite this variation, the negative relationships presented above are still statistically significant, which means the proportion of points considered does have a significant impact on the size of the region of similarity identified, whether that is at the continental scale or a more regional scale (i.e., within Germany). The goal of this study was to identify if missing point data results in an extension of identified areas of similarity, not to predict the magnitude of this effect. Given that this magnitude is likely dependent on the original signal strength, future studies should use archaeologically-relevant data to assess it.

Conclusions

The effect of spatial averaging on our understanding of areas of similarity has two important implications for archaeology. First, the amount of spatial averaging is inversely proportional to the amount of data remaining. This should affect both older periods and geographic regions with histories of significant landscape changes. We should expect to have fewer and fewer data points as
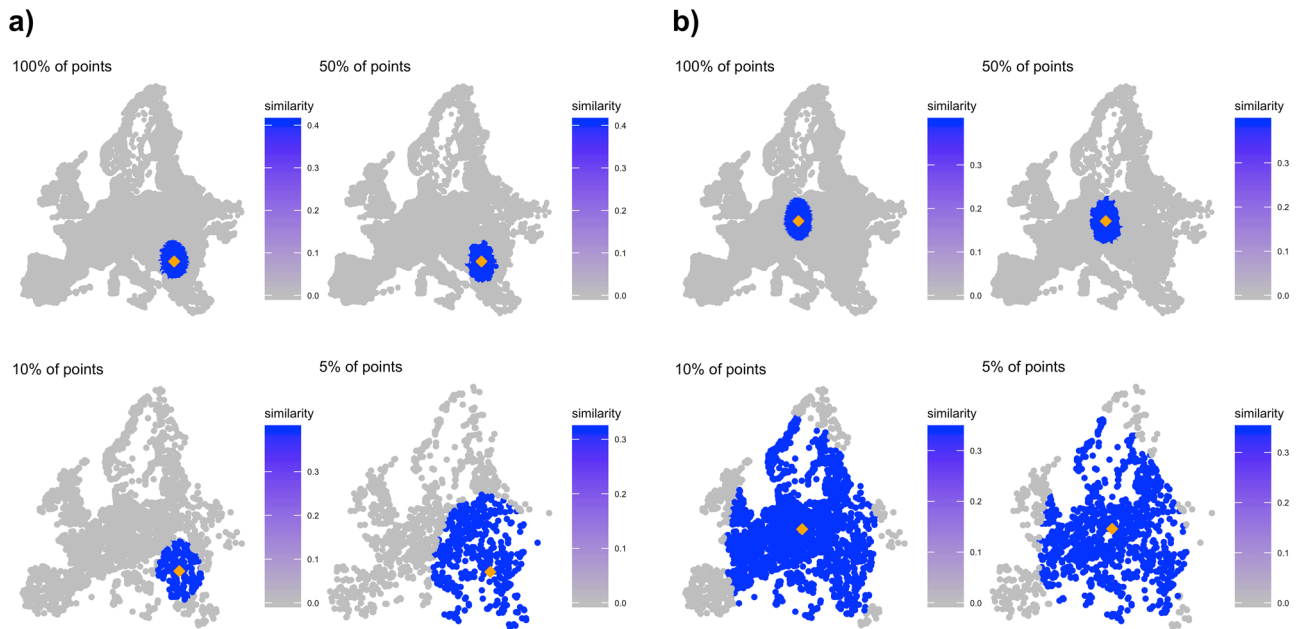
Fig. 5 Distance to first minimum of similarity with each point within that distance coloured in blue. The left four images are based on a point starting in southeastern Europe, whereas the right four images are based on a point in central Europe. Map projection is Lambert Azmuthal Equal Area Europe (ESPG:3035).

we consider older and older periods of prehistory (Surovell et al., 2009). This means that for older archaeological deposits, especially those from the Stone Age, we are likely vastly overestimating regions of similarity just by virtue of not having the complete data set of points, particularly when comparing similarity on an attribute basis. For large parts of the Stone Age, this effect is compounded by the common use of lithics as the single available type of material culture. Likewise, we can expect the amount of erosion, sediment accumulation, and overprinting to produce different amounts of spatial averaging in different regions (e.g., Wilson, 1988; Fanning and Holdaway, 2004; van Leusen et al., 2011; Lovis et al., 2012; Iovita et al., 2014).

Second, the geographic location of the first point of comparison matters. This makes sense if similarity is not originally uniformly distributed, and especially if the information loss will affect the underlying distribution at different rates. As stated above, this is likely to create undetectable biases in classifications, as chance, research history, the availability of funding, and many other factors unrelated to ancient behaviour will determine where any particular type of material culture is first found.

Our report of spatial averaging adds an important dimension to the recent focus on time averaging in archaeological deposits. Several recent papers highlight its role in the formation and interpretive potential of the archaeological record (Davies et al., 2016, 2018; Dibble et al., 2017; Coco et al., 2020; Rezek et al., 2020). Furthermore, the effects of time averaging as a result of the way archaeologists group data has become better established (Miller-Atkins and Premo, 2018; Perreault, 2018). For example, Miller-Atkins and Premo use agent-based simulations to show that considering time-averaged assemblages can increase the apparent spatial spread of a cultural signature beyond the actual occupation region of a population with that particular cultural variant (Miller-Atkins and Premo, 2018). The effects of spatial averaging due to missing data points are similar to those demonstrated for time-averaged assemblages. The results from the study presented here demonstrates that spatial averaging separately from time averaging can also increase the area within a

region of similarity. Further research should look into how the combined effects of time and spatial averaging impact archaeological patterns.

Given that both temporal and spatial averaging are inherent biases of the archaeological record, we need to consider more carefully how we use similarity to demonstrate the cultural relatedness of any two sites in archaeology. This is not to say that archaeologists should not be organising the archaeological records into units of analysis based on similarity. Such a goal is probably unrealistic and impractical. Instead, the results presented here demonstrate that the similarity noted in the archaeological record does not necessarily map onto the original spatial organisation of cultures at the time of interest due to time-dependent taphonomic information loss. Thus, we as archaeologists need to rethink using the size of a defined material cultural area to discuss population and cultural dynamics. The inherent spatial and temporal incompleteness of the archaeological record and its demonstrated effects on interpretations of cultural areas clearly shows that archaeological "cultural" areas are not true representations of *actual living cultures*. As such, we need to further examine what behavioural and cultural signals our regions of similarity actually represent and better incorporate that into our interpretations of the archaeological record.

Data availability

References

Astorga A, Oksanen J, Luoto M, Soininen J, Virtanen R, Muotka T (2012) Distance decay of similarity in freshwater communities: do macro-and microorganisms follow the same rules? Glob Ecol Biogeogr 21:365–375. https://doi.org/10.1111/j.1466-8238.2011.00681.x

Barth F (ed.) (1981) Ethnic groups and boundaries. Process and Form in Social Life. London, Routledge Kegan & Paul, pp. 198–227

Bjorholm S, Svenning J-C, Skov F, Balslev H (2008) To what extent does Tobler's 1st law of geography apply to macroecology? A case study using American palms (Arecaceae). BMC Ecol 8:11. https://doi.org/10.1186/1472-6785-8-11

Coco E, Holdaway S, Iovita R (2020) The effects of secondary recycling on the technological character of lithic assemblages. J Paleolit Archaeol 3:453–474. https://doi.org/10.1007/s41982-020-00055-4

Contreras DA, Meadows J (2014) Summed radiocarbon calibrations as a population proxy: a critical evaluation using a realistic simulation approach. J Archaeol Sci 52:591–608. https://doi.org/10.1016/j.jas.2014.05.030

Crema ER, Kerig T, Shennan S (2014) Culture, space, and metapopulation: a simulation-based study for evaluating signals of blending and branching. J Archaeol Sci 43:289–298. https://doi.org/10.1016/j.jas.2014.01.002

Davies B, Holdaway S, Fanning PC (2016) Modelling the palimpsest: an exploratory agent-based model of surface archaeological deposit formation in a fluvial arid Australian landscape. The Holocene 26:450–463

Davies B, Holdaway S, Fanning PC (2018) Modeling relationships between space, movement, and lithic geometric attributes. Am Antiq 83:444–461. https://doi.org/10.1017/aaq.2018.23

Derungs C, Köhl M, Weibel R, Bickel B (2018) Environmental factors drive language density more in food-producing than in hunter–gatherer populations. Proc R Soc B Biol Sci 285:20172851. https://doi.org/10.1098/rspb.2017.2851

Dibble HL, Holdaway S, Lin SC, Braun DR, Douglass MJ, Iovita R et al. (2017) Major fallacies surrounding stone artifacts and assemblages. J Archaeol Method Theory 24:813–851. https://doi.org/10.1007/s10816-016-9297-8

Eerkens JW, Bettinger RL, McElreath R (2006) Cultural transmission, phylogenetics, and the archaeological record. In: Lipo CP, O'Brien MJ, Collard M, Shennan SJ (eds) Mapping our ancestors: phylogenetic methods in anthropology and prehistory. Aldline Transaction, pp. 169–183

ESRI (2016) ArcGIS desktop. Environmental Systems Research Institute, Redlands, CA

European Soil Data Center (2004) European Soil Database v2.0

Fanning PC, Holdaway SJ (2004) Artifact visibility at open sites in western New South Wales. Australia. J Field Archaeol 29:255–271. https://doi.org/10.1179/jfa.2004.29.3-4.255

Furholt M (2008) Pottery, cultures, people? The European Baden material re-examined. Antiquity 82:617–628. https://doi.org/10.1017/S0003598X0009726X

Grove M, Blinkhorn J (2020) Neural networks differentiate between Middle and Later Stone Age lithic assemblages in eastern Africa. PLoS ONE 15:e0237528. https://doi.org/10.1371/journal.pone.0237528

Hodder I (2012) Entangled: an archaeology of the relationships between humans and things. John Wiley & Sons

Iovita R, Doboş A, Fitzsimmons KE, Probst M, Hambach U, Robu M et al. (2014) Geoarchaeological prospection in the loess steppe: Preliminary results from the Lower Danube Survey for Paleolithic Sites (LoDanS). Quat Int 351:98–114. https://doi.org/10.1016/j.quaint.2013.05.018

Johnson DKN, Siripong A, Brown AS (2006) The demise of distance? the declining role of physical proximity for knowledge transmission. Growth Change 37:19–33. https://doi.org/10.1111/j.1468-2257.2006.00304.x

Kimes T, Haselgrove C, Hodder I (1982) A method for the identification of the location of regional cultural boundaries. J Anthropol Archaeol 1:113–131. https://doi.org/10.1016/0278-4165(82)90017-4

van Leusen M, Pizziolo G, Sarti L (eds) (2011) Hidden landscapes of mediterranean Europe. BAR International Series. Archaeopress, Oxford

Lovis WA, Monaghan GW, Arbogast AF, Forman SL (2012) Differential temporal and spatial preservation of archaeological sites in a Great Lakes coastal zone. Am Antiq 77:591–608. https://doi.org/10.2307/23486290

Lucy S (2005) Ethnic and cultural identities. Archaeology of Identity. Routledge, London, pp. 86–109

Lycett SJ (2016) The Importance of a "Quantitative Genetic" approach to the evolution of artifact morphological traits. In: Mendoza Straffon L (ed) Cultural phylogenetics: concepts and applications in archaeology, interdisciplinary evolution research. Springer International Publishing, Cham, pp. 73–93

Lycett SJ (2019) Confirmation of the role of geographic isolation by distance in among-tribe variations in beadwork designs and manufacture on the High Plains. Archaeol Anthropol Sci 11:2837–2847. https://doi.org/10.1007/s12520-018-0742-3

Miller-Atkins G, Premo LS (2018) Time-averaging and the spatial scale of regional cultural differentiation in archaeological assemblages. STAR Sci Technol Archaeol Res 4:12–27. https://doi.org/10.1080/20548923.2018.1504490

Nekola JC, White PS (1999) The distance decay of similarity in biogeography and ecology. J Biogeogr 26:867–878. https://doi.org/10.1046/j.1365-2699.1999.00305.x

O'Brien MJ, Buchanan B, Eren MI (eds) (2018) Convergent evolution in stone-tool technology. The MIT Press

Perreault C (2018) Time-averaging slows down rates of change in the archaeological record. J Archaeol Method Theory 25:953–964. https://doi.org/10.1007/s10816-018-9364-4

Perreault C (2019) The quality of the archaeological record. University of Chicago Press. https://doi.org/10.7208/chicago/9780226631011.001.0001

R Core Team (2019) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria

Renfrew C (1977) Alternative models for exchange and spatial distribution. In: Earle TK, Ericson JE (eds) Exchange systems in prehistory. Elsevier, pp. 71–90

Reynolds N, Riede F (2019) House of cards: cultural taxonomy and the study of the European Upper Palaeolithic. Antiquity 93:1350–1358. https://doi.org/10.15184/aqy.2019.49

Rezek Z, Holdaway SJ, Olszewski DI, Lin SC, Douglass M, McPherron S et al. (2020) Aggregates, formational emergence, and the focus on practice in stone artifact archaeology. J Archaeol Method Theory. https://doi.org/10.1007/s10816-020-09445-y

Riede F, Hoggard C, Shennan S (2019) Reconciling material cultures in archaeology with genetic data requires robust cultural evolutionary taxonomies. Palgrave Commun 5:55. https://doi.org/10.1057/s41599-019-0260-7

Ross RM, Atkinson QD (2016) Folktale transmission in the Arctic provides evidence for high bandwidth social learning among hunter–gatherer groups. Evol Hum Behav 37:47–53. https://doi.org/10.1016/j.evolhumbehav.2015.08.001

Shennan SJ (1994) Archaeological approaches to cultural identity. Taylor & Francis Group, London, United States

Shennan SJ, Crema ER, Kerig T (2015) Isolation-by-distance, homophily, and "core" vs. "package" cultural evolution models in Neolithic Europe. Evol Hum Behav 36:103–109. https://doi.org/10.1016/j.evolhumbehav.2014.09.006

Sprafke T (2016) Löss in Niederösterreich-Archiv quartärer Klima- und Landschaftsveränderungen. Dissertation. Würzburg, Julius-Maximilians-Universität Würzburg

Sui DZ (2004) Tobler's first law of geography: a big idea for a small world? Ann Assoc Am Geogr 94:269–277. https://doi.org/10.1111/j.1467-8306.2004.09402003.x

Surovell TA, Brantingham PJ (2007) A note on the use of temporal frequency distributions in studies of prehistoric demography. J Archaeol Sci 34:1868–1877. https://doi.org/10.1016/j.jas.2007.01.003

Surovell TA, Byrd Finley J, Smith GM, Brantingham PJ, Kelly R (2009) Correcting temporal frequency distributions for taphonomic bias. J Archaeol Sci 36:1715–1724. https://doi.org/10.1016/j.jas.2009.03.029

Tobler W (1970) A computer movie simulating urban growth in the detroit region. Econ Geogr 46:234. https://doi.org/10.2307/143141

Tobler W (2004) On the first law of geography: a reply. Ann Assoc Am Geogr 94:304–310. https://doi.org/10.1111/j.1467-8306.2004.09402009.x

Vermeersch PM (2019) Radiocarbon Palaeolithic Europe Database, Version 25

Villa P, Roebroeks W (2014) Neandertal demise: an archaeological analysis of the modern human superiority complex. PLoS ONE 9:e96424. https://doi.org/10.1371/journal.pone.0096424

Wetzel CE, Bicudo D, de C, Ector L, Lobo EA, Soininen J, Landeiro VL et al. (2012) Distance decay of similarity in neotropical diatom communities. PLoS ONE 7:1–8. https://doi.org/10.1371/journal.pone.0045071

Wilson MC (1988) Geoarchaeology and archaeological visibility in the northern Mandara Mountains and Mora Plain (Cameroon): preliminary results. In: Barreteau D, Tourneaux H (eds) Le milieu et les hommes. pp. 9–50

Acknowledgements

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to R.I.