Modeling Temporal Uncertainty in Historical Datasets

Vojtěch Kaše[1,2,*], Adéla Sobotková[2] and Petra Heřmánková[2]

[1]Department of Philosophy, University of West Bohemia, Czech Republic
[2]Department of History and Classical Studies, Aarhus University, Denmark

Abstract

This paper explores several approaches to assess temporal trends within archaeological and historical datasets containing records marked with significant extent of uncertainty accompanying their dating. We evaluate the strengths and pitfalls of these methodologies by employing two datasets: one comprising ancient shipwrecks and the other ancient Greek inscriptions. While these objects can, in principle, be precisely dated to specific years, they are often assigned broader date ranges, spanning centuries or longer historical periods. We propose that the most promising approaches involve using these date ranges as defining probabilities. By randomly assigning specific dates based on these probabilities, we enable hypothesis testing for temporal trends. As we want to encourage other scholars to employ the methods we propose, we offer a detailed description of the implementation of these methods using functions from the Python *tempun* package.

Keywords

temporal uncertainty, historical data analysis, archaeological data analysis, Monte Carlo simulations, computational history,

1. Introduction

In this paper, we introduce three computational approaches to deal with temporal uncertainty in historical datasets for the purposes of exploratory data analysis and hypothesis testing. By historical datasets we mean datasets consisting of data points representing objects or events from the past, such as collections of material artifacts, corpora of textual sources, or lists of historical events studied by archaeologists or historians. Such datasets are often marked by some level of uncertainty accompanying dating of individual data points, as creators or curators of the datasets are not able to say precisely when an object was produced or when an event occurred. In what follows, we will focus on one particular form of such uncertainty, which is called *within-phase uncertainty* [1]. In this case, instead of dating an event by means of specific date (e.g. assigning it to a specific year), the dating is expressed by means of a range or interval (e.g. with reference to a century).

Our aim is to demonstrate the strengths and pitfalls of the three most commonly applied approaches and to show how they can be employed in a straightforward manner using the Python programming language and the *tempun* package [2]. We will place particular emphasis

CHR 2023: Computational Humanities Research Conference, December 6 – 8, 2023, Paris, France
*Corresponding author.
kase@kfi.zcu.cz (V. Kaše); adela@cas.au.dk (A. Sobotková); petra.hermankova@cas.au.dk (P. Heřmánková)
https://vojtechkase.cz (V. Kaše)
0000-0002-6601-1605 (V. Kaše); 0000-0002-4541-3963 (A. Sobotková); 0000-0002-6349-0540 (P. Heřmánková)

on the third and most complex approach, which relies on Monte Carlo simulation, as we consider it the most promising, especially for hypothesis testing. To facilitate our comparisons of the three approaches, we will utilize two example datasets: the database of ancient Mediterranean shipwrecks up to 1500 CE, compiled by Strauss and digitized from Parker's previous work [3, 4], and the GIST dataset of ancient Greek inscriptions [5]. However, since our primary focus is on methodology, we will not conduct in-depth data analysis or test specific hypotheses using these datasets in the present paper.

2. The example datasets

The shipwreck database is available in a spreadsheet format with 1,784 records and 16 attributes. These attributes include details about the provenance, destination, and cargo of the ships, among other information. For our topic, two attributes are of crucial importance: "post_0" and "ante_0," which represent the starting and ending points of the estimated temporal interval during which the shipwreck occurred.

The GIST (Greek Inscriptions in Space and Time) database of ancient Greek inscriptions is also available in a tabular format and covers 217,863 records with 29 column attributes. The GIST dataset is primarily based on an online collection of ancient Greek inscriptions published by the Packard Humanities Institute (PHI) and available via the Ithaca project as I.PHI [6]. A key feature of the dataset is that, wherever available, the spatial and temporal metadata are expressed in a machine-readable form. Thus, in the case of temporal information, the estimated date of production is again expressed by means of two attributes, called here 'not_before' and 'not_after', having the same meaning as 'post_0' and 'ante_0' in the case of the shipwrecks; both delineate the dating interval. For instance, the date of origin of an inscription which has been in the original collection dated to the 5th century BCE (e.g. PH174) is now in a machine-readable form expressed by means of an interval with limit points -500 and -401.

While in the case of shipwrecks the interval expresses a time range within which it is estimated that the wreckage occurred, in the case of inscriptions it is an interval within which the inscription was erected. In both cases, the interval captures a one time event, which might be in principle dated to a singular year, or even a day. Ultimately, the annual temporal resolution is probably the most appropriate one for working with data from the ancient Mediterranean context.

A first issue to explore in the case of a dataset with dating information encoded by means of an interval is the distribution of durations of these intervals. Figure 1 introduces histograms of durations for both datasets. In the case of inscriptions, we see that there is a higher ratio of inscriptions within the two leftmost bins reserved to the shortest durations up to 15 years. These are the most precisely dated inscriptions in the dataset. In the case of shipwrecks, we see that such precise dates are rather exceptional.

It has been proposed that regardless of the issues with temporal uncertainty, the datasets of shipwrecks and inscriptions offer valuable information concerning economic and cultural development of the ancient Mediterranean world. Both have been used as proxies for economic trends, as a higher amount of shipwrecks or inscriptions dated to a certain period are generally considered to indicate higher intensity of trade or higher overall economic performance of the

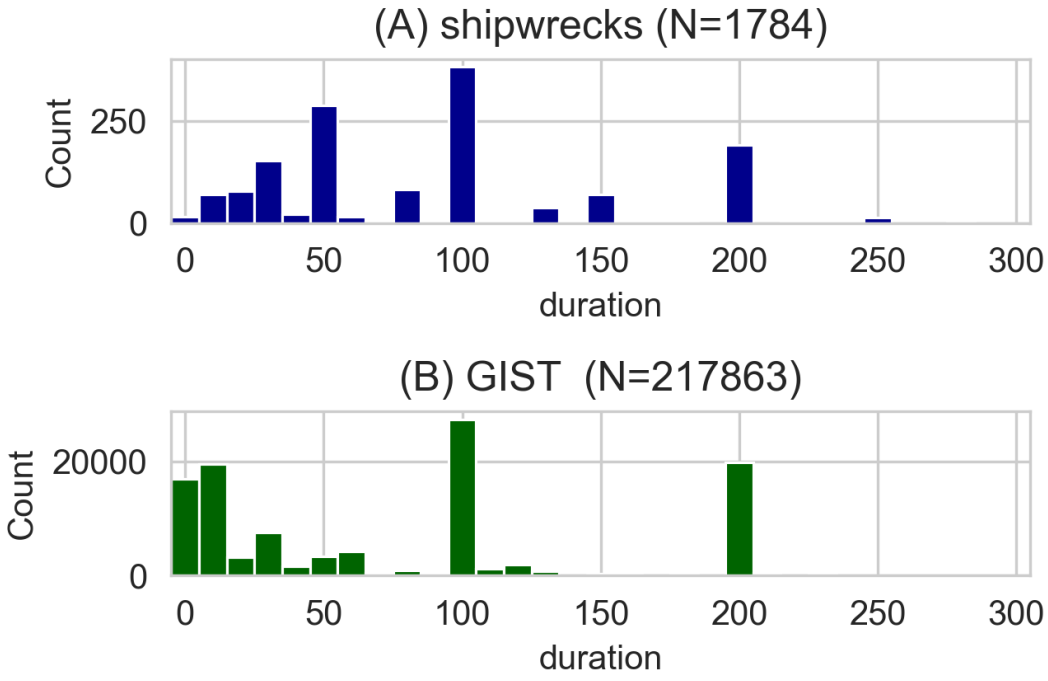

Figure 1: Distributions of durations of dating intervals for individual records. The le□most bin covers durations of up to 5 years. All other bins cover 10 years. For example, the sixth bin from the le□ represents records with dating range duration between 46 and 55 years (It is a half-century whether it is numerically expressed as being formed by 49, 50, or 51 years, as there are o□en inconsistencies in how these dates are numerically expressed in digitized datasets. For instance, while sometimes an object dated to the 1st half of the 2nd century CE is coded by the interval (101,150) with duration of 50 years, in other cases we can find it translated as (100-150), which is 51 years. Our bins guarantee that both these variants will be within the same bin).

area under scrutiny in this period in comparison to a period from which we have a smaller number of records within these databases.

Thus, for instance, Hopkins relied on an older version of the shipwrecks' dataset to argue that "in the period of Roman imperial expansion and in the High Empire (200 BC – AD 200), there was more sea-borne trade in the Mediterranean than ever before, and more than there was for the next thousand years" [7, pp. 105-106] and concluded "that trade in the third century A.D. declined" [7, pp. 115] (for the same argument, see also [8]). Similarly, Kaše and Glomb [**?**] argued that the temporal distribution of ancient Greek inscriptions can serve as a useful proxy for the economic development in the eastern Mediterranean region. But how do we evaluate such claims using a dataset with dates expressed by means of dating ranges?

All the below introduced analyses and visualizations were implemented using the Python 3 programming language [9], relying especially on the *tempun* package. *tempun* is a specialized library designed to efficiently handle temporal uncertainty in historical datasets. We make our

example analyses available for reuse by other scholars via Github: https://github.com/sdam-au/tempun_demo.

3. Midpoints

Perhaps the most straightforward approach to analyze temporal trends in historical data characterized by a substantial extent of within-phase temporal uncertainty is the one based on midpoint values, i.e. arithmetic mean of the two numeric values defining the range. Thus, an inscription dated to the 3rd century BCE, which is expressed by means of an interval delimited by the values -300 and -201, is treated as being dated to the year 250 BCE. Such dates are finally used within some temporal distribution plots using stable bins such as centuries, half centuries, quarter centuries etc.

This approach has at least one serious issue: In case that a dataset contains a lot of data points dated on the century basis, you will bias the dataset by overestimating the number of objects dated to the middle of centuries. Further, the dataset might also contain a substantial amount of data points dated with reference to historical periods such as "Roman Imperial period", conventionally delimited by an interval from 31 BCE to 410 CE and covering a period longer than four hundred years with mid-point at 190 CE. Again, using the mid-range values will substantially bias the data; in this case it will cause overestimation of the number of records dated to the end of the 2nd century CE.

One way to diminish the impact of the bias caused by extensive and imprecise ranges is to exclude all data points with temporal range (duration) greater than a certain threshold (e.g. 100 or 50 years). This approach might be useful for specific applications, but it implies a substantial information loss, as broadly dated records get completely omitted, while we should expect that they still somehow contribute to the general pattern.

Figure 2 demonstrates this approach while clearly revealing its limitations. The figure comprises four histograms showing midpoint dates of records from the GIST dataset. Subplots (A) and (B) are based on a bin width of 50 years, while subplots (C) and (D) use a narrower bin width of 25 years. In subplot (A), we depict all records within the dataset that have a properly delimited dating range (i.e., both "not_before" and "not_after" values are available). On the other hand, subplot (B) employs the same dataset but includes only data points with dating ranges shorter than 100 years, excluding inscriptions dated on a centurial or longer basis. As a result, this filtering reduces the size of the GIST dataset to approximately half of its original size.

Subplots (C) and (D) offer a more granular overview of the temporal distribution of the data due to their narrower histogram bins of 25 years each. Subplot (D) further restricts the data by excluding inscriptions with intervals equal to or longer than 50 years, keeping only about 40 % of all dated inscriptions. Across the subplots, temporal trends show notable differences.

By comparing subplots (A) and (B) and subplots (C) and (D), we can identify entirely distinct patterns in the data from the 2nd and 3rd centuries CE. During this period, ancient Greek inscriptions were primarily produced within the Eastern part of the Roman Empire, which experienced its peak economic performance, followed by a decline starting at some point during the third century, exhibiting substantial regional differences. It is, therefore, not surprising that the 2nd century CE contains the highest number of dated epigraphic records. However,

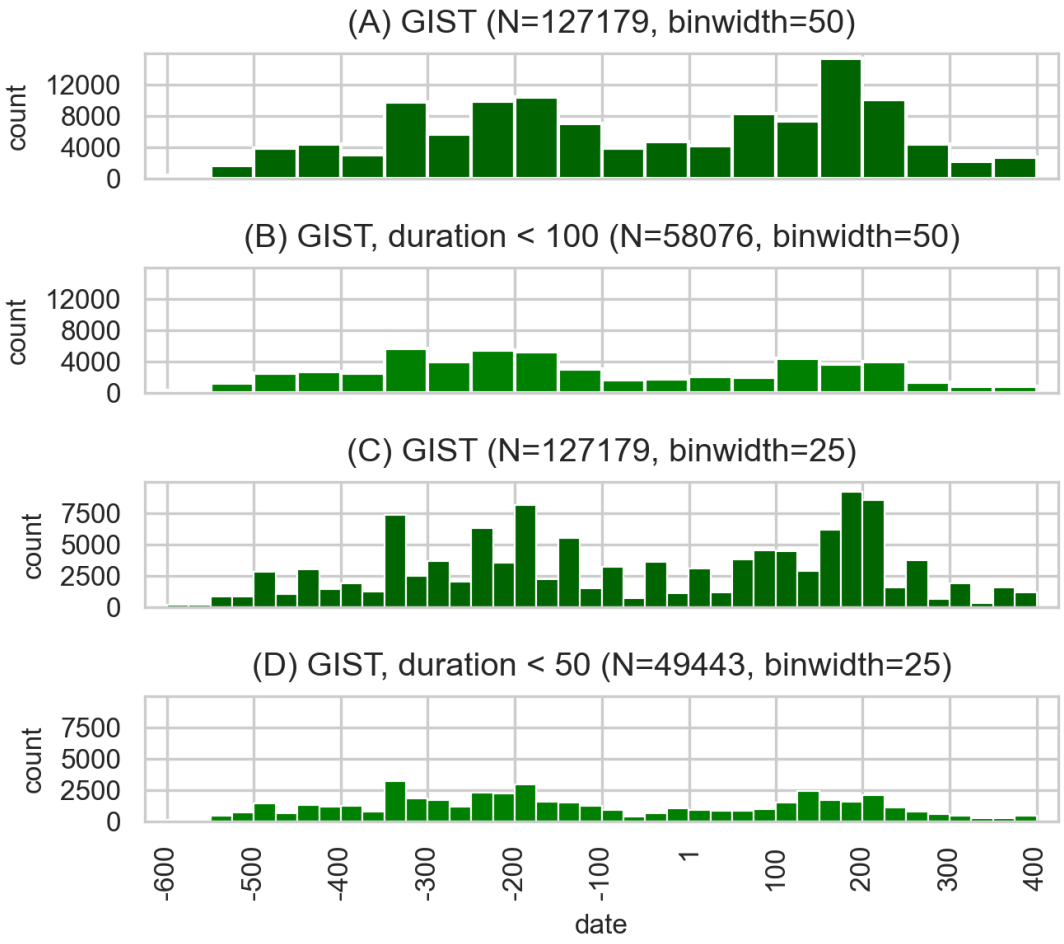

Figure 2: Histograms of temporal distribution of GIST using the midpoints-based approach. Subplots di□er in the length of the bins (subplots (A) and (B) vs (C) and (D)) and in filtering of records within certain duration threshold (subplots (B) and (D)).

while subplot (A) suggests a significantly higher number of inscriptions from the second half of that century, subplot (B) presents a reverse trend. This indicates that the peak in (A) was, to some extent, an artificial product of midpoints derived from widely dated inscriptions. Drawing on the higher granularity of subplot (C), one might easily form a false impression that a more nuanced picture of the period is obtained. However, this picture becomes entirely blurred when compared with subplot (D).

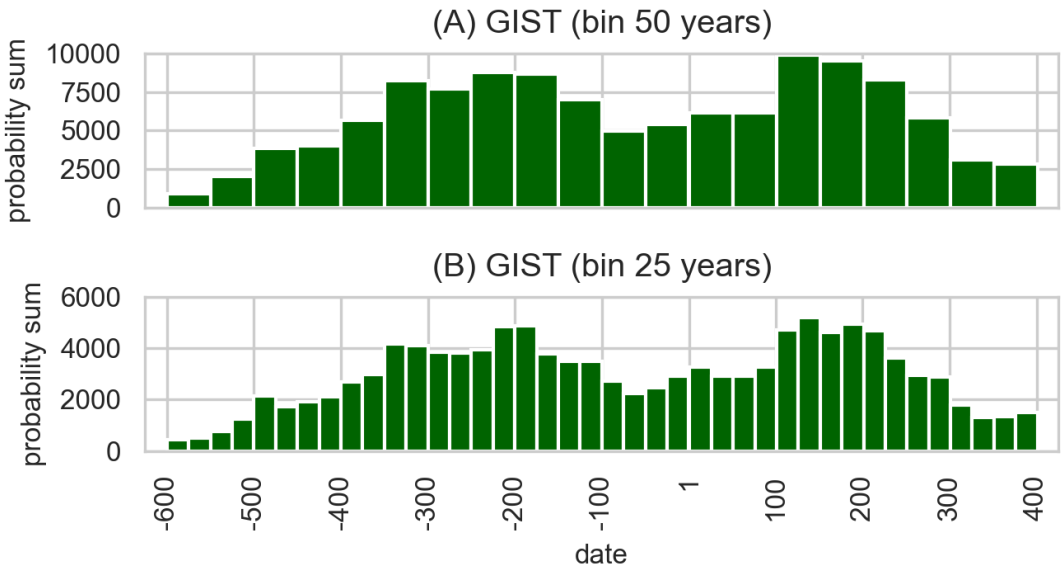

Figure 3: Aoristic sum based histograms of the GIST dataset employing di□erent binwidth.

4. Aoristic Sum

An alternative, which overcomes some of the aforementioned limitations, is the Aoristic Sum (AS) based approach, developed originally in criminology and geography [10] and later on adopted in archaeology (see [11]). AS converts date ranges into probabilities, providing a more nuanced representation of temporal distribution. For instance, an inscription dated to the 3rd and 4th centuries CE (i.e., numerically ranging from 201 to 400) would have a 0.5 probability of being erected in either of the two centuries, a 0.005 probability of being produced within any singular year within this range, and a 0 probability of being produced outside of this range. Essentially, for each record within the dataset, we generate its unique probability distribution based on the dating range. We then plot the overall cumulative distribution by combining the probability distributions from all individual records. In other words, first, for each time bin we sum up probabilities of all records within it and, second, plot these sums for all bins in form of a histogram.

In Figure 3 we see that the temporal distribution is now rather smooth. Focusing on the 2nd century CE, we see much more balanced values across the bins on both subplots than in the case of the midpoints. Thus, by avoiding the biases listed above, a plot based on AS is a valuable contribution to the visual exploration of the data. But how do we draw on this approach when proceeding further to cross-data comparisons or hypothesis testing? This is where the Monte Carlo Simulation approach deserves our attention.

5. Monte Carlo Simulations

The Monte Carlo Simulation (MCS) approach to address within-phase temporal uncertainty in historical datasets [12] relies on the same assumptions as the Aoristic Sum method, treating date ranges as expressing probabilistic distributions. However, MCS derives these distributions as secondary outcomes of random number generation, rather than directly calculating them from the data. This technical distinction requires a more detailed explanation. In the following sections, we will demonstrate the efficacy of the MCS approach by exploring our two example datasets using specific functions from the *tempun* package, as *tempun* provides a range of useful functions for both the AS and MCS methodologies.

In the case of MCS, the initial step involves drawing on the dating intervals to generate random dates within them in the format of singular years. This task is accomplished using the function `model_date()`, which requires two mandatory parameters: "start" to specify the beginning of the dating interval, and "stop" to indicate the end of the interval. Additionally, there are several optional parameters that can be utilized for specific requirements, such as "size", specifying the number of the random dates we attempt to generate, or "seed", guaranteeing that we will obtain identical results once we again execute the command.

By default, the random number generator underlying the `model_date()` function follows a uniform distribution. This means that all numbers within the interval have an equal probability of being generated, which aligns with one of the shared assumptions of the AS approach. This characteristic becomes evident when we set the size parameter to a high value (e.g., 10,000) and subsequently plot the distribution of the resulting data. In a standard setting, we generate a fixed number of random dates for all records within a dataset in a single step. In the code snippet Listing 1, we apply the `model_date()` function to all records from the GIST dataset. The GIST dataset is loaded into the Python environment as a dataframe object, with the "not_before" and "not_after" column attributes defining the dating interval for each record. The `model_date()` function processes the values of these attributes for each row record and assigns to each of them 100 random dates using the "size" parameter. To ensure reproducibility of our analyses, we also specify the seed for each row using their numerical ID inherited from PHI.

```
GIST["random_dates"] = GIST.apply(lambda row: tempun.model_date(start=row["
    not_before"], stop=row["not_after"], size=100, seed=row.index), axis=1)
```

Listing 1: Assignment of 100 random dates to each record within the GIST dataset

The resulting data is saved into a new dataframe column called "random_dates", where each row entry is represented by a list of 100 numbers. Collectively, the data in this column can be viewed as a matrix. In this object-date matrix, row vectors correspond to individual observations within the source dataset, while columns correspond to individual simulated date variants. For example, in Table 1, we display a few of the 100 modeled dates from a random sample of 5 dated inscriptions from the GIST dataset. Additionally, for reference, we include the PHI identifier of each record, its original verbal dating, and the "not_before" and "not_after" attributes.

Now, focusing exclusively on the random dates data from the entire dataset and treating them as a matrix, we can proceed to various forms of analysis of temporal trends within the data. First and foremost, each of the 100 column vectors in the matrix can be regarded as a time series on its own, with its unique temporal distribution. Our primary question is to what extent

Table 1
GIST sample: 5 records with random dates

PHI_ID	raw_date	not_before	not_a□er	random_dates
218434	Rom. Imp. period	-31.0	410.0	[306, 386, 82, 123, 7, 189, 99, 119, 80, 206, 1...
12996	s. II p.	101.0	200.0	[171, 136, 125, 169, 198, 197, 131, 128, 109, 1...
299245	2nd/3rd c. AD	101.0	300.0	[144, 119, 159, 116, 151, 225, 146, 151, 190, 1...
294030	Hellenistic period	-330.0	-30.0	[-158, -325, -64, -204, -280, -268, -52, -35, -...
207841	ca. 250-225 BC	-252.0	-223.0	[-237, -232, -243, -231, -247, -231, -230, -225...

the individual temporal distributions are similar to each other, as this will reflect the impact of temporal uncertainty accompanying dating of individual records on the overall temporal trends. To address this question, we can plot all the histograms cumulatively and observe differences within individual time blocks. However, bar-style histograms, commonly used for this type of visualization, may not be the most suitable choice. Instead, line-style histograms appear to be more useful in this regard.

With the help of the *tempun* package, you can easily generate a line-style cumulative histogram plot directly from the "random_dates" attribute of a dataframe dataset. This can be achieved by utilizing the `timeblocksplot_from_randoms()` function. Unlike plotting separate line histograms for each time series, this function consolidates all data within each time bin. Within each bin, the 2.5 % upper and lower margins are visually indicated using gray coloring, while all values within the 95 % interval are represented by a user-defined color value.

Figure 4 allows a straightforward visual comparison of the three aforementioned techniques for exploring the diachronic distribution of historical datasets with range-defined temporal information, using the Ancient Shipwrecks dataset as an example. In Figure 1(A), we have already demonstrated that this dataset exhibits a significant extent of temporal uncertainty, with a relatively high number of records lacking precise dating. In Figure 4(A) we apply the midpoints-based approach without any filtering threshold for vaguely dated wreckage events. Therefore, we can expect that these records significantly bias the overall distribution. This becomes visible when we apply the Aoristic Sum methodology, as we do in Figure 4(B). This reveals an overall decline in the number of documented shipwrecks starting in the 1st half of the 3rd century CE rather than its later half, as suggested by the unfiltered midpoints data. This observation echoes those made by Wilson [13], who used a previous version of the same shipwreck dataset together with the aoristic approach to highlight weaknesses in interpretations based on midpoints' distribution visualizations.

In Figure 4(C), we move a step forward, as we adopt the MCS timeblocks visualization method described above, putting together temporal distributions of 100 simulated time series. Unlike the AS methodology, this approach allows us to directly detect the periods most affected by temporal uncertainty in the underlying data. For instance, in Figure 4(B), there appears to be a higher number of shipwrecks from the 1st half of the 1st century CE than from its later half. This trend is apparent in subplot Figure 4(C) as well, but here the fat lines with gray margins also show that at least in case of some simulated time series data, the trend is in fact reversed. As we show below, this might be subjected to hypothesis testing.

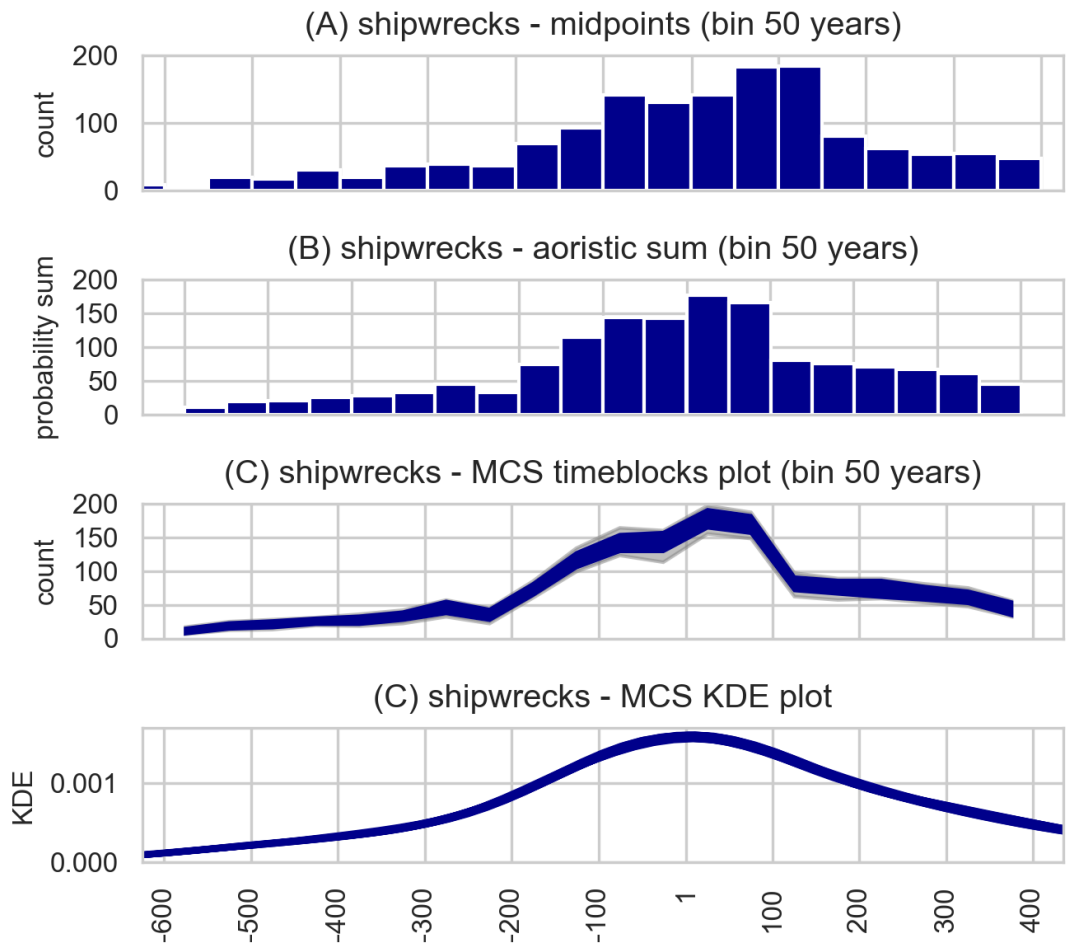

Figure 4: Overview of di□erent visualization approaches using the shipwreck dataset.

Finally, on Figure 4(D), we employ yet another visualization function of the simulated time series data implemented in the *tempun* package: the `kdeplot_from_randoms()` function. This visualization produces a cumulative kernel density estimate plot with each time series represented by one line. The bandwidth of the KDE is calculated automatically on the basis of Scott's rule [14]. As these lines are highly similar in this case, the final product looks like one fat line. This method allows us to inspect trends which escape the bins employed in the previous cases.

The visualizations above fall under the umbrella of exploratory data analysis, but perhaps the most promising possibility of employing the MCS approach is in the context of hypothesis testing. For instance, we can ask whether the temporal distribution of the simulated time series from one subset of the data differs significantly from temporal distribution characterizing the rest of the dataset or another subset of the data. Thus, applying the MCS approach on a dataset of Latin Inscriptions of the Roman Empire (LIRE) [15], Glomb et al ([16] focused on the

distribution of a subset of inscriptions naming the god Asclepius while testing the hypothesis that these inscriptions (and hence the religious healing cult they are related to) was especially thriving during the period of the so called Antonine Plague (dated approximately to 165 - 180 CE), as proposed by a number of scholars (e.g. [17]). First, Glomb at el. tested whether the distribution of the inscriptions naming Asclepius differs significantly from a distribution of a random control sample of inscriptions of the same size and with the same proportions of individual inscription types (e.g. epitaphs, votive inscriptions etc.). For that purpose, they employed the two-sample Kolmogorov-Smirnov test (2sKS) [18] to compare against each other 1,000 pairs of simulated time series data from the both groups of inscriptions and calculated the average *KS statistic* and *p* values. The test revealed that the temporal distribution of inscriptions naming Asclepius was not significantly different from the distribution of its control sample.

Another option is to test statistically - in face of the temporal uncertainty within the underlying data - whether a number of records from a certain period is higher than a number of records from another period. On the most basic level, we can calculate the proportion of cases in which the trend goes in one direction. Thus, returning back to the GIST dataset, comparing statistically the data from the bins for the the first half and the second half of the 1st century CE, we could find that the number of inscriptions assigned to the earlier bin is higher in case of 68 out of 100 simulated time series. In other words, there is a non-negligible probability equal to 0.32 that the trend went in the opposite direction (c.f. Figure 5(A)). Accordingly, we should hesitate to formulate any strong claims concerning the distributional trend within this period. But, once we move one century forward and repeat the same test, we see that the number of inscriptions from the earlier half of the 2nd century is higher in 100 % of cases. Thus, there is a discernible decline starting at this point.

Finally, this logic might be further elaborated while focusing on some internal properties of the data, such as word frequencies. It is in this respect where the advantage of MCS over AS is most apparent. Thus, using the LIRE dataset of ancient Latin inscriptions from the Roman Empire, Kaše et al. [19] analyzed the changing frequencies of occupation terms on the inscriptions. They first mapped the time series simulation data back on the inscriptions. For each time-series variant and for each of the 50-years-long timeblocks from 50 BCE to 350 CE, they randomly sampled 1000 inscriptions (with replacement), and then calculated the number of preselected occupation terms mentioned on these inscriptions. For visualizations of the resulting data they again used the cumulative line-style histogram. This analysis relied on the `sim_data_by_function()` and `plot_timeblocks_data()` functions from the *tempun* package. In Figure 5(B), we see this same approach applied on the GIST dataset, inspecting the changing frequency of the Greek term *dikaiosynē* (righteousness). As above, this approach might be combined with hypothesis testing.

6. Conclusion

In this paper, we introduced three most commonly applied approaches for dealing with within-phase temporal uncertainty in historical datasets. The first introduced approach, based on the midpoint values, is the most straightforward one, but it is also associated with some substantial issues, as it tends to bias the data. The approach based on Aoristic Sum is free of the biggest

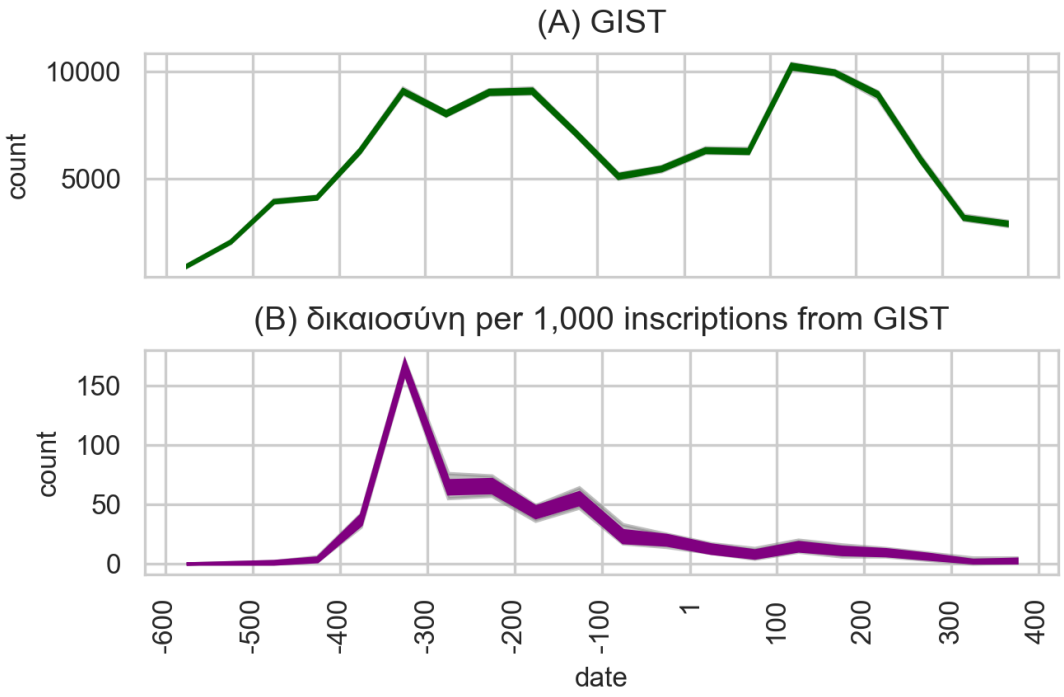

Figure 5: (A) Temporal distribution of GIST using the MCS dates and (B) temporal distribution of the term dikaiosynē (righteousness) within the GIST dataset relying on the same simulated dates and the same visualization approach.

problems of the midpoints' based approach, but has its own limitation, as it does not allow us to identify in a straightforward fashion the periods most affected by temporal uncertainty. Finally, the Monte Carlo approach allows us to move forward with hypothesis testing, even on the level of secondary features within a dataset (such as word frequencies in the case of inscriptions). Our implementation of the Aoristic Sum and MCS approach, including production of the figures, relied heavily on the *tempun* Python package. Furthermore, we demonstrate the usefulness of the *tempun* package for the purposes of hypothesis testing in the studies referred to above. All our examples were elaborated within one Jupyter notebook script which we also share publicly for future reuse by other scholars.

Acknowledgments

This research was funded by the Aarhus University Forskningsfond Starting grant no. AUFF-E-2018-7-22 awarded to the 'Social Complexity in the Ancient Mediterranean' (SDAM) project.

References

[1] E. R. Crema, K. Kobayashi, A Multi-Proxy Inference of Jōmon Population Dynamics Using Bayesian Phase Models, Residential Data, and Summed Probability Distribution of 14C Dates, J. Archaeol. Sci. 117 (2020) 10536. doi:10.1016/j.jas.2020.105136.

[2] Vojtěch Kaše, tempun (v0.2.2), 2023. URL: https://doi.org/10.5281/zenodo.8179346. doi:10.5281/zenodo.8179346.

[3] J. Strauss, Shipwrecks Database (v1.0), 2013. URL: http://oxrep.classics.ox.ac.uk/databases/shipwrecks_database/.

[4] A. J. Parker, Ancient Shipwrecks of the Mediterranean & the Roman provinces, Tempus Reparatum, Oxford, 1996.

[5] V. Kaše, P. Heřmánková, GIST (v0.2), 2022. URL: https://doi.org/10.5281/zenodo.7185509. doi:10.5281/zenodo.7185509.

[6] T. Sommerschield, Y. Assael, B. Shillingford, M. Bordbar, J. Pavlopoulos, M. Chatzipanagiotou, I. Androutsopoulos, J. Prag, N. de Freitas, I.PHI Dataset: Ancient Greek Inscriptions, 2021. URL: https://github.com/sommerschield/iphi.

[7] K. Hopkins, Taxes and Trade in the Roman Empire (200 B.C.-A.D. 400), The Journal of Roman Studies 70 (1980) 101–125.

[8] F. de Callataÿ, The Greco-Roman Economy in the Super Long-run: Lead, Copperand Shipwrecks, Journal of Roman Archaeology 18 (2005) 361–372.

[9] G. Van Rossum, F. L. Drake, Python 3 Reference Manual, CreateSpace, Scotts Valley, CA, 2009.

[10] J. H. Ratcliffe, Aoristic Analysis: The Spatial Interpretation of Unspecific Temporal Events, International Journal of Geographical Information Science 14 (2000) 669–679. doi:10.1080/136588100424963.

[11] I. Johnson, Aoristic Analysis: Seeds of a New Approach to Mapping Archaeological Distributions through Time, in: K. F. Ausserer, W. Börner, M. Goriany, L. Karlhuber-Vöckl (Eds.), [Enter the Past] the E-way into the Four Dimensions of Cultural Heritage: CAA2003, number 1227 in BAR International Series, Archaeopress, Oxford, 2004, pp. 448–452.

[12] E. R. Crema, Modelling Temporal Uncertainty in Archaeological Analysis, Journal of Archaeological Method and Theory 19 (2012) 440–461. doi:10.1007/s10816-011-9122-3.

[13] A. Wilson, Approaches to Quantifying Roman Trade, in: A. Bowman, A. Wilson (Eds.), Quantifying the Roman Economy: Methods and Problems, Oxford University Press, Oxford - New York, 2009, pp. 213–249.

[14] D. W. Scott, Multivariate Density Estimation: Theory, Practice, and Visualization, John Wiley & Sons, 2015.

[15] V. Kaše, P. Heřmánková, A. Sobotková, LIRE (v1.0.0), 2021. URL: https://doi.org/10.5281/zenodo.5074774. doi:10.5281/zenodo.5074774.

[16] T. Glomb, V. Kaše, P. Heřmánková, Popularity of the Cult of Asclepius in the Times of the Antonine Plague: Temporal Modeling of Epigraphic Evidence, Journal of Archaeological Science: Reports 43 (2022) 103466. doi:10.1016/j.jasrep.2022.103466.

[17] G. E. v. d. Ploeg, The Impact of the Roman Empire on the Cult of Asclepius, Brill, Leiden ; Boston, 2018.

[18] J. L. Hodges, The Significance Probability of the Smirnov Two-Sample Test, Arkiv för

Matematik 3 (1958) 469–486. doi:10.1007/BF02589501.
[19] V. Kaše, P. Heřmánková, A. Sobotková, Division of Labor, Specialization and Diversity in the Ancient Roman Cities: A Quantitative Approach to Latin Epigraphy, PLOS ONE 17 (2022) e0269869. doi:10.1371/journal.pone.0269869.