Journal of Archaeological Science 177 (2025) 106179

Contents lists available at ScienceDirect

Journal of Archaeological Science

journal homepage: www.elsevier.com/locate/jas

Assessing quantitative methods in archaeology via simulated datasets: The Archaeoriddle challenge. Concept, project and motivations

A. Cortell-Nicolau[a,n,*], S. Carrignon[a,o], I. Rodíguez-Palomo[a], D. Hromada[b], R. Kahlenberg[c], A. Mes[a], D. Priss[d], P. Yaworsky[e,f,g], X. Zhang[h], L. Brainerd[a], J. Lewis[a], D. Redhouse[a], C. Simmons[a], M. Coto-Sarmiento[i,p,q], D. Daems[j], A. Deb[k], D. Lawrence[b], M. O'Brien[l], F. Riede[e,f], X. Rubio-Campillo[m], E. Crema[a]

[a] Department of Archaeology, University of Cambridge, Downing Street, Cambridge, CB2 3DZ, United Kingdom
[b] Fakultät Gestaltung, Universtät der Künste Berlin, Str. des 17, Juni 118, 10587, Berlin, Germany
[c] Department of Archaeology, Durham University, Lower Mount Joy, South Rd, Durham, DH1 3LE, United Kingdom
[d] Department of Geography, Durham University. Lower Mount Joy, South Rd, Durham, DH1 3LE, United Kingdom
[e] Department of Archaeology and Heritage Studies, School of Culture and Society, Aarhus University, Moesgård Allé 20, building 4216, 8270, Højbjerg, Denmark
[f] Center for Ecological Dynamics in a Novel Biosphere, Department of Biology, Aarhus University, Ny Munkegade 114-116, Aarhus C, 8000, Denmark
[g] Section for Molecular Ecology and Evolution, Globe Institute, Copenhagen University, Øster Farimagsgade 5, København K, 1353, Denmark
[h] School of Architecture, Tianjin University, Beiyang Ave, Nankai District, 天津大学第21教学楼 邮政编码, 300072, Tianjin, China
[i] Centre for Urban Networks, School of Culture and Society, Aarhus University, Moesgård Allé 20, Højbjerg, DK-8270, Denmark
[j] Department of Arts, Culture History and Antiquity, Vrije Universiteit Amsterdam, De Boelelaan 1085, 1081 HV, Amsterdam, Netherlands
[k] Department of Cognitive Science, Central European University, Quellenstraße 51, 5th floor, 1100, Wien, Austria
[l] Department of History, Philosophy and Geography, Texas A&M University–San Antonio, San Antonio, TX, 78224, United States
[m] Facultat d'Educació, Universitat de Barcelona, Pg. de la Vall d'Hebron, 171, Horta-Guinardó, 08035, Barcelona, Spain
[n] Max Planck Institute for Evolutionary Anthropology, Deutscher Pl. 6, 04103, Leipzig, Germany
[o] Institute of Archaeology, University College London, Gower st., WC1E 6BT, London, United Kingdom
[p] School of Culture and Society – Social Resilience Lab, Aarhus University, Jens Chr. Skous Vej 7, 4. etage, Aarhus C, 8000, Denmark
[q] HUMANE, Human Ecology and Archaeology Research Group, Department of Archaeology and Anthropology, Milà i Fontanals Institution for Research and Humanities, IMF-CSIC, C. Egipciaques 15, 08001, Barcelona, Spain

ARTICLE INFO

Keywords:
Archaeological methods
Rabbithole
Interaction processes
Tactical simulation
Tournament

ABSTRACT

Compared to what is found in many other scientific disciplines, archaeological data are typically scarce, biased and fragmented. This, coupled with the fact that archaeologists can rarely test their hypotheses using experimental design, makes archaeological inference and our ability to assess the robustness of quantitative methods used to make such inferences challenging.

Archaeoriddle is a project that was born as an attempt to compare archaeological methods in an artificial scenario where the behaviour to be reconstructed was known. In this project we organised an experiment where a virtual archaeological record generated from a simulated interaction between hunter–gatherers and early farmers in a fictional landscape was shared with interested participants. Three archaeological questions were posed and the participants were challenged to answer them with the data that the developer team made available. The model and the generative processes behind the virtual record were known to the developers of the virtual world (*Rabbithole*) but not to the participants. Additionally, players were allowed to sample only a subset of the data from *Rabbithole*, mimicking real-life archaeological research and sampling efforts.

The long-term aim of the project is to assess how different methods performed under a controlled environment since, in this case, we knew the correct answers to the questions posed. This experience provided us with some insights into (1) how efficient various archaeological methods are in answering complex questions; (2) the degree of interest from archaeologists in improving their analytical techniques; and (3) the potential of archaeological method when free from external constraints (e.g. budget, fieldwork, etc.).

---

* Corresponding author. Department of Archaeology, University of Cambridge, Downing Street, Cambridge, CB2 3DZ, United Kingdom.
E-mail address: alfredo_cortell_nicolau@eva.mpg.de (A. Cortell-Nicolau).

https://doi.org/10.1016/j.jas.2025.106179
Received 17 August 2024; Received in revised form 12 February 2025; Accepted 16 February 2025
Available online 22 February 2025
0305-4403/© 2025 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

The last few decades have witnessed a dramatic increase in the quantity, quality, and diversity of analytical methods in archaeology. Although considerable methodological advances were made during the 1960s through the 1980s (Binford and Binford, 1966; Clark, 1982; Clarke, 1968; Thomas, 1978), the enthusiasm associated with those early attempts stalled in the 1990s (Cogwill, 2001; O'Brien et al., 2005; O'Brien and Thomas, 2022; although see Aldenderfer, 1998 for a more optimistic view), perhaps a result, at least in part, of the rise of post-processualism (Earle et al., 1987). Although archaeological science never completely abandoned quantitative and computational approaches, the past 10–15 years, however, have witnessed a renewed interest in them (e.g. Bevan and Lake, 2013; Crema et al., 2014; González-Pérez et al., 2023; Grosman, 2016). Contemporary archaeology is inherently interdisciplinary, thus demanding collaboration and integration with fields such as genomics (Allentoft et al., 2024; Arzelier et al., 2022; Olalde et al., 2019; Posth et al., 2023), chemistry (Dolbunova et al., 2023; Gallello et al., 2021; Wells, 2010), mathematics (Barceló and Bogdanovic, 2015; Fort, 2023; Steele, 2009), statistics (Eren and Buchanan, 2023; Eren et al., 2016; Leonard and Jones, 1989), ecology and environmental studies (Ellis et al., 2021; Hardesty, 1980; Piskin et al., 2018; Rick and Sandweiss, 2020; Xu et al., 2020), biology (Laland and O'Brien, 2010; Murray et al., 2021; Piperno et al., 2017; Prentiss, 2021) and cultural evolutionary studies, especially cultural transmission (e.g. Cochrane and Gardner, 2011; O'Brien, 2008; O'Brien and Lyman, 2000). This interdisciplinary framework often entails borrowing and adapting methods originally developed for different purposes and in different scientific fields (e.g. phylogenetic analysis, network analysis, artificial intelligence) but also the creation of tailored solutions for specific archaeological problems (e.g. seriation, palaeodemography, phase characterisation).

Determining the extent to which these solutions can provide conceptually robust inference and the ability to test scientific hypotheses remains a subject of debate, given that archaeology studies 'unobservable behaviour patterns […] from indirect traces in bad samples' (Clarke, 1973, 17). As a result, most of these methods rely on untested assumptions that can profoundly impact the inferential process. The *Archaeoriddle* competition described in section 3 was designed to address this issue by testing the robustness of different analytical methods on a simulated dataset. Because the behavioural processes generating the artificial record were known, we can determine the extent to which a particular method is able to reconstruct the generative processes behind the observed data.

Archaeoriddle relies on 'research gamification', inspired by previous work in the social and behavioural sciences (Axelrod, 1980; Rendell et al., 2010; but see also Miu et al., 2024 for a similar, more recent approach), where participants searched for solutions to specific questions as they made their way through tournaments. Similar attempts to investigate the robustness of inferential techniques have been carried out in statistical science. For example, Silberzahn and colleagues (2018) compared statistical analyses on the extent of racial bias in football referees by contrasting the analyses carried out by 61 groups of scientists of the same dataset. Breznau et al. (2022) carried out a similar experiment, this time with a contribution of 161 groups of scientists, who examined a dataset on immigration and social policies. These studies demonstrate that different statistical approaches on the same data can lead to divergent conclusions based on prior assumptions, selection of variables and other factors. A similar approach could be carried out on archaeological datasets, but because there are few examples of the same dataset being re-analysed in multiple studies in order to assess independently the robustness of each approach, such an experiment becomes much harder to carry out. An alternative solution, adopted in the Archaeoriddle project, is to carry out similar studies on *simulated* datasets. To our knowledge, with the notable exception of a similar attempt carried out by Henry Harpending during a workshop on prehistoric demography and population genetics in Harpending (2008), our work represents the first attempt to explore the robustness of archaeological methods by means of a combination of simulated dataset and participatory research.

The volume of the material produced during the three-year project duration, with the involvement of several researchers, is substantive. Therefore, we have decided to divide its output into two separate publications. In this first paper we present the project, its concept, motivation and theoretical framework, the outline of the modelling approaches used, the resources it produced (including website, an R package, bookdown project, media content and reproducibility content), the results of participation and a brief comment on the participants' contributions. In the second paper we will present a thorough analysis of the methods proposed by the participants. Here, each proposal will be tested under varying modelling assumptions, data collection protocols and data quality in order to exhaustively assess their performance and robustness.

2. Tactical simulation and archaeology

Archaeoriddle is an example of what some archaeologists refer to as 'tactical simulation'. In contrast to hypothesis testing or heuristic modelling, tactical simulation refers to the process of producing an *in silico* dataset with the objective of assessing either the quality and performance of the method or the data supporting any specific inferential context (Lake, 2014). Often referred to by different names (e.g. toy model, methodological simulation, artificial scenario or benchmark analysis), the practice of evaluating the performance of a particular method using simulated datasets is widely employed outside archaeology (e.g. Benazzo et al., 2017; Bi et al., 2022; Smith et al., 2002). Tactical simulation was introduced into archaeology by Orton (1973) and subsequently used in numerous other studies. For example, Kintigh (1984), and later Rhode (1988), examined simulated data to explore the relation between artefact diversity and sample size. Mithen (1988) studied the relationship between different subsistence strategies and priorities (e.g. maximisation of energy intake, minimisation of foraging time and risk control) and faunal-assemblage diversity by developing and examining the output of an agent-based model. In the subsequent decade, Yorston and colleagues (1990) explored the impact of post-depositional movement and ploughing in spatial analyses; Paine and Harpending (1996) examined the robustness of methods for inferring fertility from age-at-death ratios using simulated osteoarchaeological assemblages; and Varien and Potter (1997) studied the extent to which one can infer the occupational length of sites based on patterns in discarded cooking vessels.

Many of the themes explored at the end of the twentieth century continue to play a prominent role in archaeology. The impacts of taphonomy, sample size and time averaging (Brantingham et al., 2007; Daems et al., 2024; Davies et al., 2016; Premo, 2014; Rubio Campillo et al., 2012; Surovell et al., 2009; Surovell and Brantingham, 2007), for example, continue to be of a vital concern, but perhaps the most prominent feature characterising the use of tactical simulation is the assessment of new techniques, either independently developed or adopted for archaeological applications. Examples include the application of phylogenetic and related methods (Currie et al., 2010; Klopfstein et al., 2017; Nunn et al., 2006), geometric morphometrics (Cortell-Nicolau et al., 2023; Courtenay, 2022; Klingenberg, 2022) and various approaches for inferring cultural transmission (Carrignon et al., 2023; Lipo et al., 1997; Premo, 2014) and mobility (Loog et al., 2017). A research area that has seen particularly fruitful applications of tactical simulation is the analysis of time-frequency data, in particular radiocarbon dates used to infer past population dynamics (Bevan and Crema, 2021; Carleton and Groucutt, 2021; Crema, 2022; Crema and Shoda, 2021; Timpson et al., 2021).

A special subset of tactical simulation is what Buck and Meson (2015) referred to as '*what-if*' experiments. These are often used in the context of Bayesian inference to explore the sensitivity of the output to a specific choice of priors (a form of what is known as 'prior-predictive checks') or the impact of sample size on the precision of parameter estimates (Bayesian power analyses). Again, most archaeological examples of this method involve radiocarbon dates (Bayliss et al., 2007; Bayliss and Bronk Ramsey, 2004; Bayliss and Woodman, 2009; Christen and Buck, 1998; Crema et al., 2022; Crema and Shoda, 2021; Holland-Lulewicz and Ritchison, 2021), although the basic principles are the same for any application involving Bayesian analysis. It is worth noting that, in contrast to most other forms of tactical simulation, the simulated data-generating process and the inferential engine are effectively the same—that is, they use the same probability distributions and parameters. As a result, *what-if* experiments focus on a narrower and more specific range of objectives than other kinds of tactical simulations, where there is no symmetry between the inferential machinery and the data-generating process.

3. Archaeoriddle

Archaeoriddle is a tactical simulation in which the robustness of different methods and approaches are explored through participatory research. The project was divided into two stages. First, the modelling group (see below) developed a complete, easy-to-use open-source framework that combines multiple models to simulate human demographic, ecological and cultural dynamics and generate a realistic archaeological record from these simulated dynamics. Second, we conducted an open competition in which a specific output of the framework was selected, and specific questions were posed to the scientific community to test the abilities of different methods to retrieve the original parameters and processes that generated the selected dynamics.

The framework relies on an agent-based model to create a virtual world called *Rabbithole*, where a simulated population of hunter–gatherers—the *Rabbit-skinners*—interacts with a population of early farmers—the *Poppy-chewers*. The rationale for creating *Rabbithole* was to propose a series of archaeological questions related to an interaction process in the virtual world where participants would submit methodological proposals to answer such questions. The project was first presented at the Computer Applications and Quantitative Methods in Archaeology conference (CAA) at the University of Oxford in 2022, and soon after we started receiving proposals. Early work included the development of a dedicated ShinyApp (https://theia.arch.cam.ac.uk/archaeoriddle) with detailed information about the project and links for downloading the data. Results, along with solutions to the questions posed, were compiled during a workshop at the European Association of Archaeologists (EAA) annual meeting in Belfast in 2023. We selected this format in order to give time to the participants to prepare their proposals and to give non-coders the chance to participate. However, other formats (e.g. hackathon) were considered and are indeed an interesting way forward for potential future expansions of the project.

We generated artificial archaeological data in the form of.

- A sample of radiocarbon dates from multiple site-contexts.
- Cultural affiliation (whether the occupation was by foragers or farmers).
- A raster map representing spatial variation in occupation suitability, i.e. a resource map, based on arbitrary metrics where this occupation suitability was the same for both foragers and farmers (see below).

We employed different subroutines that simulated taphonomic loss to generate our artificial archaeological record. The model (section 4) was implemented as a spatially situated agent-based model, and participants were asked to answer the following questions.

RQ1. What was the relationship between *Rabbit-skinners* (foragers) and *Poppy-chewers* (farmers)? Was it hostile or peaceful?

RQ2. What was the population trajectory of each group?

RQ3. What was the rate of dispersal of the *Poppy-chewers*?

Participants were provided with a set of identical data, including a DEM raster map representing *Rabbithole* and the geographic coordinates of 39 settlements and 120 radiocarbon dates, each matched to one of the two cultural affiliations. Sampling locations were obtained for five out of a total of 100 cells (Fig. 1). Participants were then asked to identify five more cells from which to obtain additional simulated data, providing them an opportunity to devise the most suitable sampling strategy. Given that the full simulation model included outputs of all the behavioural processes, the 'correct' answers to all three questions were known to the developers of the model but not to the participants. This provided an opportunity to compare different methods over the same archaeological context and, at least partially, the same dataset, while simultaneously being able to compare the results against *what really happened*.

4. The Archaeoriddle model

The model is the result of a collaboration among members of the Computational and Digital Archaeology Laboratory (CDAL) at the University of Cambridge. The model and all post-processing functions, were written in R (R Core Team, 2023) and are contained within a stand-alone R package,[1]together with the detailed descriptions of the project in a reproducible book generated by the bookdown R package (Xie, 2016), whose rendered output can be found at https://thearchaeoriddle.org. The simulation was structured in five stages: (1) the virtual landscape and the initial conditions, (2) population growth, dispersal and interaction, (3) formation of the archaeological record, (4) taphonomic loss and (5) generation of the archaeological remains. In order to ensure the reproducibility of this article as it stands here, and due to potential further development of the documentation reported above, all of the material is additionally stored at the zenodo repository: https://zenodo.org/records/14024547. The repository also contains the participants' proposals as they were presented in the tournament. These proposals have also been provided by the participants in individual repositories (section 5)

4.1. *Virtual landscape and initial conditions*

For the first stage, we created a virtual map with an area corresponding approximately to a real-world surface of 305,000 km^2, including a large body of water to explore how this could affect expansion and interaction patterns. The map was generated by creating a raster that represented different elevations using two-dimensional autocorrelated Perlin noise (Perlin, 1985), which is commonly used for the virtual generation of digital elevation models (DEMs). Dimensions of the map and resolution of the DEM raster were selected to create a realistic environment while at the same time keeping reasonable computation times to facilitate reproducibility. The next step consisted of positioning the initial settlements on the map, assigning their cultural identities and establishing their population size and growth dynamics. We started with a population of 30 communities, with hunter–gatherers taking most of the space and farmers confined to the southwestern quadrant. We then generated the resource map defining areas of maximum environmental suitability. Random points were arbitrarily selected as the centres of maximum environmental suitability, or fitness, from which resources decayed exponentially with distance.

4.2. *Population dynamics*

We selected different per-settlement carrying capacities for both groups. The maximum for hunter–gatherers was set at 30 (see Bird et al., 2019) individuals and for farmers at 120 individuals (a conservative

---

[1] https://github.com/acortell3/archaeoriddle/.

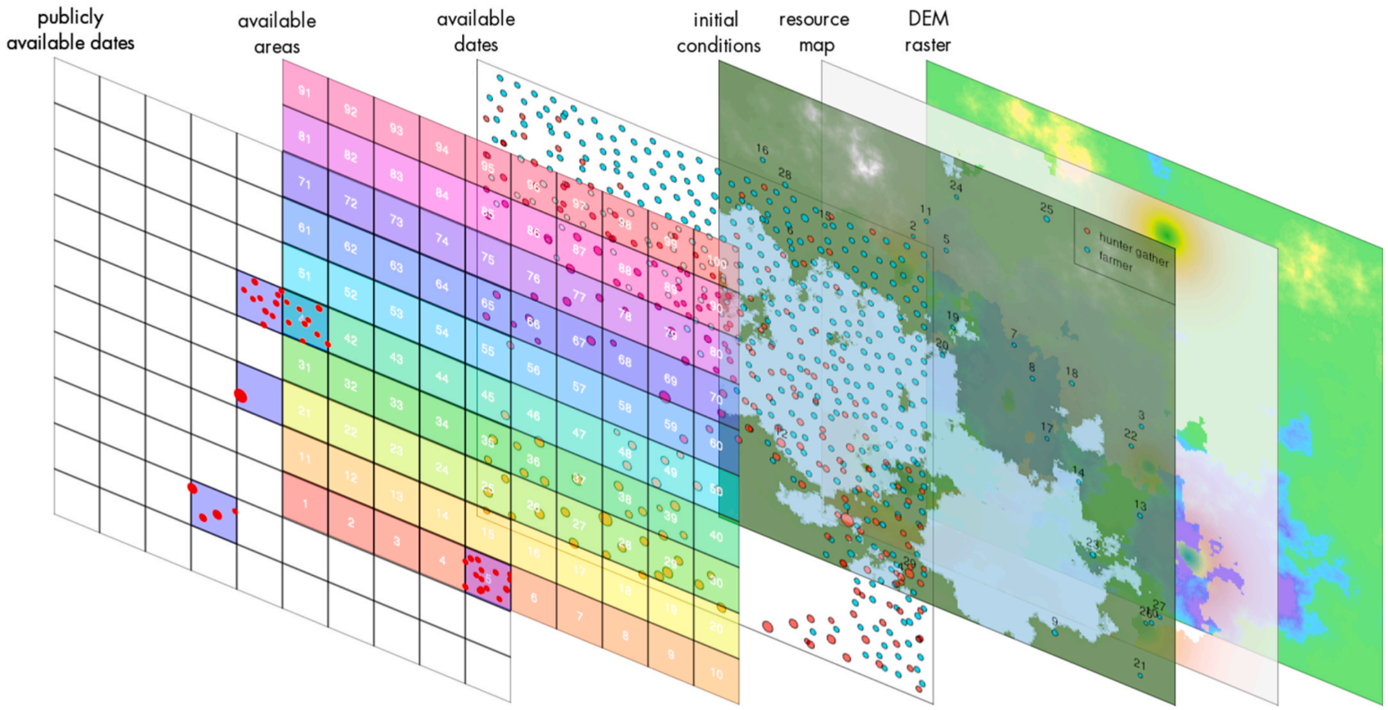

Fig. 1. The different rasters, elevation models and layers that create *Rabbithole*. The *publicly available dates* layer represents the settlements for which dates were provided to all participants. The *available areas* consist of cells that divide the world and that users could select to obtain dates of the settlement located within the cell they chose. All the available dates generated during the simulation are represented in the *available dates* layer with their associated culture, while the *initial conditions* layer shows where the initial sites were located. The *resource map* is a raster that defines how likely new settlements are to be created and will also modulate the carrying capacity of the settlements. Meanwhile, the *DEM Raster* defines the elevation of the world, thus allowing the participants to determine slope, underwater area and higher regions. The two rasters were publicly available to anyone, but the *initial conditions* layer was not.

estimate from Bocquet-Appel, 2011), given that hunter–gatherer settlements tend to be smaller and house fewer people than farmer sites. These culture-dependent baseline carrying capacities were then adjusted in light of resource availability at the site locations, extracted from the resource map. Reaching the carrying capacity would result in increased mortality and/or group fission. Additionally, we created both an initial age-structured population for each settlement and a set of age-specific probabilities that included (1) the probability of giving birth (based on the number of fertile women) within each group, (2) a decrease in population when it exceeded carrying capacity, (3) a probability of settlement fission when the community exceeded its carrying capacity and (4) the probability of joining a new settlement after a fission event.

Recall that the first research question asked participants whether the relationship between farmers and foragers was peaceful or hostile. We opted for the latter, and to achieve this we implemented the following rules. We defined a probability of hostile contact between the two populations that depended on the population size of a settlement and its proximity to a settlement of the opposing group. This meant that larger settlements would have a higher probability of engaging in conflict with an opposing group over broader areas, whereas conflict would be unlikely for very small, isolated settlements or those surrounded by friendly neighbours. In case of conflict, the outcome would be a probability determined by the size of the groups involved (inspired by the linear model described in Rubio-Campillo, 2016) with no major technological differences between the two communities. Although in a real-world environment, there would probably be an aggregation of communities, perhaps based on kinship, in the case of conflict (Turchin, 2013), we did not develop this option in order to keep the model simple. The subroutine is thus stochastically conditioned by demographic processes only (growth and fission to new locations). Given that both carrying capacity and growth rates were higher for farming populations, their likelihood of prevailing in a conflict was assured, as was the extinction of the hunter–gatherers.

4.3. Formation of the archaeological record

Creation of the archaeological record was modelled as an outcome of anthropogenic (humans producing waste) and natural forces (e.g. deposition rates). For the anthropogenic elements, we designed a stratigraphic model in which the archaeological record was deposited at each time step, with its vertical density dependent on natural-deposition rates. We decided not to include per-site variation in post-depositional disturbance. The record generated at that stage consisted of only radiocarbon dates, each associated with a culture affiliation, and a geographic location. The deposition of simulated material culture could be considered in future developments of the model but is not currently implemented. Each group produced a number of non-calibrated radiocarbon dates based on the potentially generated faunal remains. A group would produce a specific amount of faunal waste at each time step t based on population size at that time step, their caloric intake and their economic and social regime (e.g. representativeness of fauna within the diet or amount of food consumed within the settlement). Finally, and considering deposition rates, the faunal waste was deposited at specific times and then covered by layers of sediment (Fig. 2).

4.4. Taphonomic loss

We created an algorithm to emulate the taphonomic loss characteristic of most prehistoric archaeological records. We initially considered using existing models of taphonomic loss developed by Surovell (Surovell et al., 2009; Surovell and Brantingham, 2007; Surovell and Pelton, 2016), but we opted to create our own because Surovell's techniques were applied to bulk assemblages whereas we were modelling the loss of only one specific class of artefacts: animal bones. We considered two types of loss, short-term and long-term. Both are

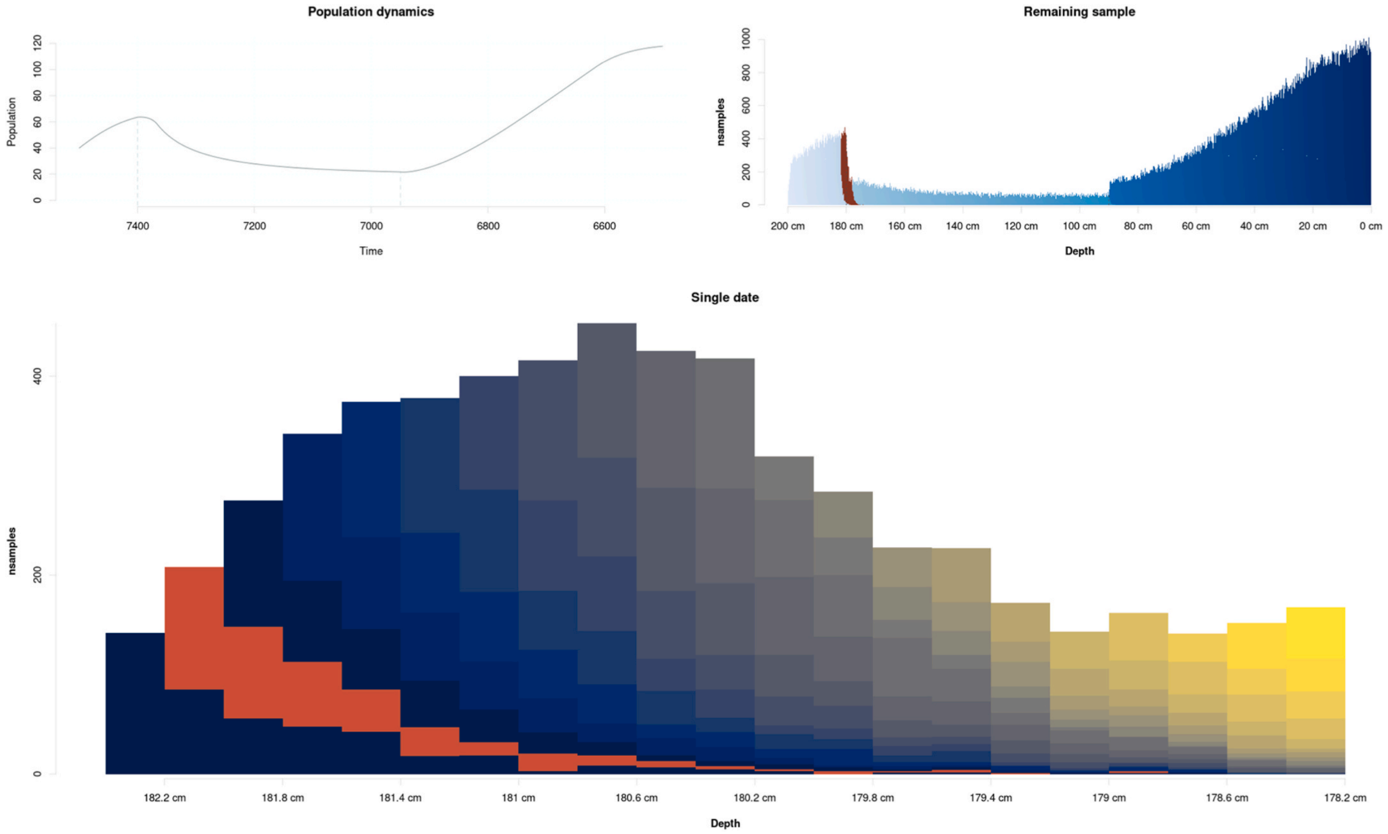

Fig. 2. General simulation of record loss and taphonomic deposition. We consider three phases for this process (1) high population, (2) population crisis and (3) strong population increase, each with different patterns of fauna consumption (parameters in published code). The top-left panel shows the simulated population curve; the top-right panel the archaeological waste produced by that population under different parameters after taphonomic loss, where the colour gradient indicates the different years of deposition, and the bottom panel illustrates a snapshot (highlighting in red the deposition belonging to one specific year) of how that record would be distributed along the sequence as it is buried, according to pre-specified deposition rates. This can be used in further exercises to explore potential post-depositional effects. A detailed explanation regarding this simulation can be found in the bookdown project and the github repository.

modelled as the proportion of the record remaining at time *t* from the record present at *t–1*, where each *t* is one year/one simulation time-step. The difference is that, for the short-term loss, we considered specific high-impact activities that might occur immediately after the initial deposition when the bones are still exposed. These might include cleaning of the area by later groups and scavenging by animals. This implements an increased proportional loss of record, but it acts only one time on the original deposition. Conversely, for the long-term loss, we considered factors such as soil pH and bioturbation in its myriad forms (Wood and Lee johnson, 1978). In this case, the yearly loss of the record is minimal, but the algorithm acts at every *t* on the remaining record from the moment of deposition up until the site is excavated. Thus, it can have a large cumulative impact despite the yearly probability of loss being close to zero.

4.5. Generation of archaeological remains

The final step involved combining the previous four steps. All the different subroutines described above were executed at each time step for a total of 1000 steps, with each step representing a year. The taphonomic loss was then applied from the moment of deposition of a dating element until the present, where, from the moment of abandonment of the site (or end of the simulation for that matter) the only process occurring was taphonomic loss. This produced an archaeological record for each site. Archaeological sites might or might not have existed throughout the complete time span of the simulation because of intersite conflicts but also because of their internal demographic dynamics, which were mediated by their carrying capacity, the proximity to high-sustainability environmental spots and their own demographic stochastic dynamics when populations were low. As a result, we produced a distribution map for each time step, where we logged the age- and sex-structured population matrix for each site at each step, alongside the results of any conflict.

Participants, however, received a single map, or more exactly, data from 10 cells on that single map (five common to all of them and five chosen by the participants), where the existing sites might or might not have existed during the full span of the simulation. To assess this possibility (how long did each site exist), participants had radiocarbon dates, which depended broadly on the population size at each time step, as explained in section 4. It follows that, for equal population, sites with longer time spans would have a higher number of radiocarbon dates, given that each bone assemblage and potential date(s) were generated at each time step. Finally, to stay within a game-like setup and to facilitate discussion, we used a random-name generator for the sites present within the common cells. Participants were encouraged to give their own names to the sites they found in their cells.

5. Proposals and discussion

The five proposals that we received used different methods to answer the research questions: agent-based modelling (ABM), species-distribution modelling, hierarchical Bayesian-phase modelling, point-process modelling (PPM), and a qualitative assessment of the data provided. It is worth noting that there was no overlap in the choice of the methods employed by the participants, hence we were not able to determine the extent to which a particular technique could have been employed differently to answer the same set of questions. It also follows that the exact choices made by the participants do not necessarily reflect the *only* way to use a particular technique to examine the provided dataset and hence our experiment cannot be used to make general claims about the robustness of the techniques explored by the participants.

Before commenting on the proposals, we list the answers to the three research questions.

RQ1. The *Poppy-chewers* and *Rabbit-skinners* had a hostile relationship.

RQ2. *Poppy-chewers* followed an exponential population growth, whereas the *Rabbit-skinners* eventually stalled and then decreased as a function of the other group's increase (Fig. 3).

RQ3. The rate of dispersal varied across the landscape, with an average of ~0.62 km/year, faster for the sea crossing (~1.11 km/year) and slowest (~0.57 km/year) in the northeastern quadrant.

5.1. *The proposals*

Although an exhaustive presentation of the methods proposed and how they behave under different assumptions of data quality will be the focus of the upcoming paper, we briefly describe them here in no particular order. Summaries were provided by the participants, along with links to the proposals.

Priss and Kahlenberg's proposal (https://doi.org/10.5281/zenodo.14062675) was based on ABM combined with exploratory data analysis. It first studied the land of *Rabbithole* and calibrated the dates provided, which were then used to infer trajectories of dispersal and study site preference using ArcGIS pro and R. The suggested dispersal rates (120 km/100a) over land and water were close to the Archaeoriddle solution. The proposal then used the results of its analyses to fit the ABM, which was built using NetLogo. For the ABM, moving groups of hunter–gatherers and farmers, as well as the different settlements, were treated as agents, with starting values obtained from the exploratory data analysis and relevant literature. Additionally, it introduced behavioural rules, including movement and site preference, reaction to population threshold or reaction to interaction. After running the model several times with different parameters, it correctly predicted a hostile relation between the two groups and inferred an initial location of the *Rabbit-skinners* in the northeastern quadrant and the *Poppy-chewers* in the southwestern quadrant, while also detecting the latter's northward movement. The expansion rate was not captured by the initial assumption of logistic growth in the areas already densely inhabited by *Poppy-chewers*, which makes sense considering the different population trajectories of the two groups.

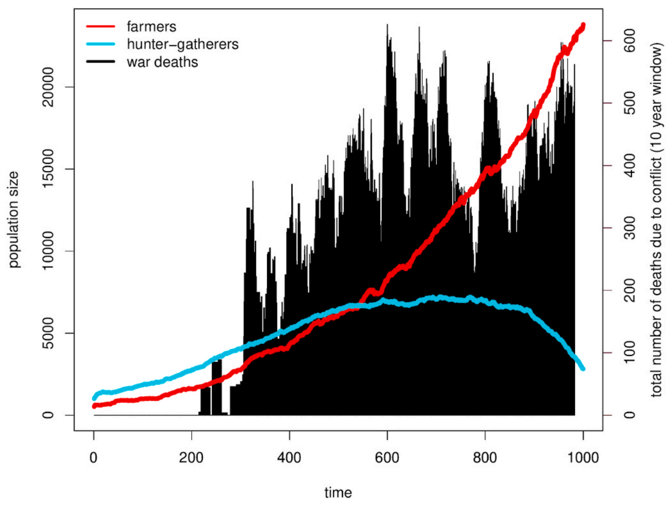

Fig. 3. Population trajectories of the hunter–gatherers (represented in blue) and the farmers (represented in red) are depicted alongside black bars, which represent the summed deaths per conflict over a 10-year moving window. Initially, farmers constitute a minority, and for the first 200 years, no conflicts arise as the two groups flourish separately. However, after this period, groups of migrants begin to settle in closer proximity to one another, leading to an increase in deaths due to conflict, while both populations continue to grow. Over time, the farmers, with their higher growth rates and their ability to sustain larger settlements, gradually begin to supplant the hunter–gatherers, whose population starts to decline.

Zhang's point process model (https://doi.org/10.5281/zenodo.12803445) focused initially on the sampling process to build first-order models combining fitness and the available cells to predict potential occupation. It then focused on the question of whether there was conflict between the two groups. After building different archaeological phases, it computed the clustering patterns of the groups under the assumption that higher clustering could lead to higher conflict (Field, 2004). Following this, it computed the interaction distance between groups of settlements of hunter–gatherers and farmers through a multitype Strauss model. According to the results, hostilities increased over time. In the original model, the rules for hostility were not time dependent, but we have seen that, even if the rules for conflict did not change, as the population grew and the number of settlements increased, there was more probability of contact and thus more probability of conflict, which finally resulted in an increased mortality.

Yaworsky (https://doi.org/10.5281/zenodo.8260754) used species-distribution modelling in R to develop a four-stage research design. The first stage focused on determining which additional cells would result in a representative sample of the range variables (elevation and resource quality) present within the data. The second stage focused on data exploration to identify internal temporal, spatial and farmer and forager patterns. In the third stage, it generated summed probability distributions from the calibrated radiocarbon dates to generate relative estimates of population size through time for foragers and farmers. The fourth stage combined these data into a spatiotemporal species-distribution model, where both time and space were explicit predictors used to estimate the distribution of farmers and foragers in 100-year intervals. The spatiotemporal species distribution was successful in reproducing the directionality of the farming dispersal (from south to north) as well as the decline in hunter–gatherer populations.

Mes (https://doi.org/10.5281/zenodo.14218979) first developed a strategy to aid in the selection of additional data using a friction calculation that considered the distance from a putative origin region, the elevation of the region and its environmental suitability. Using R, analysis focused on capturing local complexity in the dispersal of *Poppy-chewers* in the study area. To track this, it used a hierarchical Bayesian phase model that was informed by all the selected settlements, both the ones held in common with other proposals and the ones obtained after calculations for additional sampling. This method allowed uncertainty to be introduced. It divided *Rabbithole* into 25 subareas and calculated the times of arrival of the *Poppy-chewers* for each area, including their high-probability density intervals, resulting in a successful approach to expansion rates.

Hromada (http://doi.org/10.13140/RG.2.2.10753.47207) applied what he described as the NALANA method—a "naive, laic, and narrative" approach—to address the question: Was the relationship between farmers and gatherers peaceful or hostile? Hromada's approach combined logical reasoning with more data-driven analysis proposed by GPT-4 (https://doi.org/10.5281/zenodo.14207474). A key observation has been an abrupt replacement of hunter–gatherer settlements by farming communities, suggesting conflict rather than peaceful coexistence. The analysis led to the formulation of the "Persistence Disruption Conjecture," which posits that in contexts where cultural territories overlap, a significant reduction in the settlement persistence of one group often signals dominance by the other. While the broader applicability of this conjecture has yet to be verified or falsified by exposure to real archaeological data, Hromada's NALANA approach explores the value of blending computational methods proposed by generative AI with more human narrative hermeneutics and logical reasoning.

5.2. Joint consideration (by all authors)

As mentioned before, although a detailed analysis of the methods performance, their robustness and the questions addressed will be further detailed in the upcoming paper paired to this one, a brief comment here is due.

One noticeable aspect of the modelling was the effort that went into selecting the additional cells, where participants used different approaches to select the best cells to sample. Traditionally, archaeology has relied on the recovery of samples that were found by chance, either through fortuitous discoveries, for example, through information from local individuals or construction projects (Friman and Lagerås, 2023; Takata and Yanase, 2021) or via carefully designed field surveys (Bintliff, 2023; Cegielski et al., 2023; Mueller, 1974; Schiffer et al., 1978). Surveys, however, are often limited by a series of constraints (Bevan and Conolly, 2004) —ground cover being a major one—which *Rabbithole* was free of. Since these were not an issue for participants, they were able to develop thoughtful sampling methods without actual physical constraints. We did not consider differential preservation biases when setting up the model, which means that the taphonomic loss rate had the same probability for every site regardless of location. Finally, we did not include the effects of modern land use on the archaeological record of *Rabbithole*.

As shown in Table 1, not all five proposals addressed each of the three research questions, although most attempted to answer the first one: Was the relationship between the hunter–gatherers and farmers hostile or peaceful? This is interesting because unless there is clear evidence of violence among groups of farmers (Alt et al., 2020; Meyer et al., 2015) or between farmers and hunter–gatherers (Roksandic et al., 2006)—for example, bone trauma or the presence of embedded projectile points (information that was not provided by the Archaeoriddle sample)—archaeologists are purposely cautious when inferring conflict. There are, however, proxies that can be used to infer intergroup violence. For example, Priss and Kahlenberg successfully developed their own ABM considering potential conflict by including their own behavioural rules, whereas Zhang's analysis was based on the assumption that higher clustering would lead to greater conflict, and Hromada used a qualitative approach based on assumptions about settlement duration. As an important aside, participants were able to openly propose their findings because they were not writing history but rather playing a game, and they knew (1) that they would eventually know the actual processes involved and (2) that the negative consequences of being wrong were greatly minimised, having nothing of their own research at stake. Thus, participants were free to take more risks in their analyses.

Another area in which participants were successful was the directionality of population expansion (despite the fact that this was not one of the RQs), although accurate rates of spread were particularly challenging to estimate. It is true that, in this regard, if we compare the ratio of data available against the total geographic frame, the participants had fewer data than what are common in archaeological practice, since the cells provided covered only minimal parts of *Rabbithole*. This might have hindered the exploration of this research question.

A puzzling aspect of the Archaeoriddle project was the small number of proposals that were submitted, despite a cash prize (although significantly less than those offered in similar challenges from other disciplines). This was unexpected, given the initial interest expressed by numerous colleagues, the amount of advertising that was done and the fact that our questions were related to Kintigh et al.'s (2014) *grand challenges* (more in particular, our questions and their consequences could be related to Kintigh et al.'s A4, A7, B1, B2, B3, C2, E2 and E5). To assess why participation was low, we surveyed a number of archaeologists who regularly contribute to methodological development. We received 31 answers (all answers are available in SI1 and a minimal visualisation is provided here: https://thearchaeoriddle.org/survey_analysis.html). Fourteen addressed why they did not participate, of which one did not specify, one did not understand the purpose of the project, one did not know about the project, two did participate and nine stated that, despite finding the project interesting, they did not have the time, which couples with the significant amount of work that each active participant put in developing their answer. As for the reasons that brought researchers to at least consider participating, we obtained 15 responses. Thirteen mentioned as their main motivation the scientific interest of the project, whereas one researcher was interested in the financial prize and another thought that it was a good and fast way to improve their curriculum vitae. Two other aspects of the challenge are worth noting. One is the strong gender bias in the responses (12 men and three women) and the other is career stage, where, of those who did include this information, six respondents were at a faculty level, four at the postdoctoral level and five at predoctoral level, showing essentially an equal research interest in the matter for any career stage. In terms of participants, however, one was faculty level, another one postdoctoral level and the rest were all PhD students.

These responses caused us to carefully consider factors that might complicate any project of this nature. First, it was essentially computational archaeologists who engaged with the project and, in this regard, the critical mass of computational archaeologists is probably still not large enough to support an endeavour such as Archaeoriddle. Indeed, and although the current increase on the use of computational methods by archaeologists is out of discussion (e.g. reproducibility advisors, open-data policies, CAA conferences, etc.), archaeology retains overall a rather qualitative approach. How can we combine quantitative formal methods with a solid theoretical understanding of the discipline and its specific problems in terms of data, research questions and interpretation is one of the main upcoming theoretical challenges for the future of the discipline. Assessing why the project did not extend to the wider archaeological community remains speculative, but it could be related either to the developers' own network, theoretical stands or simply pure interest. A proper answer to this question exceeds the scope of the Archaeoriddle project. Adding to this, and despite the best of intentions, the project was not the prime interest of possible entrants and was placed on the back burner. On the positive side, *Rabbithole* remains an open resource for the archaeological community, and further exploration of different questions might continue in the future with different challenges, thus increasing the body of literature supporting these synthetic experiments.

Table 1
Summary of participants' proposals, based on survey SI2 sent to all participants. ABM refers to Agent-Based Modelling, PPM to Point-Process Modelling and SDM to Spatial Distribution Modelling.

Proposal	Question answered			Approach	Different than authors' usual methods	Main programming language
	RQ1	RQ2	RQ3			
Priss & Kahlenberg	partially	Yes	Yes	ABM	No	NetLogo
Hromada	No	No	Yes	qualitative	Yes	Natural language
Zhang	Yes	No	No	PPM	No	R
Yaworsky	Yes	Yes	Yes	SDM & others	No	R
Mes	No	No	Yes	Bayesian hierarchical phase modelling	No	R

The range of methods used and the details provided illustrate how archaeologists are aware of the necessity to improve their inferential tools. In this regard, it is interesting to see that the answers to the question "which inferential methods do you use more frequently in your own work?" (in the survey SI1) spanned a variety of potential responses, most of which do not relate to the methods used by the participants. This reinforces the argument that quantitative and computational methods in archaeology are currently far from being standardised. This is particularly relevant when compared to the fact that most of the participants (except one not from the field of archaeology) chose the methods they used in the challenge on the basis of what they were most familiar with (Table 1). It is difficult to say at this stage whether choosing known methods was due to the nature of this project (e.g. limited time) or if it is a more general tendency within the archaeological discipline, but in any case it certainly demands a reflection on how archaeologists choose the most proper methods to answer their research questions. We hope that Archaeoriddle is a first step in that direction and the subsequent paper builds on this framework to allow further exploration of these problems.

6. Conclusion and future work

Archaeoriddle provided a playful yet serious opportunity to test different approaches to datasets generated through a process with a known outcome. This is a valuable tool for archaeologists, who can usually only indirectly assess the validity of their proposals. Moreover, the quality of the proposals received shows a commitment to improving archaeological methods and a strong willingness to push toward transparency, openness and reproducibility. Nonetheless, the low number of participants also indicates that there is still a substantial amount of work to do in terms of increasing methodological awareness within the archaeological community. Regardless of the complexity of the model and questions posed, the experimental design for this specific use of tactical simulation is simple: One researcher/group develops a model and research questions, and another researcher/group, unaware of the actual parameters, design or rationale of the model, develops the methods to address the questions.

Toy models are considered as vital tools for the generation of novel ideas (Marzuoli, 2008). The Archaeoriddle framework, if extended, can exponentially increase the power of archaeological inference as more questions, research topics and datasets are brought into play. Indeed, it provides high levels of flexibility for the design of complex research questions while also bringing, for the participants, an environment in which to test their methods and check results against real outcomes. The choice of using an open-source R package ensures that anyone interested in developing different and more-complex versions of the model can do so. New versions might include algorithms for differential taphonomic loss patterns, implementing different ways of assessing cultural drift on material culture or analysing complex socio-economic interaction among settlements. The data used during the original competition have already been shared and can be used as an educational tool in a similar fashion to other projects such as AtlantGIS (https://github.com/kacebe/AtlantGIS). Moreover, the complete process to develop new *'Rabbitholes'* to explore different questions and test different methods, has been shared online as an interactive book produced with the bookdown R package (Xie, 2016), which is fully replicable and modifiable. Thus, interested researchers can develop their own environment and replicate or modify it according to their specific needs. In the original challenge described here, we gave only partial data to the participants. However, to assess how the poor quality of archaeological data (Perreault, 2019) affects actual archaeological inference, one could gradually increase the amount of data available, which we will cover in the second paper.

Finally, Archaeoriddle can be used as a teaching resource, enabling students to understand the specific ways in which archaeological data is generated and understood, as well as encouraging not only experimentation and computational method development, but also broader methodological and theoretical debates not necessarily confined to the computational community. This is the presentation and first publication of a project that intends to be collaborative, and we invite any interested researchers to engage with us and to propose additional solutions, methods and modifications. Our goal is to provide a collaborative tool that can help the archaeological community better assess archaeological data and methods, teach these methods to the broadest audience and allow archaeologists to propose better inferential frameworks.

CRediT authorship contribution statement

A. Cortell-Nicolau: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **S. Carrignon:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **I. Rodíguez-Palomo:** Validation, Software, Resources, Formal analysis, Data curation. **D. Hromada:** Writing – review & editing, Resources, Methodology, Investigation, Formal analysis. **R. Kahlenberg:** Writing – review & editing, Resources, Methodology, Investigation, Formal analysis. **A. Mes:** Writing – review & editing, Resources, Methodology, Investigation, Formal analysis. **D. Priss:** Writing – review & editing, Resources, Methodology, Funding acquisition, Formal analysis. **P. Yaworsky:** Writing – review & editing, Resources, Methodology, Investigation, Formal analysis. **X. Zhang:** Writing – review & editing, Resources, Methodology, Investigation, Formal analysis. **L. Brainerd:** Resources. **J. Lewis:** Writing – review & editing, Resources. **D. Redhouse:** Software, Resources. **C. Simmons:** Resources. **M. Coto-Sarmiento:** Writing – review & editing. **D. Daems:** Writing – review & editing. **A. Deb:** Writing – review & editing, Resources. **D. Lawrence:** Writing – review & editing. **M. O'Brien:** Writing – review & editing, Writing – original draft. **F. Riede:** Writing – review & editing. **X. Rubio-Campillo:** Resources, Conceptualization. **E. Crema:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Investigation, Formal analysis, Conceptualization.

Data availability statement

The data that support the findings of this study are openly available at https://github.com/acortell3/archaeoriddle and https://doi.org/10.5281/zenodo.14024547. The data has been shared without restrictions and we have not used third party data.

Declaration of competing interest

The authors declare that they have no competing interest.

Acknowledgments

This work has been supported by the following projects: MSCA-IF ArchBiMod project H-2020-MSCA-IF-2020 actions (Grant no. 101020631), funded by the European Commission and SRG2223\230262 BA/Leverhulme Small Research Grants, funded by the British Academy.

We also want to thank everyone who has helped us develop this project, including all the videos and media content. Thanks so much to (in alphabetical order) Miranda Evans, Elena Grau-Espinosa, Chris Stevens, Carmen Ting, Vaneshree Vidyarthi, Jasmine Vieri, Camilla Zeviani and Xuspi (the dire wolf).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jas.2025.106179.

Reproducibility report

The Associate Editor for Reproducibility downloaded all materials and could reproduce the results presented by the authors.

References

Aldenderfer, M., 1998. Quantitative methods in archaeology: a review of recent trends and developments. J. Archaeol. Res. 6, 91–120. https://doi.org/10.1023/A:1022893621306.

Allentoft, M.E., Sikora, M., Fischer, A., Sjögren, K.-G., Ingason, A., Macleod, R., Rosengren, A., Schulz Paulsson, B., Jørkov, M.L.S., Novosolov, M., Stenderup, J., Price, T.D., Fischer Mortensen, M., Nielsen, A.B., Ulfeldt Hede, M., Sørensen, L., Nielsen, P.O., Rasmussen, P., Jensen, T.Z.T., Refoyo-Martínez, A., Irving-Pease, E.K., Barrie, W., Pearson, A., Sousa da Mota, B., Demeter, F., Henriksen, R.A., Vimala, T., McColl, H., Vaughn, A., Vinner, L., Renaud, G., Stern, A., Johannsen, N.N., Ramsøe, A.D., Schork, A.J., Ruter, A., Gotfredsen, A.B., Henning Nielsen, B., Brinch Petersen, E., Kannegaard, E., Hansen, J., Buck Pedersen, K., Pedersen, L., Klassen, L., Meldgaard, M., Johansen, M., Uldum, O.C., Lotz, P., Lysdahl, P., Bangsgaard, P., Petersen, P.V., Maring, R., Iversen, R., Wåhlin, S., Anker Sørensen, S., Andersen, S. H., Jørgensen, T., Lynnerup, N., Lawson, D.J., Rasmussen, S., Korneliussen, T.S., Kjær, K.H., Durbin, R., Nielsen, R., Delaneau, O., Werge, T., Kristiansen, K., Willerslev, E., 2024. 100 ancient genomes show repeated population turnovers in Neolithic Denmark. Nature 625, 329–337. https://doi.org/10.1038/s41586-023-06862-3.

Alt, K.W., Tejedor Rodríguez, C., Nicklisch, N., Roth, D., Szécsényi Nagy, A., Knipper, C., Lindauer, S., Held, P., de Lagrán, Í.G.M., Schulz, G., Schuerch, T., Thieringer, F., Brantner, P., Brandt, G., Israel, N., Arcusa Magallón, H., Meyer, C., Mende, B.G., Enzmann, F., Dresely, V., Ramsthaler, F., Guillén, J.I.R., Scheurer, E., López Montalvo, E., Garrido Pena, R., Pichler, S.L., Guerra, M.A.R., 2020. A massacre of early Neolithic farmers in the high Pyrenees at Els Trocs, Spain. Sci. Rep. 10, 2131. https://doi.org/10.1038/s41598-020-58483-9.

Arzelier, A., Rivollat, M., De Belvalet, H., Pemonge, M.-H., Binder, D., Convertini, F., Duday, H., Gandelin, M., Guilaine, J., Haak, W., Deguilloux, M.-F., Pruvost, M., 2022. Neolithic genomic data from southern France showcase intensified interactions with hunter–gatherer communities. iScience 25, 105387. https://doi.org/10.1016/j.isci.2022.105387.

Axelrod, R., 1980. Effective choice in the prisoner's dilemma. J. Conflict Resolut. 24, 3–25.

Barceló, J.A., Bogdanovic, I., 2015. Mathematics and Archaeology. CRC Press, Boca Raton.

Bayliss, A., Bronk Ramsey, C., 2004. Pragmatic bayesians: a decade of integrating radiocarbon dates into chronological models. In: Buck, C.E., Millard, A.R. (Eds.), Tools for Constructing Chronologies: Crossing Disciplinary Boundaries, Lecture Notes in Statistics. Springer, London, pp. 25–41. https://doi.org/10.1007/978-1-4471-0231-1_2.

Bayliss, A., Ramsey, C.B., Plicht, J. van der, Whittle, A., 2007. Bradshaw and bayes: towards a timetable for the neolithic. Camb. Archaeol. J. 17, 1–28. https://doi.org/10.1017/S0959774307000145.

Bayliss, A., Woodman, P., 2009. A new bayesian chronology for mesolithic occupation at mount sandel, northern Ireland. Proc. Prehist. Soc. 75, 101–123. https://doi.org/10.1017/S0079497X00000311.

Benazzo, A., Trucchi, E., Cahill, J.A., Maisano Delser, P., Mona, S., Fumagalli, M., Bunnefeld, L., Cornetti, L., Ghirotto, S., Girardi, M., Ometto, L., Panziera, A., Rota-Stabelli, O., Zanetti, E., Karamanlidis, A., Groff, C., Paule, L., Gentile, L., Vilà, C., Vicario, S., Boitani, L., Orlando, L., Fuselli, S., Vernesi, C., Shapiro, B., Ciucci, P., Bertorelle, G., 2017. Survival and divergence in a small group: the extraordinary genomic history of the endangered Apennine brown bear stragglers. Proc. Natl. Acad. Sci. 114, E9589–E9597. https://doi.org/10.1073/pnas.1707279114.

Bevan, A., Conolly, J., 2004. GIS, archaeological survey, and landscape archaeology on the Island of Kythera, Greece. J. Field Archaeol. 29, 123–138. https://doi.org/10.1179/jfa.2004.29.1-2.123.

Bevan, A., Crema, E.R., 2021. Modifiable reporting unit problems and time series of long-term human activity. Philos. Trans. R. Soc. B Biol. Sci. 376, 20190726. https://doi.org/10.1098/rstb.2019.0726.

Bevan, A., Lake, M.W. (Eds.), 2013. Computation Approaches to Archaeological Spaces. Routledge, London.

Bi, R., Zhang, Y., Jiang, X., Yang, H., Yan, K., Han, M., Li, W., Zhong, H., Tan, X., Xia, L., Sun, X., Xiang, S., 2022. Simulation and techno-economical analysis on the pyrolysis process of waste tire. Energy 260, 125039. https://doi.org/10.1016/j.energy.2022.125039.

Binford, L.R., Binford, S.R., 1966. A preliminary analysis of functional variability in the Mousterian of Levallois facies. Am. Anthropol. 69, 238–295.

Bintliff, J., 2023. Agricultural intensification and the evidence from offsite survey archaeology. J. World PreHistory 36, 109–145. https://doi.org/10.1007/s10963-023-09176-4.

Bird, D.W., Bird, R.B., Codding, B.F., Zeanah, D.W., 2019. Variability in the organization and size of hunter–gatherer groups: foragers do not live in small-scale societies. J. Hum. Evol. 131, 96–108. https://doi.org/10.1016/j.jhevol.2019.03.005.

Bocquet-Appel, J.-P., 2011. When the world's population took off: the springboard of the neolithic demographic transition. Science 333, 560–561. https://doi.org/10.1126/science.1208880.

Brantingham, P.J., Surovell, T.A., Waguespack, N.M., 2007. Modeling post-depositional mixing of archaeological deposits. J. Anthr. Archaeol. 26, 517–540.

Breznau, N., Rinke, E.M., Wuttke, A., Nguyen, H.H.V., Adem, M., Adriaans, J., Alvarez-Benjumea, A., Andersen, H.K., Auer, D., Azevedo, F., Bahnsen, O., Balzer, D., Bauer, G., Bauer, P.C., Baumann, M., Baute, S., Benoit, V., Bernauer, J., Berning, C., Berthold, A., Bethke, F.S., Biegert, T., Blinzler, K., Blumenberg, J.N., Bobzien, L., Bohman, A., Bol, T., Bostic, A., Brzozowska, Z., Burgdorf, K., Burger, K., Busch, K.B., Carlos-Castillo, J., Chan, N., Christmann, P., Connelly, R., Czymara, C.S., Damian, E., Ecker, A., Edelmann, A., Eger, M.A., Ellerbrock, S., Forke, A., Forster, A., Gaasendam, C., Gavras, K., Gayle, V., Gessler, T., Gnambs, T., Godefroidt, A., Grömping, M., Groß, M., Gruber, S., Gummer, T., Hadjar, A., Heisig, J.P., Hellmeier, S., Heyne, S., Hirsch, M., Hjerm, M., Hochman, O., Hövermann, A., Hunger, S., Hunkler, C., Huth, N., Ignácz, Z.S., Jacobs, L., Jacobsen, J., Jaeger, B., Jungkunz, S., Jungmann, N., Kauff, M., Kleinert, M., Klinger, J., Kolb, J.-P., Kołczyńska, M., Kuk, J., Kunißen, K., Kurti Sinatra, D., Langenkamp, A., Lersch, P.M., Löbel, L.-M., Lutscher, P., Mader, M., Madia, J.E., Malancu, N., Maldonado, L., Marahrens, H., Martin, N., Martinez, P., Mayerl, J., Mayorga, O.J., McManus, P., McWagner, K., Meeusen, C., Meierrieks, D., Mellon, J., Merhout, F., Merk, S., Meyer, D., Micheli, L., Mijs, J., Moya, C., Neunhoeffer, M., Nüst, D., Nygård, O., Ochsenfeld, F., Otte, G., Pechenkina, A.O., Prosser, C., Raes, L., Ralston, K., Ramos, M.R., Roets, A., Rogers, J., Ropers, G., Samuel, R., Sand, G., Schachter, A., Schaeffer, M., Schieferdecker, D., Schlueter, E., Schmidt, R., Schmidt, K.M., Schmidt-Catran, A., Schmiedeberg, C., Schneider, J., Schoonvelde, M., Schulte-Cloos, J., Schumann, S., Schunck, R., Schupp, J., Seuring, J., Silber, H., Sleegers, W., Sonntag, N., Staudt, A., Steiber, N., Steiner, N., Sternberg, S., Stiers, D., Stojmenovska, D., Storz, N., Striessnig, E., Stroppe, A.-K., Teltemann, J., Tibajev, A., Tung, B., Vagni, G., Van Assche, J., van der Linden, M., van der Noll, J., Van Hootegem, A., Vogtenhuber, S., Voicu, B., Wagemans, F., Wehl, N., Werner, H., Wiernik, B.M., Winter, F., Wolf, C., Yamada, Y., Zhang, N., Ziller, C., Zins, S., Żółtak, T., 2022. Observing many researchers using the same data and hypothesis reveals a hidden universe of uncertainty. Proc. Natl. Acad. Sci. 119, e2203150119. https://doi.org/10.1073/pnas.2203150119.

Buck, C.E., Meson, B., 2015. On being a good Bayesian. World Archaeol. 47, 567–584. https://doi.org/10.1080/00438243.2015.1053977.

Carleton, W.C., Groucutt, H.S., 2021. Sum things are not what they seem: problems with point-wise interpretations and quantitative analyses of proxies based on aggregated radiocarbon dates. Holocene 31, 630–643. https://doi.org/10.1177/0959683620981700.

Carrignon, S., Alexander Bentley, R., O'Brien, M.J., 2023. Estimating two key dimensions of cultural transmission from archaeological data. J. Anthropol. Archaeol. 72, 101545. https://doi.org/10.1016/j.jaa.2023.101545.

Cegielski, W., Snitker, G., Barton, C.M., Bernabeu Aubán, J., Cortell-Nicolau, A., Pardo-Gordó, S., Bergin, S., Castillo, A.D., 2023. A multi-method approach with machine learning to evaluating the distribution and intensity of prehistoric land use in Eastern Iberia. Quat. Int. https://doi.org/10.1016/j.quaint.2023.08.011.

Christen, J.A., Buck, C.E., 1998. Sample selection in radiocarbon dating. J. R. Stat. Soc. Ser. C Appl. Stat. 47, 543–557. https://doi.org/10.1111/1467-9876.00127.

Clark, G.A., 1982. Quantifying archaeological research. In: Schiffer, M.B. (Ed.), Advances in Archaeological Method and Theory. Academic Press, San Diego, pp. 217–273. https://doi.org/10.1016/B978-0-12-003105-4.50011-X.

Clarke, D., 1973. Archaeology: the loss of innocence. Antiquity 47, 6–18. https://doi.org/10.1017/S0003598X0003461X.

Clarke, D., 1968. Analytical Archaeology. Methuen, London.

Cochrane, E.E., Gardner, A. (Eds.), 2011. Evolutionary and Interpretative Archaeologies: A Dialogue, UCL Institute of Archaeology Publications. Routledge.

Cogwill, G.L., 2001. Past, present and future of quantitative methods in United States archaeology. In: Stančič, Z., Veljanovski, T. (Eds.), Computing Archaeology for Understanding the Past. CAA 2000. Computer Applications and Quantitative Methods in Archaeology. Proceedings of the 28th Conference. Archaeopress, Oxford, pp. 35–40. Ljubljana, April 2000.

Cortell-Nicolau, A., García-Puchol, O., Juan-Cabanilles, J., 2023. The geometric microliths of cueva de la cocina and their significance in the mesolithic of Eastern Iberia: A morphometric study. Quat. Int. https://doi.org/10.1016/j.quaint.2023.05.014.

Courtenay, L.A., 2022. Can we restore balance to geometric morphometrics? A theoretical evaluation of how sample imbalance conditions ordination and classification. Evol. Biol. https://doi.org/10.1007/s11692-022-09590-0.

Crema, E.R., 2022. Statistical inference of prehistoric demography from frequency distributions of radiocarbon dates: a review and a guide for the perplexed. J. Archaeol. Method Theor 29, 1387–1418. https://doi.org/10.1007/s10816-022-09559-5.

Crema, E.R., Edinborough, K., Kerig, T., Shennan, S.J., 2014. An Approximate Bayesian Computation approach for inferring patterns of cultural evolutionary change. J. Archaeol. Sci. 50, 160–170. https://doi.org/10.1016/j.jas.2014.07.014.

Crema, E.R., Shoda, S., 2021. A Bayesian approach for fitting and comparing demographic growth models of radiocarbon dates: a case study on the Jomon-Yayoi transition in Kyushu (Japan). PLoS One 16, e0251695. https://doi.org/10.1371/journal.pone.0251695.

Crema, E.R., Stevens, C.J., Shoda, S., 2022. Bayesian analyses of direct radiocarbon dates reveal geographic variations in the rate of rice farming dispersal in prehistoric Japan. Sci. Adv. 8, eadc9171. https://doi.org/10.1126/sciadv.adc9171.

Currie, T.E., Greenhill, S.J., Mace, R., 2010. Is horizontal transmission really a problem for phylogenetic comparative methods? A simulation study using continuous cultural traits. Philos. Trans. R. Soc. B Biol. Sci. 365, 3903–3912. https://doi.org/10.1098/rstb.2010.0014.

Daems, D., Coco, E., Gillreath-Brown, A., Kafetzaki, D., 2024. The effects of time-averaging on archaeological networks. J. Archaeol. Method Theor 31, 473–506. https://doi.org/10.1007/s10816-023-09608-7.

Davies, B., Holdaway, S.J., Fanning, P.C., 2016. Modelling the palimpsest: an exploratory agent-based model of surface archaeological deposit formation in a fluvial arid Australian landscape. Holocene 26, 450–463. https://doi.org/10.1177/0959683615609754.

Dolbunova, E., Lucquin, A., McLaughlin, T.R., Bondetti, M., Courel, B., Oras, E., Piezonka, H., Robson, H.K., Talbot, H., Adamczak, K., Andreev, K., Asheichyk, V., Charniauski, M., Czekaj-Zastawny, A., Ezepenko, I., Grechkina, T., Gunnarssone, A., Gusentsova, T.M., Haskevych, D., Ivanischeva, M., Kabaciński, J., Karmanov, V., Kosorukova, N., Kostyleva, E., Kriiska, A., Kukawka, S., Lozovskaya, O., Mazurkevich, A., Nedomolkina, N., Piličiauskas, G., Sinitsyna, G., Skorobogatov, A., Smolyaninov, R.V., Surkov, A., Tkachov, O., Tkachova, M., Tsybrij, A., Tsybrij, V., Vybornov, A.A., Wawrusiewicz, A., Yudin, A.I., Meadows, J., Heron, C., Craig, O.E., 2023. The transmission of pottery technology among prehistoric European hunter–gatherers. Nat. Hum. Behav. 7, 171–183. https://doi.org/10.1038/s41562-022-01491-8.

Earle, T.K., Preucel, R.W., Brumfiel, E.M., Carr, C., Limp, W.F., Chippindale, C., Gilman, A., Hodder, I., Johnson, G.A., Keegan, W.F., Knapp, A.B., Potter, P.B., Rolland, N., Rowlett, R.M., Trigger, B.G., Zeitlin, R.N., 1987. Processual archaeology and the radical critique [and comments and reply]. Curr. Anthropol. 28, 501–538. https://doi.org/10.1086/203551.

Ellis, E.C., Gauthier, N., Klein Goldewijk, K., Bliege Bird, R., Boivin, N., Díaz, S., Fuller, D. Q., Gill, J.L., Kaplan, J.O., Kingston, N., Locke, H., McMichael, C.N.H., Ranco, D., Rick, T.C., Shaw, M.R., Stephens, L., Svenning, J.-C., Watson, J.E.M., 2021. People have shaped most of terrestrial nature for at least 12,000 years. Proc. Natl. Acad. Sci. 118, e2023483118. https://doi.org/10.1073/pnas.2023483118.

Eren, M., Buchanan, B. (Eds.), 2023. Defining and Measuring Diversity in Archaeology: Another Step toward an Evolutionary Synthesis of Culture. Berghahn, New York.

Eren, M.I., Chao, A., Chiu, C.-H., Colwell, R.K., Buchanan, B., Boulanger, M.T., Darwent, J., O'Brien, M.J., 2016. Statistical analysis of paradigmatic class richness supports greater paleoindian projectile-point diversity in the Southeast. Am. Antiq. 81, 174–192.

Field, J.S., 2004. Environmental and climatic considerations: a hypothesis for conflict and the emergence of social complexity in Fijian prehistory. J. Anthropol. Archaeol. 23, 79–99. https://doi.org/10.1016/j.jaa.2003.12.004.

Fort, J., 2023. Neolithic transitions: diffusion of people or diffusion of culture? In: Bunde, A., Caro, J., Chmelik, C., Kärger, J., Vogl, G. (Eds.), Diffusive Spreading in Nature, Technology and Society. Springer International Publishing, Cham, pp. 327–346. https://doi.org/10.1007/978-3-031-05946-9_16.

Friman, B., Lagerås, P., 2023. From neolithic boom-and-bust to iron age peak and decline: population and settlement dynamics in southern Sweden inferred from summed radiocarbon dates. Eur. J. Archaeol. 26, 168–188. https://doi.org/10.1017/eaa.2022.43.

Gallello, G., Ramacciotti, M., García-Puchol, O., Chenery, S., Cortell-Nicolau, A., Cervera, M.L., Diez-Castillo, A., Pastor, A., McClure, S.B., 2021. Analysis of stratigraphical sequences at Cocina Cave (Spain) using rare earth elements geochemistry. Boreas 50, 1190–1208. https://doi.org/10.1111/bor.12530.

González-Pérez, C., Martín-Rodilla, P., Pereira-Fariña, M. (Eds.), 2023. Discourse and Argumentation in Archaeology: Conceptual and Computational Approaches. Springer, Cham.

Grosman, L., 2016. Reaching the point of No return: the computational revolution in archaeology. Annu. Rev. Anthropol. 45, 129–145. https://doi.org/10.1146/annurev-anthro-102215-095946.

Hardesty, D.L., 1980. The use of general ecological principles in archaeology. In: Schiffer, M.B. (Ed.), Advances in Archaeological Method and Theory. Academic Press, San Diego, pp. 157–187. https://doi.org/10.1016/B978-0-12-003103-0.50009-5.

Harpending, H., 2008. Experimental workshop: do the existing methods really work? In: Matsumura, S., Forster, P., Renfrew, C. (Eds.), Simulations, Genetics and Human Prehistory. McDonald Institute Monographs. McDonald Institute for Archaeological Research. University of Cambridge, Cambridge, pp. 191–198.

Holland-Lulewicz, J., Ritchison, B.T., 2021. How many dates do I need?: using simulations to determine robust age estimations of archaeological contexts. Adv. Archaeol. Pract. 9, 272–287. https://doi.org/10.1017/aap.2021.10.

Kintigh, K.W., 1984. Measuring archaeological diversity by comparison with simulated assemblages. Am. Antiq. 49, 44–54. https://doi.org/10.2307/280511.

Kintigh, K.W., Altschul, J.H., Beaudry, M.C., Drennan, R.D., Kinzig, A.P., Kohler, T.A., Limp, W.F., Maschner, H.D.G., Michener, W.K., Pauketat, T.R., Peregrine, P., Sabloff, J.A., Wilkinson, T.J., Wright, H.T., Zeder, M.A., 2014. Grand challenges for archaeology. Am. Antiq. 79, 5–24. https://doi.org/10.7183/0002-7316.79.1.5.

Klingenberg, C.P., 2022. Methods for studying allometry in geometric morphometrics: a comparison of performance. Evol. Ecol. 36, 439–470. https://doi.org/10.1007/s10682-022-10170-z.

Klopfstein, S., Massingham, T., Goldman, N., 2017. More on the best evolutionary rate for phylogenetic analysis. Syst. Biol. 66, 769–785. https://doi.org/10.1093/sysbio/syx051.

Lake, M.W., 2014. Trends in archaeological simulation. J. Archaeol. Method Theor 21, 258–287. https://doi.org/10.1007/s10816-013-9188-1.

Laland, K.N., O'Brien, M.J., 2010. Niche construction theory and archaeology. J. Archaeol. Method Theor 17, 303–322. https://doi.org/10.1007/s10816-010-9096-6.

Leonard, R.D., Jones, G.T. (Eds.), 1989. Quantifying Diversity in Archaeology. Cambridge University Press, Cambridge.

Lipo, C.P., Madsen, M.E., Dunnell, R.C., Hunt, T., 1997. Population structure, cultural transmission, and frequency seriation. J. Anthropol. Archaeol. 16, 301–333. https://doi.org/10.1006/jaar.1997.0314.

Loog, L., Mirazón Lahr, M., Kovacevic, M., Manica, A., Eriksson, A., Thomas, M.G., 2017. Estimating mobility using sparse data: application to human genetic variation. Proc. Natl. Acad. Sci. 114, 12213–12218.

Marzuoli, A., 2008. Toy models in physics and the reasonable effectiveness of mathematics. In: Lupacchini, R., Corsi, G. (Eds.), Deduction, Computation, Experiment: Exploring the Effectiveness of Proof. Springer, Milan, Milano, pp. 49–64. https://doi.org/10.1007/978-88-470-0784-0_3.

Meyer, C., Lohr, C., Gronenborn, D., Alt, K.W., 2015. The massacre mass grave of Schöneck-Kilianstädten reveals new insights into collective violence in Early Neolithic Central Europe. Proc. Natl. Acad. Sci. 112, 11217–11222. https://doi.org/10.1073/pnas.1504365112.

Mithen, S.J., 1988. Simulation as a methodological tool: inferring hunting goals from faunal assemblages. In: Rugles, C.L.N., Rahts, S.P.Q. (Eds.), Computer and Quantitative Methods in Archaeology 1987. BAR International Series. B.A.R., Oxford, pp. 118–137.

Miu, E., Rendell, L., Bowles, S., Boyd, R., Cownden, D., Enquist, M., Eriksson, K., Feldman, M.W., Lillicrap, T., McElreath, R., Murray, S., Ounsley, J., Lala, K.N., 2024. The refinement paradox and cumulative cultural evolution: complex products of collective improvement favor conformist outcomes, blind copying, and hyper-credulity. PLoS Comput. Biol. 20, e1012436. https://doi.org/10.1371/journal.pcbi.1012436.

Mueller, J.W., 1974. The use of sampling in archaeological survey. Mem. Soc. Am. Archaeol. 28, 1–91 i-xi.

Murray, J.K., Benitez, R.A., O'Brien, M.J., 2021. The extended evolutionary synthesis and human origins: archaeological perspectives. Evol. Anthropol. Issues News Rev. 30, 4–7. https://doi.org/10.1002/evan.21837.

Nunn, C.L., Mulder, M.B., Langley, S., 2006. Comparative methods for studying cultural trait evolution: a simulation study. Cross-cult. Res. 40, 177–209. https://doi.org/10.1177/1069397105283401.

O'Brien, M.J. (Ed.), 2008. Cultural Transmission and Archaeology: Issues and Case Studies. Society for American Archaeology.

O'Brien, M.J., Lyman, R.L., 2000. Applying Evolutionary Archaeology: A Systematic Approach. Springer US, Boston, MA. https://doi.org/10.1007/b100324.

O'Brien, M.J., Lyman, R.L., Schiffer, M.B., 2005. Archaeology as a Process: Processualism and its Offspring. University of Utah Press, Salt Lake City.

O'Brien, M.J., Thomas, D.H., 2022. Foreword. In: Eren, M., Buchanan, B. (Eds.), Defining and Measuring Diversity in Archaeology: Another Step toward an Evolutionary Synthesis of Culture. Berghahn, New York, pp. XX–XXIV.

Olalde, I., Mallick, S., Patterson, N., Rohland, N., Villalba-Mouco, V., Silva, M., Dulias, K., Edwards, C.J., Gandini, F., Pala, M., Soares, P., Ferrando-Bernal, M., Adamski, N., Broomandkhoshbacht, N., Cheronet, O., Culleton, B.J., Fernandes, D., Lawson, A.M., Mah, M., Oppenheimer, J., Stewardson, K., Zhang, Z., Arenas, J.M.J., Moyano, I.J.T., Salazar-García, D.C., Castanyer, P., Santos, M., Tremoleda, J., Lozano, M., Borja, P.G., Fernández-Eraso, J., Mujika-Alustiza, J.A., Barroso, C., Bermúdez, F.J., Mínguez, E.V., Burch, J., Coromina, N., Vivó, D., Cebrià, A., Fullola, J.M., García-Puchol, O., Morales, J.I., Oms, F.X., Majó, T., Vergès, J.M., Díaz-Carvajal, A., Ollich-Castanyer, I., López-Cachero, F.J., Silva, A.M., Alonso-Fernández, C., Castro, G.D. de, Echevarría, J.J., Moreno-Márquez, A., Berlanga, G.P., Ramos-García, P., Ramos-Muñoz, J., Vila, E.V., Arzo, G.A., Arroyo, Á.E., Lillios, K.T., Mack, J., Velasco-Vázquez, J., Waterman, A., Enrich, L.B. de L., Sánchez, M.B., Agustí, B., Codina, F., Prado, G. de, Estalrrich, A., Flores, Á.F., Finlayson, C., Finlayson, G., Finlayson, S., Giles-Guzmán, F., Rosas, A., González, V.B., Atiénzar, G. G., Pérez, M.S.H., Llanos, A., Marco, Y.C., Beneyto, I.C., López-Serrano, D., Tormo, M.S., Valera, A.C., Blasco, C., Liesau, C., Ríos, P., Daura, J., Michó, M.J. de P., Diez-Castillo, A.A., Fernández, R.F., Farré, J.F., Garrido-Pena, R., Gonçalves, V.S., Guerra-Doce, E., Herrero-Corral, A.M., Juan-Cabanilles, J., López-Reyes, D., McClure, S.B., Pérez, M.M., Foix, A.O., Borràs, M.S., Sousa, A.C., Encinas, J.M.V., Kennett, D.J., Richards, M.B., Alt, K.W., Haak, W., Pinhasi, R., Lalueza-Fox, C., Reich, D., 2019. The genomic history of the Iberian Peninsula over the past 8000 years. Science 363, 1230–1234. https://doi.org/10.1126/science.aav4040.

Orton, C., 1973. The tactical use of models in archaeology - the SHERD project. In: Renfrew, C. (Ed.), The Explanation of Culture Change. Duckworth, London, pp. 137–139.

Paine, R.R., Harpending, H., 1996. Assessing the reliability of paleodemographic fertility estimators using simulated skeletal distributions. Am. J. Phys. Anthropol. 101, 151–159. https://doi.org/10.1002/(SICI)1096-8644(199610)101:2<151::AID-AJPA3>3.0.CO;2-5.

Perlin, K., 1985. An image synthesizer. ACM SIGGRAPH Comput. Graph. 19, 287–296. https://doi.org/10.1145/325165.325247.

Perreault, C., 2019. The Quality of the Archaeological Record. The University of Chicago Press, Chicago.

Piperno, D.R., Ranere, A.J., Dickau, R., Aceituno, F., 2017. Niche construction and optimal foraging theory in Neotropical agricultural origins: a re-evaluation in consideration of the empirical evidence. J. Archaeol. Sci. 78, 214–220. https://doi.org/10.1016/j.jas.2017.01.001.

Piskin, E., Marciniak, A., Bartkowiak, M. (Eds.), 2018. Environmental Archaeology: Current Theoretical and Methodological Approaches. Springer, New York.

Posth, C., Yu, H., Ghalichi, A., Rougier, H., Crevecoeur, I., Huang, Y., Ringbauer, H., Rohrlach, A.B., Nägele, K., Villalba-Mouco, V., Radzeviciute, R., Ferraz, T., Stoessel, A., Tukhbatova, R., Drucker, D.G., Lari, M., Modi, A., Vai, S., Saupe, T., Scheib, C.L., Catalano, G., Pagani, L., Talamo, S., Fewlass, H., Klaric, L., Morala, A., Rué, M., Madelaine, S., Crépin, L., Caverne, J.-B., Bocaege, E., Ricci, S., Boschin, F., Bayle, P., Maureille, B., Le Brun-Ricalens, F., Bordes, J.-G., Oxilia, G., Bortolini, E., Bignon-Lau, O., Debout, G., Orliac, M., Zazzo, A., Sparacello, V., Starnini, E., Sineo, L., van der Plicht, J., Pecqueur, L., Merceron, G., Garcia, G., Leuvrey, J.-M., Garcia, C.B., Gómez-Olivencia, A., Połtowicz-Bobak, M., Bobak, D., Le Luyer, M., Storm, P., Hoffmann, C., Kabaciński, J., Filimonova, T., Shnaider, S., Berezina, N., González-Rabanal, B., González Morales, M.R., Marín-Arroyo, A.B., López, B., Alonso-Llamazares, C., Ronchitelli, A., Polet, C., Jadin, I., Cauwe, N., Soler, J., Coromina, N., Rufí, I., Cottiaux, R., Clark, G., Straus, L.G., Julien, M.-A., Renhart, S., Talaa, D., Benazzi, S., Romandini, M., Amkreutz, L., Bocherens, H., Wißing, C., Villotte, S., de Pablo, J.F.-L., Gómez-Puche, M., Esquembre-Bebia, M.A., Bodu, P., Smits, L., Souffi, B., Jankauskas, R., Kozakaitė, J., Cupillard, C., Benthien, H., Wehrberger, K., Schmitz, R.W., Feine, S.C., Schüler, T., Thevenet, C., Grigorescu, D., Lüth, F., Kotula, A., Piezonka, H., Schopper, F., Svoboda, J., Sázelová, S., Chizhevsky, A., Khokhlov, A., Conard, N.J., Valentin, F., Harvati, K., Semal, P., Jungklaus, B., Suvorov, A., Schulting, R., Moiseyev, V., Mannermaa, K., Buzhilova, A., Terberger, T., Caramelli, D., Altena, E., Haak, W., Krause, J., 2023. Palaeogenomics of upper palaeolithic to neolithic European hunter–gatherers. Nature 615, 117–126. https://doi.org/10.1038/s41586-023-05726-0.

Premo, L.S., 2014. Cultural transmission and diversity in time-averaged assemblages. Curr. Anthropol. 55, 105–114.

Prentiss, A.M., 2021. Theoretical plurality, the extended evolutionary synthesis, and archaeology. Proc. Natl. Acad. Sci. 118, e2006564118. https://doi.org/10.1073/pnas.2006564118.

R Core Team, 2023. R: A Language and Environment for Statistical Computing.

Rendell, L., Boyd, R., Cownden, D., Enquist, M., Eriksson, K., Feldman, M.W., Fogarty, L., Ghirlanda, S., Lillicrap, T., Laland, K.N., 2010. Why copy others? Insights from the social learning strategies tournament. Science 328, 208–213. https://doi.org/10.1126/science.1184719.

Rhode, D., 1988. Measurement of archaeological diversity and the sample-size effect. Am. Antiq. 53, 708–716. https://doi.org/10.2307/281114.

Rick, T.C., Sandweiss, D.H., 2020. Archaeology, climate, and global change in the Age of Humans. Proc. Natl. Acad. Sci. 117, 8250–8253. https://doi.org/10.1073/pnas.2003612117.

Roksandic, M., Djurić, M., Rakočević, Z., Seguin, K., 2006. Interpersonal violence at lepenski vir mesolithic/neolithic complex of the iron gates gorge (Serbia-Romania). Am. J. Phys. Anthropol. 129, 339–348. https://doi.org/10.1002/ajpa.20286.

Rubio Campillo, X., Cela, J.M., Hernàndez Cardona, F.X., 2012. Simulating archaeologists? Using agent-based modelling to improve battlefield excavations. J. Archaeol. Sci. 39, 347–356. https://doi.org/10.1016/j.jas.2011.09.020.

Rubio-Campillo, X., 2016. Model selection in historical research using approximate bayesian computation. PLoS One 11, e0146491. https://doi.org/10.1371/journal.pone.0146491.

Schiffer, M.B., Sullivan, A.P., Klinger, T.C., 1978. The design of archaeological surveys. World Archaeol. 10 (1), 1–28.

Silberzahn, R., Uhlmann, E.L., Martin, D.P., Anselmi, P., Aust, F., Awtrey, E., Bahník, Š., Bai, F., Bannard, C., Bonnier, E., Carlsson, R., Cheung, F., Christensen, G., Clay, R., Craig, M.A., Dalla Rosa, A., Dam, L., Evans, M.H., Flores Cervantes, I., Fong, N., Gamez-Djokic, M., Glenz, A., Gordon-McKeon, S., Heaton, T.J., Hederos, K., Heene, M., Hofelich Mohr, A.J., Högden, F., Hui, K., Johannesson, M., Kalodimos, J., Kaszubowski, E., Kennedy, D.M., Lei, R., Lindsay, T.A., Liverani, S., Madan, C.R., Molden, D., Molleman, E., Morey, R.D., Mulder, L.B., Nijstad, B.R., Pope, N.G., Pope, B., Prenoveau, J.M., Rink, F., Robusto, E., Roderique, H., Sandberg, A., Schlüter, E., Schönbrodt, F.D., Sherman, M.F., Sommer, S.A., Sotak, K., Spain, S., Spörlein, C., Stafford, T., Stefanutti, L., Tauber, S., Ullrich, J., Vianello, M., Wagenmakers, E.-J., Witkowiak, M., Yoon, S., Nosek, B.A., 2018. Many analysts, one data set: making transparent how variations in analytic choices affect results. Adv. Methods Pract. Psychol. Sci. 1, 337–356. https://doi.org/10.1177/2515245917747646.

Smith, V.A., Jarvis, E.D., Hartemink, A.J., 2002. Evaluating functional network inference using simulations of complex biological systems. Bioinformatics 18, S216–S224. https://doi.org/10.1093/bioinformatics/18.suppl_1.S216.

Steele, J., 2009. Human dispersals: mathematical models and the archaeological record. Hum. Biol. 81, 121–140. https://doi.org/10.3378/027.081.0302.

Surovell, T.A., Brantingham, P.J., 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, T.A., Byrd Finley, J., Smith, G.M., Brantingham, P.J., 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.

Surovell, T.A., Pelton, S.R., 2016. Spatio-temporal variation in the preservation of ancient faunal remains. Biol. Lett. 12, 20150823. https://doi.org/10.1098/rsbl.2015.0823.

Takata, Y., Yanase, P., 2021. The production, preservation and dissemination of archaeological data in Japan. Internet Archaeol. https://doi.org/10.11141/ia.58.11.

Thomas, D.H., 1978. The awful truth about statistics in archaeology. Am. Antiq. 43, 231–244. https://doi.org/10.2307/279247.

Timpson, A., Barberena, R., Thomas, M.G., Méndez, C., Manning, K., 2021. Directly modelling population dynamics in the South American Arid Diagonal using 14C dates. Philos. Trans. R. Soc. B Biol. Sci. 376, 20190723. https://doi.org/10.1098/rstb.2019.0723.

Turchin, P., 2013. The puzzle of human ultrasociality: how did large-scale complex societies evolve? In: Richerson, P.J., Christiansen, M.H. (Eds.), Cultural Evolution: Society, Technology, Language, and Religion. The MIT Press, pp. 61–73. https://doi.org/10.7551/mitpress/9780262019750.003.0004.

Varien, M.D., Potter, J.M., 1997. Unpacking the discard equation: simulating the accumulation of artifacts in the archaeological record. Am. Antiq. 62, 194–213. https://doi.org/10.2307/282506.

Wells, E.C., 2010. Sampling design and inferential bias in archaeological soil chemistry. J. Archaeol. Method Theor 17, 209–230. https://doi.org/10.1007/s10816-010-9087-7.

Wood, W.R., Lee johnson, D., 1978. A survey of disturbance processes in archaeological site formation. In: Schiffer, M.B. (Ed.), Advances in Archaeological Method and Theory. Academic Press, San Diego, pp. 315–381. https://doi.org/10.1016/B978-0-12-003101-6.50015-9.

Xie, Y., 2016. Bookdown: Authoring Books and Technical Documents with R Markdown. Chapman and Hall/CRC, New York. https://doi.org/10.1201/9781315204963.

Xu, C., Kohler, T.A., Lenton, T.M., Svenning, J.-C., Scheffer, M., 2020. Future of the human climate niche. Proc. Natl. Acad. Sci. 117, 11350–11355. https://doi.org/10.1073/pnas.1910114117.

Yorston, R.M., Gaffney, V.L., Reynolds, P.J., 1990. Simulation of artefact movement due to cultivation. J. Archaeol. Sci. 17, 67–83. https://doi.org/10.1016/0305-4403(90)90015-W.