Reconstructing the full temporal range of archaeological phenomena from sparse data

Alastair Key [1], David Roberts [2], and Ivan Jarić [3,4],

Corresponding author: AK2389@cam.ac.uk

[1] Department of Archaeology, University of Cambridge, Cambridge, CB2 3DZ, UK
[2] Durrell Institute of Conservation and Ecology, School of Anthropology and Conservation, University of Kent, Canterbury, Kent, CT2 7NR, UK
[3] Biology Centre of the Czech Academy of Sciences, Institute of Hydrobiology, České Budějovice, Czech Republic
[4] University of South Bohemia, Faculty of Science, Department of Ecosystem Biology, České Budějovice, Czech Republic

Keywords:
Temporal distribution; modelling; end date and origin inference; optimal linear estimation; Weibull; computational archaeology; archaeological record

Abstract
Archaeologists rarely discover the first or last known occurrences of past cultural phenomena, yet 'start' and 'end' dates are central to our understanding of past human behaviour; therein lays a paradox long known within the discipline. Optimal linear estimation (OLE) has recently been used to reconstruct the full temporal range of prehistoric archaeological technologies using only the partial records available. That is, OLE has been used to reconstruct the portions of the archaeological record not yet evidenced through artefact discoveries. Here we present OLE to a wider archaeological audience and outline for the first time the model's assumptions as they pertain to archaeological phenomena. We demonstrate OLE to be an accessible, user-friendly and methodologically transparent temporal range estimation method applied via a single set of equations. Further, we present five additional frequentist techniques that enable archaeologists to account for observation reliability, search effort and extreme data scarcity when inferring temporal ranges. These methods allow archaeologists to gain a more accurate understanding of the temporal range of past human behaviour.

1. **Introduction**

Recently introduced to archaeology from conservation science and palaeontological studies (e.g., Roberts and Solow, 2003; Pimiento and Clements, 2014; Zhang et al., 2020), optimal linear estimation (OLE) modelling is proving a useful technique to infer the full chronology of archaeological phenomena (Key et al, 2021a, 2021b; Bebber and Key, 2021). This brief communication is designed to introduce the technique to a wider archaeological audience, and to discuss how the model's predictions and assumptions interact with the unique nature of archaeological phenomena. We also present several additional, related techniques that are new to archaeological science and can be used separately or in conjunction with OLE to increase the accuracy of temporal estimates.

Unlike traditional assessments which use dated artefacts as a start or end point, OLE can infer how much longer a phenomenon is likely to have persisted prior to, or after, these known dates. That is, OLE is able to reconstruct the portion of the archaeological record that has not yet been discovered and provide a more accurate account of an archaeological phenomena's temporal presence (Fig. 1). To date, the technique has been used to extend the Oldowan and Acheulean periods by tens of thousands of years (Key et al., 2021a 2021b), and push back the origin of North American copper use by several hundred years (Bebber and Key, 2021).

The importance of reconstructing the earliest and latest portions of an archaeological phenomenon's temporal range has long been recognised (e.g., Cowgill, 1972; Surovell and Brantingham, 2007; Crema, 2012). Currently, Bayesian modelling approaches are the dominant (*c.f.* Buck and Meson, 2015; Ramsey, 2015) method applied to solving temporal probability questions in archaeological research. Although techniques addressing temporal uncertainty in archaeological occurrences are also well known (Crema, 2012; Kolář et al., 2016; Baxter and Cool, 2016). Most often used to calibrate radiocarbon dates (Bayliss, 2015; Crema and Bevan, 2021), Bayesian models can provide 'start' or 'end' dates for a given phenomenon based on probability densities derived from groups of temporally bounded radiocarbon samples (e.g., Wicks et al., 2014; Bicho et al., 2015). In turn, temporal range estimations have generally been limited to archaeological phenomena with associated radiocarbon dates.

Importantly, most Bayesian techniques applied to radiocarbon probability distributions do not directly model the start or end date of a phenomena. Instead, they provide revised probability estimates reconstructed from the probability densities of known radiocarbon samples; they do not directly reconstruct yet-to-be-discovered portions of the archaeological record. In other words, Bayesian techniques estimate when known artefact records occurred. Further, Bayesian techniques do not always consider changes to artefact occurrence frequencies through time (although it is possible [e.g., Ramsey, 2015; Fernández-López de Pablo and Barton, 2015; Banks et al., 2019; Crema and Kobayashi, 2020]).) Finally, Bayesian techniques (see: Otarola-Castillo and Torquato [2018] and references therein) are underpinned by the necessity of including prior assumptions about the phenomena they are investigating (Litton and Buck, 1995), meaning that inherent to any Bayesian temporal range estimates are subjective interpretations of prior information specified by the archaeologist (Pettitt and Zilhão, 2015) (although techniques to limit the impact of unverified *a priori* information can be applied [e.g., Long and Taylor, 2015]). While useful in many contexts, Bayesian techniques require prior information that can often be unavailable, are mathematically demanding, and can be difficult to compare to more traditional frequentist methods (Pettitt and Zilhão, 2015; Brook et al., 2019). It is for these reasons that their use has been limited within conservation studies (Boakes et al., 2015).

In contrast, OLE makes very few prior assumptions, and as a frequentist method based on the extreme value theory its predictions are rooted in the temporal spacing of the data that it investigates. This means that partial temporal records (i.e., periodic occurrences of artefacts through time) are not only easily accommodated into the method but are intrinsic to its predictions. Moreover, it can be applied to any type of dating information, any archaeological phenomena, and at any timescale, so long as it is represented by lineal occurrences through time, be this highly sporadic early stone tool technologies dated through optically stimulated luminiscence (OSL), faunal, $40^{Ar}/39^{Ar}$ or radiocarbon methods, or more recent archaeological phenomena dated through historical texts or ethnographic sightings. Its implementation via a single set of equations further increases its transparency and accessibility as an analytical tool. In other words, OLE has potential to be applied within any archaeological context and by any archaeologist.

Other temporal range estimation techniques have been used in conservation science and paleobiology research, and could potentially be used to reconstruct the archaeological record. However, some have been shown to be outperformed by OLE (Rivadeneira et al., 2009) or require records to represent stationary Poisson processes (i.e., there is no decline in frequency towards extinction, or increase in frequency after invention) (e.g., Solow, 1993), which is not realistic for most archaeological phenomena (Surovell and Brantingham, 2007; Mesoudi and Lycett, 2009; Lycett, 2015; Mesoudi, 2015; Shennan, 2015). Indeed, one of the reasons that OLE can be readily applied within archaeology is that, as with species, the origination or extinction of culture is not sudden. It develops over time, meaning the likelihood of identifying sites that represent the start and end points of a phenomena are incredibly low (Surovell and Brantingham, 2007; Prasciunas and Surovell, 2015). Rather, phenomena are likely to have persisted for longer periods, and these periods will be relatively difficult to detect in the archaeological record. There are always exceptions, such as catastrophic events that quickly wipe out a population, but these will be less common.

It is important to note that although first developed for conservation science (Roberts and Solow, 2003; Solow, 2005), OLE has no parameters specific to biological organisms and can readily be applied to cultural traditions. Moreover, the analogous mechanisms underpinning biological and cultural evolution allow for similar factors to be influencing the start/end dates of both (Lycett, 2015; Mesoudi, 2015; Shennan, 2015). In addition to OLE, we also briefly describe here five other techniques that enable researchers to account for observation reliability (Jarić and Roberts, 2014; Brook et al., 2019), search effort (McCarthy, 1998), and situations with extreme data scarcity, such as datasets consisted of only two records (Solow and Roberts, 2003) or just a single record (Roberts and Jarić, 2020), when inferring temporal ranges.

2. **How does optimal linear estimation (OLE) work?**

OLE requires the oldest or youngest dated occurrences of a phenomena to be entered into the model (depending on whether it is being used to estimate a ‘start’ or an ‘end’), from which the timings and chronological spacing of these known occurrences are used to statistically estimate how much earlier or longer the phenomena is likely to have existed. Ten dates are generally recommended as being optimal (Solow, 2005; Rivadeneira et al., 2009), although lower sample sizes (e.g., n = 5) have also been demonstrated to display good accuracy (Clements et al., 2013). OLE relies on the assumption that the dates entered into the model display (at least roughly) a joint distribution with a ‘Weibull form’. As noted above, this is a valid assumption for most archaeological phenomena. In turn, the shape (form) parameters of the Weibull distribution used in the OLE model are based on the chronology (spacing) of the dates entered. An ‘end’ or ‘start’ point can then be determined and is defined as the point at which the Weibull distribution determines that another occurrence should have been found had the cultural phenomenon not ended or not yet existed (relative to the temporal direction of the model) (Fig. 1).

In case of a phenomenon ending, the model predicts that given the temporal spacing of the known archaeological record, we would have expected to have found a younger archaeological site relative to the youngest currently known, if the phenomenon still existed beyond the tail end of the model’s distribution. Given that no sites have been found, we can infer the phenomenon ends at this point. In cases of origination, the model infers that given the temporal spacing of the known archaeological record, another earlier occurrence of the phenomenon would have been expected to have been found, relative to the currently known oldest occurrence, if it existed beyond the tail of the modelled distribution.

2.1 **Assumptions**

Although OLE has few underlying assumptions compared to other temporal modelling techniques (Solow, 2005; Rivadeneira et al., 2009; Clements et al., 2013; Pettitt and Zilhão, 2015; Otarola-Castillo and Torquato, 2018), there are several to be aware of. First, OLE assumes a continuation of the phenomena in question after or before the latest or earliest (respectively) currently known occurrences. Simply put, it assumes that it is unlikely that the most recent or earliest observation of a process would be the last/first point in time when the process was active. This can be safely assumed for most archaeological phenomena (see above), but near instantaneous ends to cultural phenomena are possible. This could range from catastrophic natural disasters (e.g., volcanic eruptions) through to the rapid replacement of one artefact type with another (e.g., changes to coinage). In instances where this is assumed, other methods, designed for cases with constant sighting rates prior to a phenomenon's end, could be considered instead, such as that by Solow (1993). However, it should be noted that OLE and other frequentist methods are sensitive to, and account for, the amount of available data on the studied phenomenon. For example, in cases of well-studied and recent artefact types, where occurrences are temporally densely distributed and frequent, predicted end/start points would shift to and closely match the time of the last known record.

Second, OLE assumes all observations (in this case archaeological occurrences) to be discrete. When modelling species extinction this means that each observation is assumed to represent a different, independent sighting event. Given that artefacts represent physical manifestations of cultural information contained within a biological repository, and ultimately it is the chronology of the cultural information that is being modelled, a 'discrete' archaeological occurrence is at its most extreme an artefact/assemblage assumed to be produced by a unique individual relative to those already included in the model (Bebber and Key, 2021). This means that all artefactual occurrences should be assumed to have been made by different individuals. At its most moderate, 'discrete' may simply refer to an independent representation of the phenomena; although this entails an assumption of random sampling from a population. Several identical dates can be used so long as they all meet this criterion. Indeed, variation in data type, quality and characteristics needs to be considered carefully to avoid biases when applying OLE (as with any modelling technique). Highly clustered data make an important example, and the archaeological record is full of waste middens, battle grounds, occupancy sites, artefact caches and other phenomena that should be considered carefully in terms of what precisely they represent when included within the model. In cases where cultural phenomena are broader and the product of multiple individuals/craftsperson's, such as specific settlement types or some ceremonial outfits, then it may be more appropriate to consider discrete observations at a population level.

Third, although OLE does not assume the probability of detection to be consistent, it does assume relatively stable search effort that never equates to zero. When modelling species extinction, this means that any breaks in between dates in the model (occurrences) are not a function of irregular search effort. In other words, the model assumes that gaps in species sighting data are not because people are not looking for the species. This assumption is less relevant to studies of paleontological or archaeological phenomena, as the temporal record of fossils and artefacts through time (and therefore their spacing) are not dictated by human search effort, but geological and taphonomic factors in combination with frequency changes (Surovell and Brantingham, 2007; Surovell et al., 2009). Archaeologists can display biases in search effort towards sediments of a specific age, but assuming that there is no intentional avoidance of artefacts then this assumption should not be violated in most archaeological instances (notably, geographic biases in search effort may also influence temporal records).

Although OLE makes no other assumptions, Clements et al. (2013) highlighted two additional factors that could potentially influence the model's accuracy. The first includes sudden changes to the rate of a phenomena's decline/uptake. Within an archaeological context, this means the addition of new pressures could alter the rate of cultural decline/uptake experienced for some, but not all, occurrences in the model. For example, sudden environmental change relevant to only the latest occurrences of an 'extinction' (end) model could additionally speed up the phenomena's decline due to additional stresses on the population. The second factor concerns changes to the 'observability' of a phenomena. This does not relate to the naturally increasing/decreasing likelihood of discovering an occurrence as the phenomena become more or less widespread (Surovell and Brantingham, 2007; Mesoudi and Lycett, 2009; Jordan and Cummings, 2014). Instead, it refers to a change in relative observability. For instance, taphonomic processes can introduce natural bias to the preservation of archaeological sites, which can impact observability irrespective of the demographic and frequency trends present at the time of existence (Surovell and Brantingham, 2007; Surovell et al., 2009). Further, spatial considerations may also impact observability through changes to the accessibility of past landscapes (for example, earlier or later occurrences may be more or less likely to be discovered due to flooding events [e.g., the flooding of Doggerland]). Thus, these factors should be assumed constant when using OLE estimates (see Solow [2005], Clements et al. [2013], and Boakes et al. [2015] for further details).

Finally, there are several expectations relevant to using OLE in archaeology that are not necessarily applicable in other contexts. First, there is an inherent assumption that the dates used in the model are an accurate representation of the temporal presence of the occurrence in question (see: Crema, 2012). That is, the dates associated with an artefact/site should accurately reflect when it entered the archaeological record (or at least, as accurate as the relevant dating method can be). Finally, there should be sufficient accuracy of an archaeological occurrence's identification when including it in an OLE model. This is primarily an issue when typological artefact classifications are subjective and have potential to vary between analysts; after all, one cannot identify the final or first occurrence of a phenomena if it cannot be reliably detected. However, this is likely more an issue for some archaeological fields than others.

2.2 OLE model

The formulaic expression of OLE has recently been published in open access archaeological literature (Key et al., 2021a, 2021b), and is widely available elsewhere, including the original articles describing the technique (Roberts and Solow, 2003; Solow, 2005). Further, the R sExtinct software package provides an easily accessible means through which to run OLE (Clements, 2013). We provide an explanation of the OLE equations in the attached Supplementary Information, along with a link to the R script provided by Clements (2013). Notably, models inferring the 'start' date of archaeological phenomena need to be adjusted to run in the reverse temporal direction to those provided by Clements (2013; *cf*. Key et al., 2021b). The 10 youngest or oldest dates should be approximately used as the beginning of the period, dependent on the direction of the model. As with any frequentist model the most common time unit used is years, but decades, centuries or millennia can be used just as well.

As noted above, OLE provides temporal estimates based on extreme Weibull form distributions that are tailored relative to the dated occurrences used, their distribution, and intervals observed between these dates. The distribution's curve is then used to identify an 'end point' or 'point of origin' beyond which the phenomenon can be inferred to no longer exist, as if it did (given current artefact distributions) another artefactual occurrence would be known about between the last/earliest current known occurrence and the inferred point of end/origin. To put it another way, OLE asks, given the distribution of known occurrences, how likely is it that another does not exist in the archaeological record.

Two estimated dates are commonly produced by OLE. One represents the estimated origin (T_O) or end (T_E) date of the phenomenon in question. The other represents the upper bound of each model's confidence interval (T_{CI}), although lower bounds can also easily be produced and applied (Roberts and Solow, 2003). T_O and T_E dates are the main output of archaeological OLE models and represent the 'end' or 'origin' point in whatever dating time scale used for the input data (e.g., BP, BCE). T_{CI} dates represent the point beyond which there is a 5% or less probability (in line with $\alpha = 0.05$) that the phenomenon still existed. In other words, T_{CI} dates represent a confidence limit where in 95 out of 100 cases, the true end/origin point will be within the confidence interval.

Archaeological dating methods often come with a degree of uncertainty. That is, although the age of artefactual occurrences can be identified, they often come with error ranges, or date ranges with varying distributions of likelihood. Combined with the OLE method's use of discrete time units (e.g., specific years), it means that point estimates must be identified and used in place of these ranges. It is possible to run individual OLE models using central tendency values taken from ranges, and this will provide a reasonable estimated account of the phenomenon's temporal range (particularly if the date range has an approximate normal distribution of likelihood). Other methods for deciding on point estimates can be used, so long as they are justified (e.g., a secondary dating technique indicating that a specific portion of the first techniques range is more likely). In other instances, however, it may be favourable to account for temporal range uncertainty through repeated sampling procedures. As already applied elsewhere (Bebber and Key, 2021; Key et al., 2021), it is possible to draw dates randomly from a normal or uniform distribution (or other distribution, if preferred) within a defined range. This is repeated for all investigated occurrences and the randomly generated datasets are subsequently assessed with the OLE method. By repeating this procedure many times (e.g., 10,000) and using averages from the OLE data produced, it becomes possible to account for the temporal uncertainty associated with many archaeological dating techniques. Additional frequentist techniques that specifically account for record reliability, including dating uncertainty, are described below (Section 3.1).

2.3 How does temporal spacing of dated archaeological occurrences impact OLE estimations?

OLE has repeatedly been demonstrated to be robust within a variety of scenarios, including those that vary in temporal scale, 'sighting' probabilities, or search effort and trajectories (Rivadeneira et al., 2009; Clements et al., 2013). This means that OLE tends to remain accurate under most archaeological scenarios. As with any statistical modelling, however, the characteristics of the data entered into the model determine the output and the inferences derived thereof. Thus, it is useful to highlight how variation in sampling patterns and discovery happenstance can influence temporal range estimations. Figure 1 details a typical sequence of ten archaeological finds through time, under four different scenarios. This is a simple example and does not represent a test of robustness, as this has already been undertaken elsewhere (see Rivadeneira et al. 2009; Clements et al., 2013). Nevertheless, it is a useful illustration of the method's performance, as it highlights its sensitivity to the distribution and trends in records.

Figure 1a details a typical artefact sequence where there is a greater number of earlier artefactual occurrences (blue), but as the phenomena increasingly moves towards its end the likelihood of finding artefacts/sites decreases until there is a final known occurrence (yellow). Using this example, OLE models predict the phenomena to end at 965 years before present (BP), 35 years after the youngest known artefact (1000 BP). Using the same artefactual occurrences but with several being randomly removed (such that n = 5, while ensuring the youngest date remains the same), the model's inferred end date shifts slightly forward to 957 BP (Fig. 1b), suggesting a slightly longer persistence of the phenomenon. Of course, this is a simple example, but it highlights the ability of the method to be used with low sample sizes, albeit with increased uncertainty.

The factor most likely to influence an OLE model's prediction is the temporal density of dates close to the latest or earliest known occurrence (depending on its temporal direction). If there are greater numbers of known occurrences immediately prior to the last one, then the model will predict a closer end date (989 BP; Fig. 1c). Indeed, it follows that if there are several closely dated sites and then an absence of occurrences, it is likely that the last known record will be reasonably close to the true end date of the phenomenon. Conversely, relatively few records prior to the last known occurrence will point to a likely declining trend of occurrence, and that the phenomenon likely continued for an extended period with low chance of detection. Thus, in these instances the model has increased uncertainty and predicts a relatively longer tail of presence for the phenomenon (942 BP; Fig. 1d).

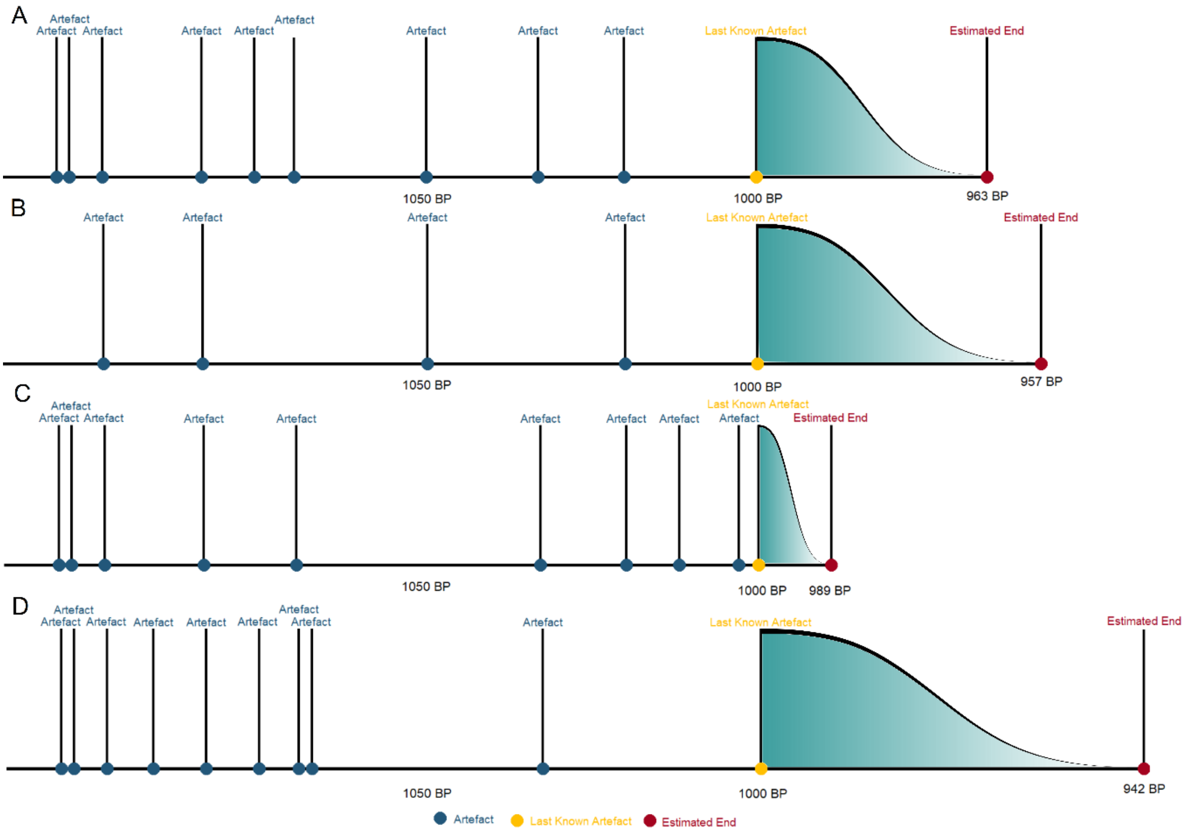

Figure 1: Examples of how OLE temporal estimates are influenced by date distributions (these illustrations are derived from real OLE estimates). Figure A represents a 'standard' archaeological scenario where more occurrences exist closer to the phenomenon's peak and then gradually decrease prior to the 'last known artefact' (n = 10). Figure B represents the same scenario, but happenstance and/or reduced search effort has resulted in only five artefacts (and therefore dates) being available (n = 5; randomly sampled from Figure A's scenario). Figure C illustrates improved surveying and/or preservation conditions, which has resulted in a greater number of instances having been found prior to the last known artefact (n = 10). This will predict a faster end to the phenomenon; the model's logic is that *if* it had continued beyond 989 BP, archaeologists should have discovered another site between 1000 – 989 BP given the density and distribution of sites found prior to 1000 BP. Figure D demonstrates the reverse scenario, where fewer dates are available around the 'last known artefact', indicating a declining trend of occurrence. In such a scenario, a quick end to the phenomenon is unlikely due to extended intervals between occurrences, and therefore the increased likelihood that artefacts do exist but remain undiscovered. Note that scenarios C and D are relatively extreme examples for demonstrative purposes.

3. Further frequentist range estimation methods

3.1 **Methods that account for record reliability**

Archaeological records often contain artefacts of variable reliability. This may stem from their state of preservation and other factors impeding identification, uncertainty of their origin (e.g., surface finds), or credibility of the person reporting the find. Whether an artefact or site is accepted as a valid record can strongly affect the inferences derived using temporal estimation methods (Roberts et al., 2010). Thus, it is important for the reliability of occurrence identification to be considered as part of some archaeological scenarios that apply these techniques.

The majority of the existing temporal estimation methods designed to handle mixed-certainty records either classify them as reliable or unreliable, or assign them a score that represents probability that a record is true, based on the record type and its characteristics (Boakes et al., 2015; Brook et al., 2019). While the common approach for treating variable reliability is to include or discard records based on their type or arbitrary inclusion probability thresholds, recently developed approaches by Jarić and Roberts (2014) and Brook et al. (2019) allow for individual record reliabilities to be directly included in existing temporal inference methods (including OLE).

The approach by Jarić and Roberts (2014) represents a simple modification of existing methods and allows each individual record to be weighted based on its reliability score. Although so far only applied to the standard Solow method (Solow, 1993), Jarić and Roberts (2014) suggest that it can also be readily applied to other frequentist methods, including those by Strauss and Sadler (1989), Solow and Roberts (2003) and McInerny et al. (2006). The modification works by replacing standard record time series, represented by a binary sequence of presences and absences, with presences expressed as probabilities that represent the reliability of each individual record (i.e., a likelihood that the given observation is true) (Jarić and Roberts, 2014). In effect, the number or records in a dataset is replaced by a sum of probability values assigned to each record, indicating the most likely number of observations. Additionally, individual record reliabilities are used to estimate the most likely endpoint of a sighting record, based on the likelihood that all later records in the dataset are false. Brook et al. (2019) introduced additional improvements to this approach by also estimating the likely year of the first true record, based on the likelihood that all preceding records are false. Together, these approaches can substantially improve the predictive power of origin and end-point (extinction) estimation methods when occurrence records display mixed-certainty.

A further simple and powerful extension to OLE and other existing temporal range estimation methods was proposed by Brook et al. (2019). The approach is based on resampling without replacement from a dataset of records, where reliability of each record is used as its probability of being sampled and included. Repeated sampling and extinction inference based on each sample record produces a range and frequency distribution around the extinction date. The method was demonstrated to be robust and to perform well, and has a number of advantages compared to existing methods: 1) it is based on resampling of records and these can thereafter be used as input for other end/origin estimation methods, including OLE; 2) it is simple and easily applicable using the freely available R script; 3) it allows direct integration of overlapping records (Brook et al., 2019).

3.2 **Methods that account for collection effort**

As discussed previously, relatively stable collection effort represents one of the key assumptions of OLE and other methods from this group. Any changes in collection effort over geological time will be confounded with true changes and trends in the presence and frequency of phenomenon studied, and might consequently lead to biased results. We note this to potentially be less of an issue for archaeology relative to other fields using these techniques (Section 2.1), but there are situations where collection efforts may be biased in favour of older or more recent sediments due to temporal or geographic biases.

When a substantial level of temporal instability in collection effort is suspected, methods such as that introduced by McCarthy (1998) should be considered. Also known as the "Partial Solow equation", this method represents a modification of the Solow method (Solow, 1993) and uses indices of collection effort made in each discrete time unit (i.e., year). It can be used to estimate the likelihood that a phenomenon has ended, and the likely extinction time, based on the proportion of collection effort made prior to the last known record and the total collection effort, as well as the total number of records.

3.3 **Methods designed for extreme data scarcity**

Archaeological records can be scarce, which can complicate or obstruct application of quantitative methods such as OLE. Overall, sample sizes below five records are not recommended for OLE and other standard methods from this group, and usually cannot be applied to less than three records. However, two extinction inference methods presented here are designed to handle such extreme data scarcity. A non-parametric method introduced by Solow and Roberts (2003), and based on the truncation point estimate by Robson and Whitlock (1964), provides inference of the extinction date based on the timing of the last two records. Here, extinction likelihood simply represents a proportion of the interval between the last two records, and the interval between the second-last record and the end of the observation period. This technique was recently applied to estimate the Lomekwian’s date of origin (Key et al., 2021b).

Roberts and Jarić (2020) further proposed an approach based on the Partial Solow equation (McCarthy, 1998) that is able to infer extinction probability and confidence intervals for phenomena known from only a single record. Using indices of collection effort, extinction likelihood represents a proportion between the collection effort made prior to the only known record and the total effort over the whole collection period (Roberts and Jarić, 2020).

Considering the scarcity of input data, both methods are very conservative, and produce wide confidence intervals. Nevertheless, for such data-deficient situations, these methods often represent the only available quantitative indices for the temporal spans of phenomena, and can be used as just one line of evidence, or as a preliminary screening tool (Roberts and Jarić, 2020). New archaeological phenomena or exceptional finds very rarely preserved (e.g., Lower and Middle Palaeolithic organic technology) would benefit most from these methods.

4. **Discussion and Conclusion**

The adage ‘garbage in, garbage out’ applies equally to optimal linear estimation as it does to any other modelling technique. Following Pettitt and Zilhão’s (2015) example, we wanted to emphasise the assumptions underlying the models detailed here and used elsewhere (Key et al., 2021a, 2021b), and to suggest points of good practice for those wishing to use them. There are no ‘one rule fits all’ approaches to meeting all assumptions, and in our own experience we have taken the decision to slightly modify how we define discrete archaeological occurrences depending on the context (e.g., Bebber and Key, 2021). However, by detailing these suggestions as the methods are introduced, we hope to create an environment where OLE can be applied to the archaeological record with minimal room for misleading and inaccurate results.

Predictive models are, of course, just that; predictions. They can be based on the best available evidence, but it is nevertheless the case that unexpected or unknowable influences could have been acting on the archaeological phenomenon in question, or that key information has been obscured to date due to human-based search biases/effort or taphonomic processes. Future discoveries therefore have the potential to overhaul any predictions made using the techniques outlined here by extending temporal ranges or increasing/decreasing predicted slopes of decline/growth. This should not be viewed as a failure on a part of the model. Instead, it is an invitation to reapply the model using these new data, and to continuously update estimated temporal ranges as new information comes to light.

OLE is an accessible, user-friendly and methodologically transparent temporal range estimation method applied via a single set of equations. It is for these reasons it has become widely used within ecological and palaeontological studies (Rout et al., 2010; Lee, 2014). We therefore hope that OLE and the other techniques described here become accepted and applied within archaeological research. Doing so will allow us to gain a better understanding the temporal presence of past human behaviour.

Acknowledgments

A.K.'s research is supported by the Royal Society. I.J. was supported by J.E. Purkyně Fellowship of the Czech Academy of Sciences.

References

Banks, W.E., Bertran, P., Ducasse, S., Klaric, L., Lanos, P., Renard, C., Mesa, M., 2019. An application of hierarchical Bayesian modeling to better constrain the chronologies of Upper Paleolithic archaeological cultures in France between ca. 32,000–21,000 calibrated years before present. Quat. Sci. Rev. 220, 118–214

Baxter, M.J. and Cool, H.E.M., 2016. Reinventing the wheel? Modelling temporal uncertainty with applications to brooch distributions in Roman Britain. J. Arch. Sci. 66, 120-127

Bayliss, A., 2015. Quality in Bayesian chronological models in archaeology. World Arch. 47 (4), 677–700

Bebber, M.R., Key, A.J.M., in press. Optimal linear estimation (OLE) modelling supports early Holocene (9000-8000 RCYBP) copper tool production in North America. Am. Antiq. DOI: TBC

Bicho, N., Marreiros, J., Cascalheira, J., Pereira, T., Haws, J., 2015. Bayesian modeling and the chronology of the Portuguese Gravettian. Quat. Intern.359-360: 499–509

Buck, C.E. and Meson, B. 2015. On being a good Bayesian. World Arch. 47 (4), 567–584

Boakes, E.H., Rout, T.M., Collen, B., 2015. Inferring species extinction: the use of sighting records. Methods in Ecol. Evo. 6 (6), 678–687

Brook, B.W., Buettel, J.C., Jarić, I., 2019. A fast re-sampling method for using reliability ratings of sightings with extinction-date estimators. Ecology 100 (9), e02787

Buck, C.E., Meson, B., 2015. On being a *good* Bayesian. World Arch.47 (4), 567–584

Clements C., 2013. sExtinct: Calculates the historic date of extinction given a series of sighting events. R package version 1.1. https://cran.r-project.org/src/contrib/Archive/sExtinct/, Accessed 16th Jul 2020

Clements, C.F., Worsfold, N.T., Warren, P.H., Collen, B., Clark, N., Blackburn, T.M., Petchey, O.L., 2013. Experimentally testing the accuracy of an extinction estimator: Solow's optimal linear estimation model. J. Animal Ecol. 82 (2), 345–354

Cowgill, G.L., 1972. Models, methods and techniques for seriation. In: Clarke, D.L. (Ed.) Models in Archaeology. Methuen and Co., London. pp. 381–424

Crema, E.R. 2012. Modelling temporal uncertainty in archaeological analysis. J. Arch. Meth. Theory 19, 440-461

Crema, E.R., Kobayashi, K. 2020. A multi-proxy inference of Jōmon population dynamics using bayesian phase models, residential data, and summed probability distribution of 14C dates. J. Arch. Sci. 117, 105136

Crema, E.R., Bevan, A. 2021. Inference from large sets of radiocarbon dates: software and methods. Radiocarbon 61 (1), 23-39

Fernández-López de Pablo, J., Barton, C.M., 2015. Bayesian estimation dating of lithic surface collections. J. Arch. Meth. Theory 22, 559–583

Jarić, I., Roberts, D.L., 2014. Accounting for observation reliability when inferring extinction based on sighting records. Biodivers. Conserv. 23 (11), 2801–2815

Jordan, P., Cummings, V., 2014. Prehistoric hunter-gatherer innovations. In: Cummings, V., Jordan, P., Zvelebil, M. (Eds.) The Oxford Handbook of the Archaeology and Anthropology of Hunter-Gatherers. Oxford University Press, Oxford. pp. 585–606

Key, A.J.M., Jarić, I., Roberts, D.L., 2021a. Modelling the end of the Acheulean at global and continental levels suggests widespread persistence into the Middle Palaeolithic. Humanities Soc. Sci. Comm. 8, 55

Key, A.J.M, Roberts, D.L, Jarić, I., 2021b. Statistical inference of earlier origins for the first flaked stone technologies. J. Human Evo. 154, 102976

Kolář, J., Macek, M., Tkáč, P., Szabó, P. 2016. Spatio-temporal modelling as a way to reconstruct patterns of past human activities. Archaeometry 58 (3), 513-528

Lee, T.E., McCarthy, M.A., Wintle, B.A., Bode, M., Roberts, D.L., Burgman, M.A., 2014. Inferring extinctions from sighting records of variable reliability. J. Applied Ecol. 51 (1), 251–258

Litton, C.D., Buck, C.E., 1995. The Bayesian approach to the interpretation of archaeological data. Archaeom. 37 (1), 1–24

Long, T., Taylor, D., 2015. A revised chronology for the archaeology of the lower Yangtze, China, based on Bayesian statistical modelling. J. Arch. Sci. 63, 115–121

Lycett, S.J., 2015. Cultural evolutionary approaches to artifact variation over time and space: basis, progress, and prospects. J. Arch. Sci. 56, 21–31

McCarthy, M.A., 1998. Identifying declining and threatened species with museum data. Bio. Conserv. 83 (1), 9–17

McInerny, G.J., Roberts, D.L., Davy, A.J., Cribb, P.J., 2006. Significance of sighting rate in inferring extinction and threat. Conserv. Bio.20 (2), 562–567

Mesoudi, A., 2015. Cultural evolution: a review of theory, findings and controversies. Evo. Biology 43, 481–497

Mesoudi, A., Lycett, S.J., 2009. Random copying, frequency-dependent copying and culture change. Evo. Human Behav. 30 (1), 41–48

Otarola-Castillo, E., Torquato, M.G., 2018. Bayesian statistics in archaeology. Annual R. Anth. 47, 435–453

Pettitt, P., Zilhão, J., 2015. Problematizing Bayesian approaches to prehistoric chronologies. World Arch. 47 (4), 525–542

Pimiento, C., Clements, C.F., 2014. When Did Carcharocles megalodon Become Extinct? A New Analysis of the Fossil Record. PLOS One 9 (1), e111086

Prasciunas, M.M., Surovell, T.A., 2015. Reevaluating the duration of Clovis: The problem of non-representative radiocarbon. In: Smallwood, A.M. and Jennings, T.A. (Eds.) Clovis: On the Edge of a New Understanding. Texas A&M University Press, College Station. pp. 21–35.

Ramsey, C.B., 2015. Bayesian approaches to the building of archaeological chronologies. In: Barcelo, J.A. and Bogdanovic, I. (Eds.) Mathematics and Archaeology. CRC Press, Boca Raton. pp. 272–292

Rivadeneira, M.M., Hunt, G., Roy, K., 2009. The use of sighting records to infer species extinctions: an evaluation of different methods. Ecology 90 (5), 1291–1300

Roberts, D.L., Solow, A.R., 2003. When did the dodo become extinct? Nature 426, 245

Roberts, D.L., Elphick, C.S., Reed, J.M., 2010. Identifying anomalous reports of putatively extinct species and why it matters. Conserv. Bio. 24 (1), 189–196

Roberts, D.L., Jarić, I., 2020. Inferring the extinction of species known only from a single specimen. Oryx 54 (2), 161–166

Robson, D.S., Whitlock, J.H., 1964. Estimation of a truncation point. Biometrika 51 (1/2), 33–39

Rout, T.M., Heinze, D., McCarthy, M.A., 2010. Optimal allocation of conservation resources to species that may be extinct. Conserv. Bio. 24, 1111–1118

Shennan, S., 2015. Demography and cultural evolution. In: Scott, R.A. and Buchmann, M.C. (Eds.) Emerging Trends in the Social and Behavioral Sciences. John Wiley and Sons, Hoboken. pp. 1–14

Solow, A.R., 1993. Inferring extinction from sighting data. Ecology 74 (3), 962–964

Solow, A.R., 2005. Inferring extinction from a sighting record. Math. Biosceinces 195 (1), 47–55

Solow, A.R., Roberts, D.L., 2003. A nonparametric test for extinction based on a sighting record. Ecology 84 (5), 1329–1332

Strauss, D., Sadler, P.M., 1989. Classical confidence intervals and Bayesian probability estimates for ends of local taxon ranges. Math. Geology 21 (4), 411–427

Surovell, T.A., Brantingham, P.J., 2007. A note on the use of temporal frequency distribution in studies of prehistoric demography. J. Arch. Sci. 34 (11), 1868–1877

Surovell, T.A., Finley, J.B., Smith, G.M., Brantingham, P.J., Kelly, R., 2009. Correcting temporal frequency distribution for taphonomic bias. J. Arch. Sci. 36 (8), 1715–1724

Wicks, K., Pirie, A., Mithen, S.J., 2014. Settlement patterns in the late Mesolithic of western Scotland: the implications of Bayesian analysis of radiocarbon dates and inter-site technological comparisons. J. Arch. Sci. 41, 406–422

Wurzer, G., Kowarik, K., Reschreiter, H., 2015. Agent-Based Modeling and Simulation in Archaeology. Springer, Cham

Zhang, H., Jarić, I., Roberts, D.L., He, Y., Du, H., Wu, J., Wang, C., Wei, Q., 2020. Extinction of one of the world's largest freshwater fishes: Lessons for conserving the endangered Yangtze fauna. Sci. Total Enviro. 710, 136242