

Portland State University

PDXScholar

Anthropology Faculty Publications and
Presentations

Anthropology

11-21-2017

Radiocarbon Test for Demographic Events in Written and Oral History

Kevan Edinborough
University College London

Marko Porčić
University of Belgrade

Andrew Martindale
University of British Columbia

Thomas J. Brown
University of British Columbia

Kisha Supernant
University of Alberta

See next page for additional authors

Follow this and additional works at: https://pdxscholar.library.pdx.edu/anth_fac



Part of the [Archaeological Anthropology Commons](#)

Let us know how access to this document benefits you.

Citation Details

Edinborough, K., Porčić, M., Martindale, A., Brown, T. J., Supernant, K., & Ames, K. M. (2017). Radiocarbon test for demographic events in written and oral history. *Proceedings of the National Academy of Sciences*, 201713012.

This Article is brought to you for free and open access. It has been accepted for inclusion in Anthropology Faculty Publications and Presentations by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.

Authors

Kevan Edinborough, Marko Porčić, Andrew Martindale, Thomas J. Brown, Kisha Supernant, and Kenneth M. Ames



Radiocarbon test for demographic events in written and oral history

Kevan Edinborough^{a,1}, Marko Porčić^b, Andrew Martindale^c, Thomas Jay Brown^c, Kisha Supernant^d, and Kenneth M. Ames^e

^aInstitute of Archaeology, University College London, London WC1H 0PY, United Kingdom; ^bDepartment of Archaeology, Faculty of Philosophy, 11000 Belgrade, Serbia; ^cDepartment of Anthropology, University of British Columbia, Vancouver, BC, Canada V6T 1Z1; ^dDepartment of Anthropology, University of Alberta, Edmonton, AB, Canada T6G 2H4; and ^eDepartment of Anthropology, Portland State University, Portland, OR 97207

Edited by James A. Brown, Northwestern University, Evanston, IL, and approved August 21, 2017 (received for review July 25, 2017)

We extend an established simulation-based method to test for significant short-duration (1–2 centuries) demographic events known from one documented historical and one oral historical context. Case study 1 extrapolates population data from the Western historical tradition using historically derived demographic data from the catastrophic European Black Death/bubonic plague (*Yersinia pestis*). We find a corresponding statistically significant drop in absolute population using an extended version of a previously published simulation method. Case study 2 uses this refined simulation method to test for a settlement gap identified in oral historical records of descendant Tsimshian First Nations communities from the Prince Rupert Harbour region of the Pacific Northwest region of British Columbia, Canada. Using a regional database of $n = 523$ radiocarbon dates, we find a significant drop in relative population using the extended simulation-based method consistent with Tsimshian oral records. We conclude that our technical refinement extends the utility of radiocarbon simulation methods and can provide a rigorous test of demographic predictions derived from a range of historical sources.

historical record | oral history | archaeology | simulation | radiocarbon

We extend an established simulation-based method to test for significant regional-scale demographic events known from documentary historical and oral historical sources. Simulation-based models based on real archaeological datasets are proving increasingly useful for identifying population-related changes in archaeological contexts (1–3). Such approaches offer a far more rigorous statistical assessment of a given demographic question than was previously possible (4).

Well-deployed simulation-based demographic approaches have two main strengths. First, data simulation can potentially account for the ubiquitous archaeological problem of finite, small sample sizes that diminish over time (5–7). Second, because simulation-based approaches can avoid qualitative assessments of patterns within summed probability distributions (SPDs), they can mitigate the thorny issue of confirmation bias. This problem is one long recognized by psychologists, wherein the influence of a favored hypothesis inadvertently biases the choice of data and model selected by a researcher (8, 9). The converse issue is one of rejection bias, where researchers reject an unfavorable model out of hand without adequately considering or even replicating it (10).

Here we attempt to explicitly avoid these biases and encourage more researchers to follow our lead when using SPDs as a proxy for demographic signatures. We extend the methodological reach of a widely cited simulation-based demographic method (1, 4). Then we test this method against the historically well-documented population decline in 14th-century Europe that was caused by the bubonic plague (*Yersinia pestis*), or Black Death (11). We find support for this particular simulation-based approach using the established (known) data of this historical context following previously contested concerns about this approach raised by an earlier study (4, 10). Using a newly collated radiocarbon dataset containing 523 results, we then apply a more

conservative version of the same simulation method to test for a shorter-duration demographic-settlement gap, known from the oral-historical record in Tsimshian territory in the Prince Rupert Harbour Region of northern British Columbia, Canada (12). The results of this test suggest that significant drops in relative population identified using the simulation-based method are also consistent with Tsimshian oral records. This paper presents the rigorous testing of an indigenous oral record against demographic data derived from a statistically robust model. The absence of such tests is a common criticism of the use of Indigenous oral records in archaeology (13, 14). As we find support for demographic events extrapolated from both oral and historical records, we conclude that these simulation-based demographic models are consistent with other lines of evidence, which suggests that our results have considerable explanatory power. To encourage more researchers to use this approach, we include the associated freeware R code and data and a summarized explanation of the methods in *Supporting Information*.

As the methods used here are advancing apace, the research lineage of our particular simulation approach is important to note. The following methodological progression is underpinned by the fundamental belief that population dynamics can be

Significance

Indigenous oral traditions remain a very controversial source of historical knowledge in Western scientific, humanistic, and legal traditions. Likewise, demographic models using radiocarbon-based simulation methods are controversial. We rigorously test the historicity of indigenous Tsimshian oral records (*adawx*) using an extended simulation-based method. Our methodology is able to detect short-duration (1–2 centuries) demographic events. First, we successfully test the methodology against a simulated radiocarbon dataset for the catastrophic European Black Death/bubonic plague (*Yersinia pestis*). Second, we test the Tsimshian *adawx* accounts of an occupational hiatus in their territorial heartland ca. 1,500–1,000 years ago. We are unable to disconfirm the oral accounts. This represents the first formal test of indigenous oral traditions using modern radiocarbon modeling techniques.

Author contributions: K.E., M.P., A.M., T.J.B., K.S., and K.M.A. designed research; K.E., M.P., A.M., T.J.B., K.S., and K.M.A. performed research; K.E. and M.P. contributed new reagents/analytic tools; K.E., M.P., A.M., T.J.B., K.S., and K.M.A. analyzed data; and K.E., M.P., A.M., T.J.B., K.S., and K.M.A. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Published under the PNAS license.

Data deposition: All extant radiocarbon data used in [Datasets S1](#) and [S2](#) have been deposited in the Canadian Archaeological Radiocarbon Database, www.canadianarchaeology.ca/. This database is an open access compilation of radiocarbon measurements that indicate the ages of samples primarily from archaeological sites in North America.

¹To whom correspondence should be addressed. Email: k.edinborough@ucl.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1713012114/-DCSupplemental.

recovered from the archaeological record, given a sufficient observed sample of dated human activity. Our position is that, while this sample itself may be a skewed approximation of true population levels and dynamics, our results will reflect the underlying population signal if they meet the strenuous criteria set by sufficiently rigorous methodological protocols.

Uncalibrated radiocarbon dates have long been used as evidence around the world for inferring general human settlement patterns (15–17), and this paper builds out of this approach. These tentative First Order approaches always come with stated cautions and caveats. Recently, given the increasing availability of computer power, the promise of the approach has encouraged a controversial demographic turn in archaeology (18–21). Initially, uncalibrated radiocarbon data were simply collated from subsets of a defined geographical region of archaeological interest and then summed over one to produce a temporally coarse-grained histogram, a time series of the relative intensity of uncalibrated radiocarbon data (17). After a comprehensive radiocarbon analysis of a well-excavated prehistoric region of southern Scandinavia by Edinborough (20, 22, 23), Shennan and Edinborough (24) summed and calibrated discrete bins of archaeological radiocarbon results from across northern and central Europe to produce a broad-scale calibrated population model spanning selected parts of the Neolithic transition there. Radiocarbon dates were binned in this way into archaeologically determined units, or phases, to avoid inadvertent sampling biases caused by oversampling of specific sites or periods. Collard et al. (19) developed the method further, summing the calibrated archaeological radiocarbon date bins, or phases, and producing a demographic boom-and-bust model for the Neolithic transition of Great Britain. The potentially confounding effects of exponential human growth rates (4, 25, 26), and archaeological site formation processes producing a general exponential taphonomic loss over time of archaeological data (6, 7), necessitated a further refinement of this method. The most sophisticated method that accounts for research bias, taphonomic loss, and the long-term population trends was developed by Shennan et al. (1). The research bias is reduced by the specific binning procedure, which gives equal weight to site phases with differential numbers of dates. To account for the effects of taphonomy and long-term population growth on the empirical curve, an exponential model was fitted to the empirical curve by regression. The resulting exponential model is used as a null model against which the empirical SPD is statistically evaluated.

Methods

To assess the statistical significance of the deviation of the empirical curve from the null model, a large number of simulated radiocarbon datasets are generated by randomly sampling calendar dates from the specified time interval according to the probabilities given by the null model. The number of dates for each simulated dataset is equal to the number of bins in the

empirical dataset. The sampled calendar dates are “back calibrated” by simulating a radiocarbon date that might have produced the particular calendar date. The back-calibrated dates are then recalibrated and summed. This procedure is repeated several 1,000 times to create a distribution of simulated values for each moment in time.

To assess the statistical significance of the empirical SPD pattern, the empirical curve is compared with the 95% percentile intervals calculated from the simulated data for each year. For time intervals where the empirical summed calibrated probability distribution is above or below the simulated 95% confidence intervals (CI), there is a statistically significant growth or decline, respectively, of population relative to the null model. Given that in 5% of cases the curve will be outside of the 95% CI limits even if the underlying population dynamics are identical to the null model, false-positive results are identified through a global significance statistic. This is calculated by first transforming both empirical and simulated probability density values into Z scores in relation to the simulated distribution for each time unit. Z scores outside the 95% CI are then summed both for the empirical and the simulated curves. The empirical sum of Z scores is compared with the distribution of summed Z scores from simulated datasets. The global significance value is the relative frequency of simulated Z score sums, which are equal to or greater than the empirical value. Recent progress in this particular research lineage now allows formal comparison of entire regional radiocarbon assemblages using different datasets (for example, from different areas of Jomon culture in Japan) that also produce global significance tests, so that interregional demographic models can be critically assessed and productively compared (27).

As it is, the Shennan et al. method [henceforth the University College London (UCL) method] tests for the departure of the empirical SPD curve from the null model SPD curve by simulating SPD curves from the null model and constructing confidence intervals for each point in time. However, this method cannot tell whether a difference in the values of the SPD between the two points on the empirical curve is significant relative to the null model when it comes to differences in the shape of the curve or parts of the curve. For example, if the true population scenario looked like the one in Fig. 1, *Left*, the UCL method would pick up the general deviation from the null model. The uniform null model is used here for simplicity, because there are no taphonomic effects given that these are simulated data. If the uniform model is applied to a set of 350 randomly simulated radiocarbon dates from the hypothetical population model, it would not be able to tell us whether the changes in the part of the curve that is already outside the confidence intervals are significant (Fig. 1, *Right*); the original method does not detect the small trough in the high population zone (vertical difference between A and B in Fig. 1, *Right*). Likewise, the method would not be able to detect the subtleties of the situation shown in Fig. 2, where 350 dates are sampled from the underlying hypothetical model and summed. The SPD curve is consistent with the null model when it comes to the range of variation for each calendar year; however, we know that the shape of the true underlying model is different from that of the uniform model. Despite this, we would not be able to detect a significant drop in the curve between points A and B in Figs. 1 or 2 using only the original UCL method.

A simple extension of the original UCL method is proposed to resolve this. The main idea of this refinement is that the significance of the relative changes in the SPD curve can be tested by statistically comparing the difference between two points on the empirical SPD curve to the distribution of differences between the points with the same coordinates in calibrated time

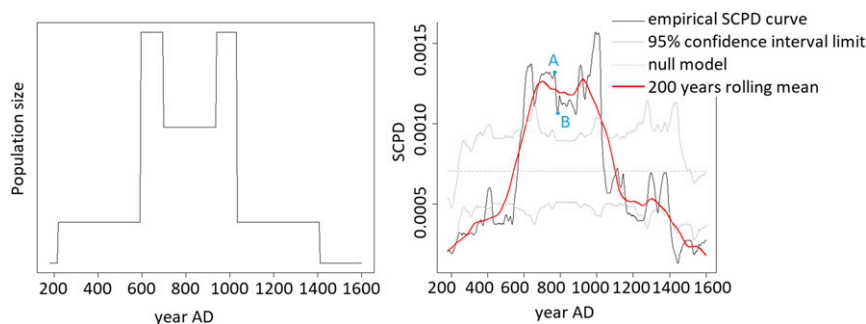


Fig. 1. If the true population scenario looked like the one on the *Left*, the UCL method would pick up the general deviation from the null model. If applied to a set of 350 randomly simulated radiocarbon dates from the hypothetical population model, the UCL method would not be able to tell us whether the changes in the part of the curve that is already outside the confidence intervals are significant (*Right*). SCPD, summed calibrated probability distribution.

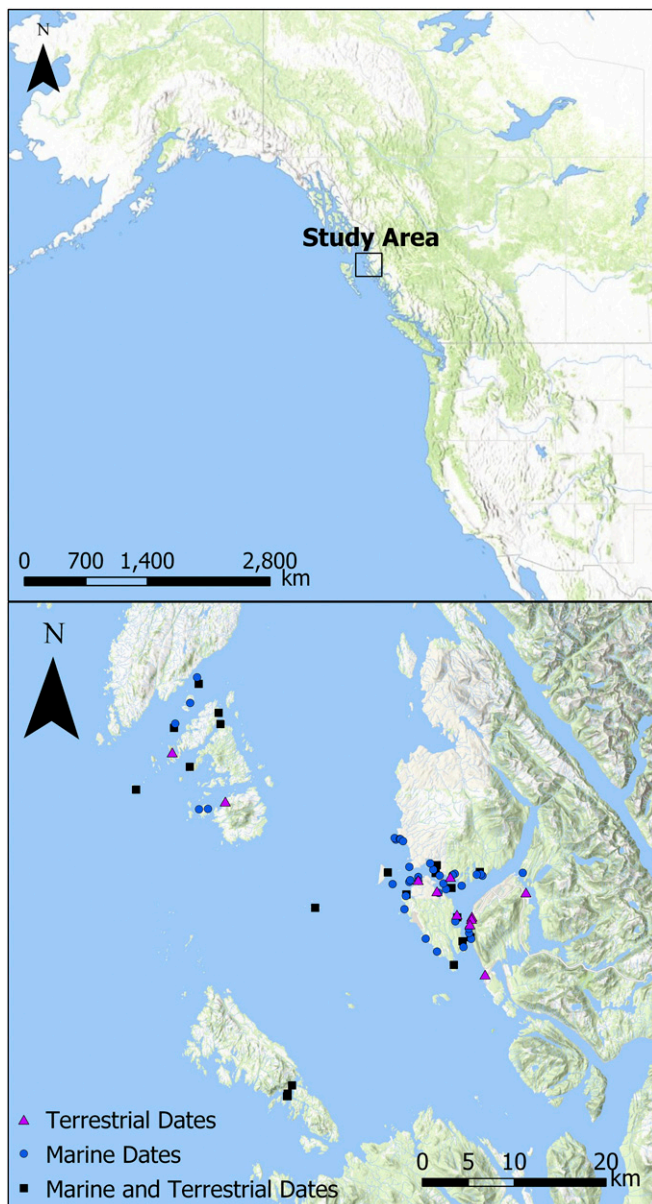


Fig. 4. The Prince Rupert Harbour area showing archaeological sites with terrestrial and marine-based radiocarbon samples.

population growth trend to be a separate demographic phenomenon to be discovered.

The red line in Fig. 5 shows a general trend of the real data by fitting a rolling 200-y average to the real data (the black line). The interval between $\sim 2,800$ and 1,500 cal B.P. remains outside of the expected confidence range, which suggests $\sim 1,300$ y of a large yet fluctuating relative population before a significant demographic drop in the region starting somewhere between 1,800 and 1,100 cal B.P. Although we do see an ~ 200 -y gap at ca. 1,000 cal B.P. (95% confidence interval delineated by the two solid gray lines) with a significant general downward trend in population, interpretative caution is required. Assuming a single marine reservoir effect (MRE) value for the entire marine radiocarbon dataset may be problematic if it insufficiently accounts for all local ΔR variation in the dataset. Although we are confident that this value is accurate for the last ca. 5,000 y following the most recent conservative calculation of a local ΔR at the site of Kitandach (273 ± 38) (35), interpretations of marine-based radiocarbon results remain variable and potentially problematic in this region, even when using the most rigorously calculated MRE values with the latest radiocarbon methods (35). Furthermore, lack of calibration data present in the smoother Marine 13 curve (34) compared with its terrestrial counterpart obscures smaller features and

smooths SPD results, despite the larger radiocarbon sample size ($n = 336$) used in this case.

Fig. 6 results are generated again using 100-y data bins for the available Prince Rupert Harbour terrestrial radiocarbon data. The real radiocarbon data crosses (D), or is marginally close to, the 95% CI (B) of the simulated data (the solid gray line) in two places. Thus, we find two candidate occupation gap horizons (B and D) indicated by this method. The global P value is highly significant ($P = 0.0095$), indicating that deviations of the empirical curve from the null model are greater than chance. The extended method shows that both gaps may be significant as the differences between points A and B and C and D are statistically significant at the 0.05 level (with Bonferroni correction the threshold would be 0.025; significance values associated with A–B and C–D differences are below this value, 0.0049 and 0.0091, respectively). The more recent gap, $\sim 1,200$ –1,000 cal B.P. at D, is in broad agreement with the results of our OxCal radiocarbon model detailed above, so we suggest that this is the best candidate gap for correspondence with other lines of material evidence for the hiatus described in the oral record (12). Additionally, the earlier gap is more likely to be a result of sampling bias (Supporting Information). Our preferred gap model of $\sim 1,200$ –1,000 cal B.P. is consistent with our Marine sample SPD model, as only the later gap (at D) persists in both datasets.

Discussion

There is wide consensus that demographic patterns are potentially visible in radiocarbon data if the data are representative of historical trends. In archaeological contexts with smaller numbers of dates, the UCL method provides a means of assessing demographic trends via a comparison between the actual data

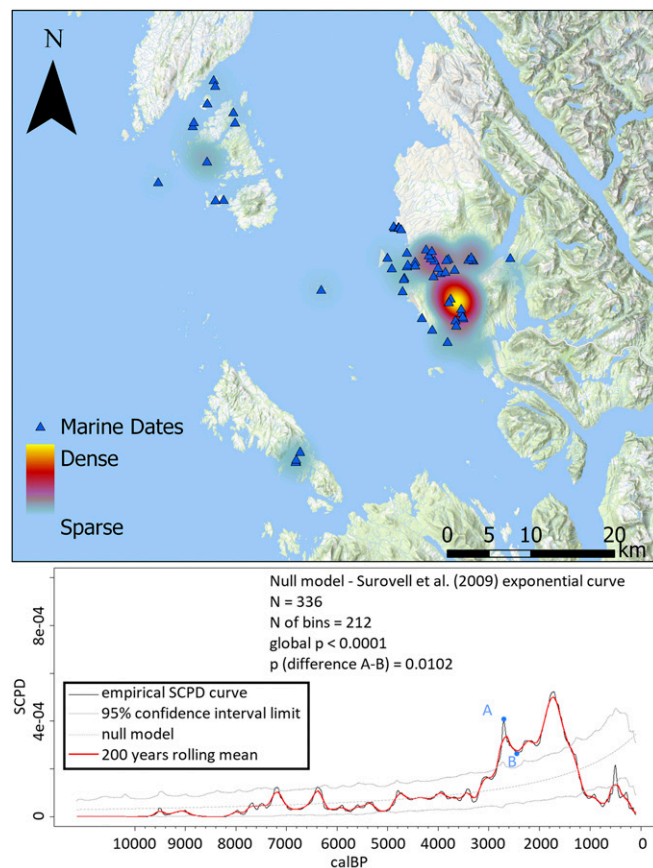


Fig. 5. Prince Rupert Harbour area (Top) with an illustrative kernel density heat map showing both distribution and relative intensity of marine-based radiocarbon results. (Bottom) Prince Rupert Harbour marine-based radiocarbon data summed with extended UCL method with 100-y data bins. Points A and B show a significant drop outside the 95% confidence intervals. SCPD, summed calibrated probability distribution.

