The Journal of Open Source Software

eratosthenes: Synchronizing archaeological chronologies with a focus on artifact types

Stephen A. Collins-Elliott [1]¶

1 Department of Classics, University of Tennessee, Knoxville, Tennessee, United States ¶ Corresponding author

DOI: 10.xxxxxx/draft

Software

- Review
- Repository
- Archive

Editor: Arfon Smith

Reviewers:

- @sebastien-plutniak
- @joeroe

Submitted: 30 June 2025
Published: unpublished

Summary

Archaeological chronology-building centers on two types of dates, absolute and relative. Absolute dates are associated with a calendrical date (even approximate), and relative refer to events only known via relationships to others (before/after). Relative events are then associated to absolute via contextual or logical associations. E.g, a *terminus post quem* (*t.p.q.*) of a datable coin found in a soil deposit ensures the deposit came after the production of the coin. Applying probabilistic methods to chronology has traditionally centered on radiocarbon dating, but the broader goal of extending formal methods to all aspects of chronology necessitate a way to address artifact typologies. Artifacts largely persist in being dated by qualitative judgment (e.g, "around the start of the 4th century BCE"). Moreover, typologies are in a constant state of revision and adjustment, such that the dates of types may differ according to authority. The reasons why a type is dated the way it is (i.e., which sites, deposits, and comparisons inform its date) can also become opaque, such that investigators inherently resist chronological revisions in order to assess its bibliography (on this conservatism, see Rotroff (2005, p. 20)). This situation presents a challenge for any researcher applying statistical methods to datable artifacts, as one must make *ad hoc* and often awkward decisions, resulting in what Lavan (2021, p. 15) has called "abonimable dates."

The R package `eratosthenes` (named after Eratosthenes of Cyrene, author of the *Chronographiai*) provides functions for chronology-building, above all to bring artifact dating within the scope of formal mathematical estimation. Hence, an investigator can obtain separate probability density functions (p.d.f.) for the production, use, and deposition of an artifact type, in addition to marginal p.d.f.s for all relative sequential events (howsover determined) and absolute constraints (howsoever defined). It uses Gibbs sampling, by now a conventional Markov Chain Monte Carlo method in archaeological chronology (Buck et al., 1996; Geman & Geman, 1984). Using `Rcpp` (Eddelbuettel & Balamuta, 2018), `eratosthenes` performs a two-step Gibbs routine, the first a preliminiary sampler to select a starting date, and a second main sampler that uses consistent batch means (CBM) as a stopping rule, given that convergence in distribution is assured (Flegal et al., 2008; Jones et al., 2006). Full reporting on the Monte Carlo standard errors (MCSE) is provided, giving an error in +/- years for each marginal density. Finally, `eratosthenes` provides functions for assessing the level of dependence of events upon each other within the joint conditional. Changes in dates brought about by any alteration in the structure of a chronology can therefore be readily evaluted. In sum, `eratosthenes` allows for expedient revision of chronologies, transparency in the definition of the full joint conditional, and statistics on the certainty of estimates via MCSE.

Statement of Need

Constructing chronologies via formal means is a major disciplinary goal, with a broad array of software developed to those ends. The CRAN Task View: Archaeological Science maintained by Ben Marwick provides a section on chronological dating software in `R` (R Core Team, 2024). The calibration of ^{14}C dates and the estimation of their posterior probability densities given constraints (Buck & Juárez, 2025) are well served by `OxCal` (Bronk Ramsey, 2009) and `BCal` (Buck et al., 1999), as well as `R` packages `oxcAAR` (Hinz et al., 2021), `Bchron` (Haslett & Parnell, 2008), `c14bazAAR` (Schmid et al., 2019), and `rcarbon` (Crema et al., 2017). General chronological modeling is served by `Chronomodel` (Lanos & Philippe, 2017), which also centers on radiocarbon dating, and `ChronoLog` (Levy et al., 2021), which focuses on discretized time intervals. The R package `ArchaeoPhases` (Philippe & Vibet, 2020) handles post-processing of samples form `OxCal`, `BCal`, and `ChronoModel`. Other chronological software has focused on modeling count data over time, such as `kairos` (Frerebeau, 2025) and `baorista` (Crema, 2025), which focuses on counts of durative/interval events.

Given that there exist substantial software for calibrating radiocarbon dates, `eratosthenes` does not aim to perform this task. Likewise, the formal process of establishing relative sequences of contexts and finds via comparison, called seriation (or ordination), is a computationally difficult problem well served by many `R` packages, such as `seriation` (Hahsler et al., 2008), `vegan` (Oksanen et al., 2024), `boral` (Hui, 2016), `ecoCopula` (Popovic et al., 2022), `VGAM` (Yee, 2004), and `lakhesis` (Collins-Elliott, Under Review).

Two needs which `eratosthenes` aims to satisfy are (1) estimation of artifact dates (production, use, deposition) via formal means and (2) tools for evaluating the dependence of events on each other, which are discussed in detail below.

Estimating Dates of Artifact Production, Use, and Deposition

Dating artifact types is handled by working backwards from depositional contexts, which is their primary point of observation. Estimating the date of artifact production requires a stated rule, with two options given in `eratosthenes`. To start, initial threshold boundaries are established between the earliest depositional context containing that artifact-type and the context immediately prior in sequence lacking it. The `"earliest"` rule samples from within those initial thresholds. The default, `"naive"` rule samples betwen a sample taken using the `"earliest"` rule as the lower bound and an upper bound formed by depositional date of that artifact. After estimating production dates, a use date is sampled between its dates of production and deposition. Figure 1 illustrates these rules.

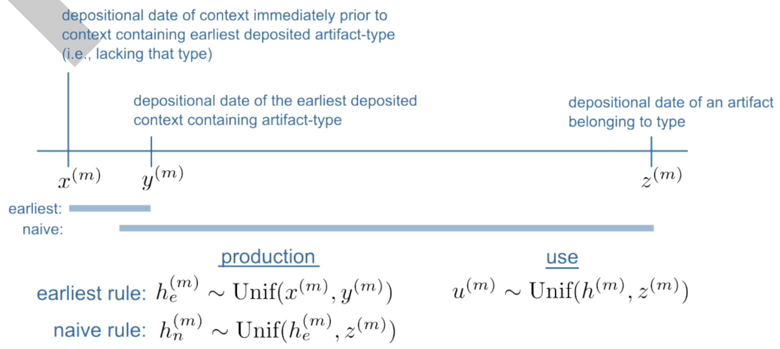

Figure 1: Rules of `earliest` and `naive` (default) for estimating the date of production of an artifact type, sampled on each m iteration of the Gibbs sampler.

Evaluating Displacement of Events

The impact of events on one another is assessed via "displacement," representing the shift in the date of a target event i should another event j be omitted from the joint conditional.

Let $\tilde{x}_i$ be the estimated marginalized Monte Carlo mean date using all events within the joint conditional, and let $\tilde{x}_i^{(-j)}$ be the Monte Carlo mean when event j has been omitted. Squared displacement of j upon i is defined as

$$\delta^2(i,j) = (\tilde{x}_i^{(-j)} - \tilde{x}_i)^2.$$

The omission of j changes the date of i by $\sqrt{\delta^2}$ amount of time (typically years). Conversely, the influence of an event j upon all others is mesaured by mean squared displacement (MSD), which is the mean of all squared displacements when event j has been omitted. Where Θ represents the set of all relative and absolute events, MSD is defined as

$$\text{MSD}(j) = \frac{1}{n-1} \sum_{i \in \Theta, i \neq j} \delta^2(i,j).$$

`eratosthenes` computes the MSD for all events via a "jackknife" or "leave one out"-style of routine, omitting each event iteratively.

Application

For the purposes of illustration, an initial application was undertaken for a small sample of events in the Mediterranean from the last four centuries BCE (data available here). Results are shown in Figure 2 for a selection of two depositional contexts and five shipwrecks. One can note, for example, the effect of using the sack of Carthage in 146 BCE as a *terminus ante quem* (*t.a.q.*) for the depositional date of context Byrsa II B 19.2 at that site (Lancel et al., 1982, p. 194).

The sack of Carthage has also been a key point for dating a particular type of ceramic transport container, the Dressel 1 amphora type. Since it has not be found in pre-destruction layers, the start of its production has been dated to the later part of the 2nd century BCE (Tchernia, 1986, p. 42). Rather than just relying on one site, however, absence from any context within the entire set of conditional events in `eda20250628` yields densities for its production, use, and deposition (Figure 3).

Future Work

Current work in progress by the author which relies upon `eratosthenes` involves the synchronism of ceramics, coinage, radiocarbon dates, depositional/seriated sequences, and historical events for the central Mediterranean in the last four centuries BCE, with datasets uploaded to the GitHub repository eratosthenes-data.

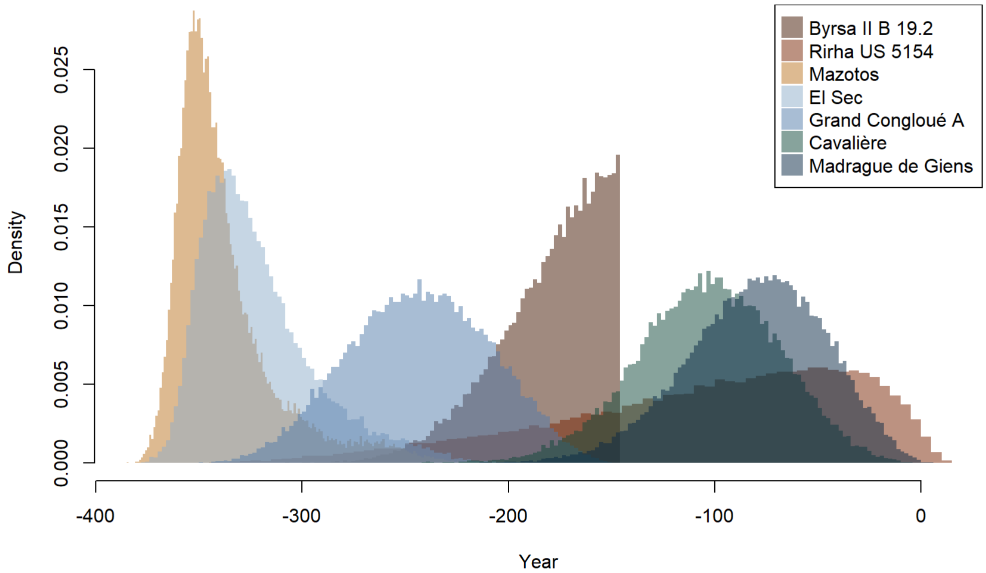

Figure 2: Marginal p.d.f.s of 2 depositional events and 5 shipwrecks from the Mediterraenan, given the joint conditional contained in eda20250628. The wreck Grand Congloué A is earlier than the traditional date: future datasets will work to revise sequencing and constraints.

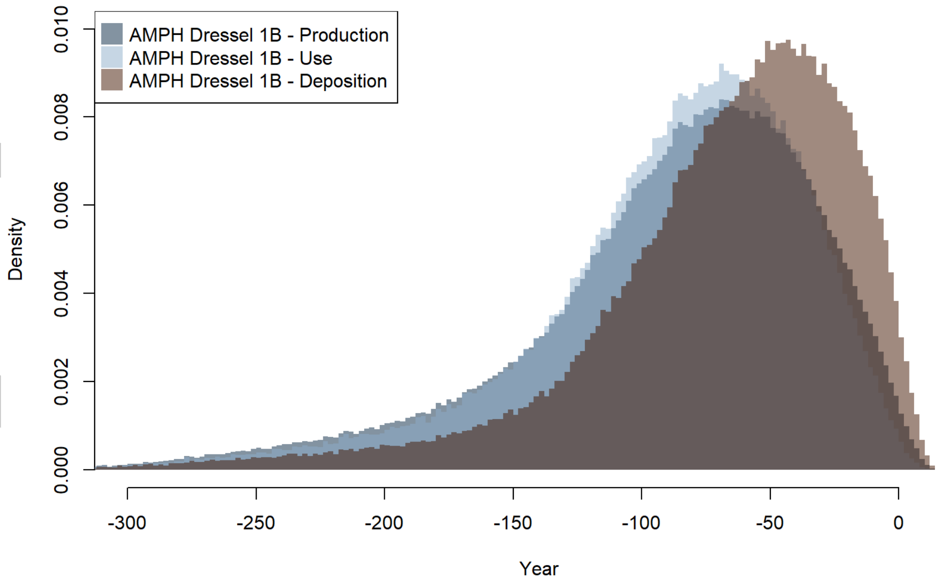

Figure 3: P.d.f.s of the production, use, and deposition of the Dressel 1B type amphora, using the "naive" production rule and given the joint conditional contained in eda20250628.

References

Bronk Ramsey, C. (2009). Bayesian Analysis of Radiocarbon Dates. *Radiocarbon*, *51*, 337–360. https://doi.org/10.1017/s0033822200033865

Buck, C. E., Cavanagh, W. G., & Litton, C. D. (1996). *Bayesian Approach to Interpreting Archaeological Data*. John Wiley; Sons.

Buck, C. E., Christen, J. A., & James, G. N. (1999). BCal: An On-line Bayesian Radiocarbon Calibration Tool. *Internet Archaeology*, *7*. https://doi.org/10.11141/ia.7.1

Buck, C. E., & Juárez, M. (2025). Bayesian radiocarbon modelling for beginners. *Archaeometry*, *67*, 31–48. https://doi.org/10.1111/arcm.12998

Collins-Elliott, S. A. (Under Review). Lakhesis: Consensus Seriation via Iterative Regression of Partial Rankings for Binary Data. *Journal of Applied Statistics*.

Crema, E. R. (2025). A Bayesian alternative for aoristic analyses in archaeology. *Archaeometry*, *67*, 7–30. https://doi.org/10.1111/arcm.12984

Crema, E. R., Bevan, A., & Shennan, S. (2017). Spatio-temporal approaches to archaeological radiocarbon dates. *Journal of Archaeological Science*, *87*, 1–9. https://doi.org/10.1016/j.jas.2017.09.007

Eddelbuettel, D., & Balamuta, J. J. (2018). Extending R with C++: A Brief Introduction to Rcpp. *The American Statistician*, *72*, 28–36. https://doi.org/10.1080/00031305.2017.1375990

Flegal, J. M., Haran, M., & Jones, G. L. (2008). Markov Chain Monte Carlo: Can We Trust the Third Significant Figure? *Statistical Science*, *23*, 250–260. https://doi.org/10.1214/08-STS257

Frerebeau, S. (2025). *Kairos: Analysis of Chronological Patterns from Archaeological Count Data*. https://CRAN.R-project.org/package=kairos

Geman, S., & Geman, D. (1984). Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *6*, 721–741. https://doi.org/10.1016/b978-0-08-051581-6.50057-x

Hahsler, M., Hornik, K., & Buchcta, C. (2008). Getting Things in Order: An Introduction to the R Package seriation. *Journal of Statistical Software*, *25*, 1–34. https://doi.org/10.18637/jss.v025.i03

Haslett, J., & Parnell, A. C. (2008). A Simple Monotone Process with Application to Radiocarbon-Dated Depth Chronologies. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *57*, 399–418. https://doi.org/10.1111/j.1467-9876.2008.00623.x

Hinz, M., Schmid, C., Knitter, D., & Tietze. (2021). *oxcAAR: Interface to 'OxCal' Radiocarbon Calibration*. https://doi.org/10.32614/CRAN.package.oxcAAR

Hui, F. K. C. (2016). Boral: Bayesian Ordination and Regression Analysis of Multivariate Abundance Data in R. *Methods in Ecology and Evolution*, *7*, 744–750. https://doi.org/10.1111/2041-210X.12514

Jones, G. L., Haran, M., Caffo, B. S., & Neath, R. (2006). Fixed-Width Output Analysis for Markov Chain Monte Carlo. *Journal of the American Statistical Association*, *101*, 1537–1547. https://doi.org/10.1198/016214506000000492

Lancel, S., Morel, J.-P., & Thuillier, J.-P. (1982). *Byrsa II: rapports préliminaires sur les fouilles 1977-1978 : niveaux et vestiges puniques*. École française de Rome.

Lanos, P., & Philippe, A. (2017). Hierarchical Bayesian modeling for combining dates in archeological context. *Journal de La Société Française de Statistique*, *158*, 72–88.

Lavan, L. (2021). Checklist of Dates for Pottery Types. In *Public Space in the Late Antique City: Volume 2* (pp. 15–82). Brill. https://doi.org/10.1163/9789004423824_015

Levy, E., Geeraerts, G., Pluquet, F., Piasetzky, E., & Fantalkin, A. (2021). Chronological networks in archaeology: A formalised scheme. *Journal of Archaeological Science*, *127*, 105225. https://doi.org/10.1016/j.jas.2020.105225

Oksanen, J., Simpson, G. L., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., Solymos, P., Stevens, M. H. H., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D., Carvalho, G., Chirico, M., De Caceres, M., Durand, S., … Weedon, J. (2024). *Vegan: Community Ecology Package*. https://doi.org/10.32614/CRAN.package.vegan

Philippe, A., & Vibet, M.-A. (2020). Analysis of Archaeological Phases Using the R Package ArchaeoPhases. *Journal of Statistical Software*, *93*(Code Snippet 1), 1–25. https://doi.org/10.18637/jss.v093.c01

Popovic, G. C., Hui, F. K. C., & Warton, D. I. (2022). Fast Model-Based Ordination with Copulas. *Methods in Ecology and Evolution*, *13*, 194–202. https://doi.org/10.1111/2041-210X.13733

R Core Team. (2024). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. http://www.r-project.org/

Rotroff, S. I. (2005). Four centuries of Athenian pottery. In V. F. Stolba & L. Hannestad (Eds.), *Chronologies of the Black Sea area in the period c. 400-100 BC*.

Schmid, C., Seidensticker, D., & Hinz, M. (2019). c14bazAAR: An R package for downloading and preparing C14 dates from different source databases. *Journal of Open Source Software*, *4*, 1914. https://doi.org/10.21105/joss.01914

Tchernia, A. (1986). *Le vin de l'Italie romaine. Essai d'histoire économique d'après les amphores*. École Française de Rome.

Yee, T. W. (2004). A New Technique for Maximum-Likelihood Canonical Gaussian Ordination. *Ecological Monographs*, *74*, 685–701. https://doi.org/10.1890/03-0078