

# Dendrochronological dates confirm a Late Prehistoric population decline in the American Southwest derived from radiocarbon dates

Erick Robinson<sup>1,\*</sup>    Kyle Bocinsky<sup>2,3</sup>    Darcy Bird<sup>4</sup>    Jacob Freeman<sup>1</sup>    Robert L. Kelly<sup>5</sup>

## Abstract

The northern American Southwest provides one of the most well-documented cases of human population growth and decline in the world. The geographic extent of this decline in North America is unknown due to the lack of high-resolution palaeodemographic data from regions across and beyond the greater Southwest, where archaeological radiocarbon data is often the only available proxy for investigating these palaeodemographic processes. Radiocarbon time series across and beyond the greater Southwest suggest widespread population collapses from AD 1300–1600. However, radiocarbon data have potential biases caused by variable radiocarbon sample preservation, sample collection, and the non-linearity of the radiocarbon calibration curve. In order to be confident in the wider trends seen in radiocarbon time series across and beyond the greater Southwest, here we focus on regions that have multiple palaeodemographic proxies and compare those proxies to radiocarbon time series. We develop a new method for time series analysis and comparison between dendrochronological data and radiocarbon data. Results confirm a multiple proxy decline in human populations across the Upland US Southwest, Central Mesa Verde, and Northern Rio Grande from AD 1300–1600. These results lend confidence to single proxy radiocarbon-based reconstructions of paleodemography outside the Southwest that suggest post-AD 1300 population declines in many parts of North America.

**Keywords:** dates-as-data, radiocarbon summed probability distributions, dendrochronology, palaeodemography, southwestern United States.

<sup>1</sup> Utah State University

<sup>2</sup> Crow Canyon Archaeological Center

<sup>3</sup> University of Montana

<sup>4</sup> Washington State University

<sup>5</sup> University of Wyoming

\* Correspondence: Erick Robinson <[erick.robinson@usu.edu](mailto:erick.robinson@usu.edu)>

## Introduction

The spatial and temporal ubiquity of archaeological radiocarbon dates makes them an unprecedented proxy for human paleodemography. The past decade has witnessed a rapid increase in the use of archaeological radiocarbon time series to reconstruct prehistoric population growth and decline in various regions of the world (Bevan et al., 2019; Crema et al., 2016; Goldberg et al., 2016; Palmisano et al., 2019; Peros et al., 2010; Russell and Steele, 2009; Shennan et al., 2013; Wang et al., 2014; Williams, 2012). One of the first aggregations of radiocarbon dates for palaeodemographic analysis was by Berry (Berry, 1982) in his study of population growth, decline and migration across the American Southwest. Radiocarbon time series were used as a complement to earlier periods that lacked robust dendrochronological records. Berry (1982: 120) noted that they were “not even remotely related to population size”, but rather the “relative probability of occupation through time” and “the major trends in these charts (i.e., peaks and troughs) should provide a reliable indication of the direction of change”. Later, Rick (1987: 56) more formally outlined the use of radiocarbon time series as a method for paleodemography by developing an ‘inference chain’ that detailed the various biases involved in their use and interpretation. Rick identified three biases: creation bias, preservation bias, and investigation bias (Rick, 1987). Over the past decade, as archaeologists have assembled and analysed large radiocarbon datasets from various regions of the world, there has been considerable focus on these biases, notably creation (Crombé and Robinson, 2014; Freeman, Byers, et al., 2018; Naudinot et al., 2014) and preservation bias (Bluhm and Surovell, 2019; Peros et al., 2010; Surovell and Brantingham, 2007; Surovell et al., 2009; Williams, 2012), as well as bias caused by the radiocarbon calibration process (Armit et al., 2013; Bamforth and Grund, 2012; Contreras and Meadows, 2014; Michczynski and Michczynska, 2006; Timpson et al., 2014; Williams, 2012). Of the three biases Rick identified, investigation bias has been the most understudied thus far. Investigation bias occurs when investigators gather more radiocarbon samples for particular periods of time or specific geographic regions, which leads to periods or regions with more or less radiocarbon representation than the average time period or region. The biased selection of radiocarbon samples would then appear to be relative increases or decreases in population when they are merely artefacts of research history and fieldwork practice. Investigation bias has been understudied because for many regions of the world radiocarbon data are the only currently available proxy evidence enabling reconstructions of paleodemography in continuous time series spanning millennia. Assessing the impact of investigation bias on radiocarbon-based reconstructions of paleodemography requires comparison with other paleodemographic proxies.

The American Southwest provides one the best available case studies for assessing the role of investigation bias in radiocarbon-based reconstructions of paleodemography. It also has one of the most well-documented cases of population growth and decline in the world (Bocinsky et al., 2016; Freeman, Baggio, et al., 2018; Hill et al., 2004; Kohler et al., 2010; Robinson et al., 2019). For almost a century, researchers have collected multiple proxies, such as dendrochronological records, settlement size records, and ceramic seriation records to address the magnitude of this population growth and decline. However, a critical knowledge gap still exists for the exact scale of population decline across and beyond the Southwest (Boyer et al., 2010; Cameron, 2010). This gap is because the less intensively researched regions lack multiple proxies for robust reconstructions of paleodemography. In these regions, radiocarbon data are the only available proxy time series for reconstructing paleodemography. Over the past six years, we have collected all available radiocarbon data across and beyond the Southwest (Robinson et al., 2019). Initial analyses of this new radiocarbon dataset have revealed extensive population decline from AD 1300–1600 across and beyond the Southwest (Freeman, Baggio, et al., 2018; Robinson et al., 2019). Therefore, the geographic extent of Late Prehistoric population decline could be much broader than previously expected.

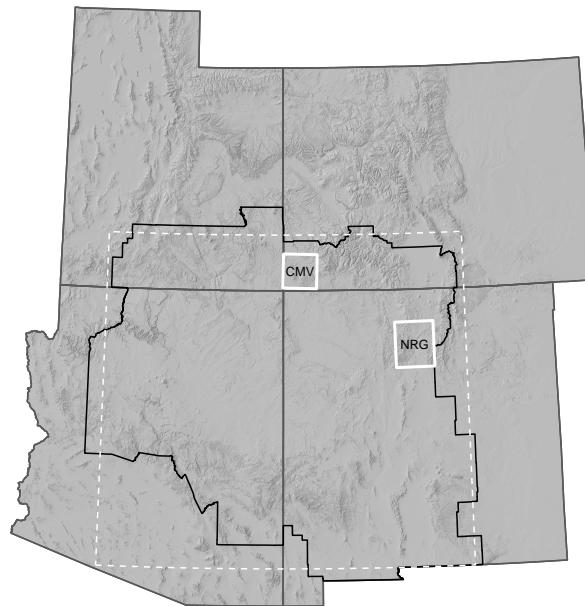
The potential importance of this widespread population decline for research on Late Prehistoric archaeology of the American West requires comparison of the sensitivity of lower-resolution radiocarbon data against other higher-resolution palaeodemographic proxies. In the most well-documented regions (such as the northern American Southwest) we expect radiocarbon records to reflect the highest amount of investigation bias, as researchers have the option of selecting from a suite of other higher-resolution proxies, including dendrochronology and ceramic seriation. Investigation bias can create an artificial ‘edge effect’ spuriously appearing as a decrease in population. Here we develop a new method for arraying dendrochronological records as time series for comparison against radiocarbon time series. We focus on the broader region of the Upland US Southwest (UUSW) for which there is a robust dendrochronological record (Bocinsky et al., 2016).

We also focus on the more concentrated and intensively investigated Central Mesa Verde (CMV) and Northern Rio Grande (NRG) regions of the UUSW (Figure 1), for which multiproxy palaeodemographic reconstructions were produced by the Village Ecodynamics Project (VEP) (Ortman, 2016; Schwindt et al., 2016). The VEP is an interdisciplinary, multi-method research initiative aimed at better understanding demography of ancestral Pueblo farmers and human-environment interaction across the UUSW. As part of the project, VEP researchers cataloged data on all known archaeological habitation sites within a 4,600-km<sup>2</sup> area in the CMV (Schwindt et al., 2016) and a 6,900-km<sup>2</sup> area in the NRG (Ortman, 2016), to which most of the CMV population migrated during the thirteenth century AD (Ortman, 2016). They integrated data including pitstructure and room count estimates, the presence of diagnostic architectural features, surface ceramic tallies, surveyor assessments, and limited dendrochronological data in an empirical Bayesian framework to develop population estimates for each region resolved to periods ranging from 20 to 125 years in length (Ortman, 2016; Schwindt et al., 2016). The VEP estimates are the highest quality regional palaeodemographic reconstructions currently available in the UUSW, and are completely independent of the radiocarbon records for the region. They are therefore useful in exploring the impact of investigation bias in the radiocarbon record.

We expect that substantial investigation bias is present in radiocarbon records if those records substantially diverge from the dendrochronological and multiproxy reconstructions. If the lower-resolution radiocarbon records reveal similar trends to the two other records, we expect investigation bias does not substantially impact the ability of radiocarbon records to be used as relative measures of the directionality of population change through time. This would, in turn, lend more confidence to radiocarbon records across the American West that reveal extensive post AD 1300 population declines but lack other palaeodemographic proxies.

## Methods

Here we compare two large collections of radiocarbon and tree-ring dates drawn from the UUSW. After harmonizing those datasets with one another and calibrating the radiocarbon dates, we calculate binned summed probability distributions across the UUSW, CMV, and NRG. The radiocarbon dates we analyse here are drawn from the National Science Foundation supported project, “Populating a Radiocarbon Database for North America,” which have been uploaded to the Canadian Archaeological Radiocarbon Database (CARD; <https://www.canadianarchaeology.ca>). Tree-ring data are



**Figure 1:** The UUSW as defined in this study (region with black border). The white dashed area represents the UUSW as defined by Bocinsky et al. (2016). VEP study areas are white boxes. CMV: Central Mesa Verde; NRG: Northern Rio Grande.

those presented by Bocinsky et al. (2016). Both datasets are available on the Digital Archaeological Record (see [Data Accessibility](#)).

## Study Area

The UUSW was geographically defined by Bocinsky et al. (2016) as the region between 105–113°W and 32–38°N. Here, we refine that definition in two ways. First, because location data for our radiocarbon database are only resolved to the county level for most sites (Robinson et al., 2019), we define our study region to include all counties whose geographic centroids fall within the region from Bocinsky et al. (2016). Second, because of the generally poor tree-ring date record in the Sonoran Desert, we exclude the Phoenix and Tucson basins by dropping Maricopa, Pinal, and Pima counties in Arizona. Figure 1 shows the original UUSW from Bocinsky et al. (2016) and revised study area used in this analysis, as well as the two VEP study areas.

## Period of Interest

Here, we focus our analyses on the period for which data are most abundant: 1750–150 BP, or AD 200–1800. This period includes the periods of the VEP population estimates for the CMV (AD 600–1280) and NRG (AD 900–1760), and extends 400 years earlier than the CMV estimates to allow us to assess trends in the radiocarbon record leading up to the VEP reconstructions, and to control for edge effects. While the VEP demographic estimates are for ancestral Pueblo populations only, it is likely that the tree-ring records include samples from Spanish colonial contexts after AD 1600 or so.

## Datasets

**Table 1:** Counts of radiocarbon and tree-ring dates and sites in this study.

Study Area	Radiocarbon		Tree-ring	
	Dates	Sites	Dates	Sites
UUSW Total	1,531	482	15,077	770
CMV	36	18	4,435	219
NRG	83	40	2,306	146

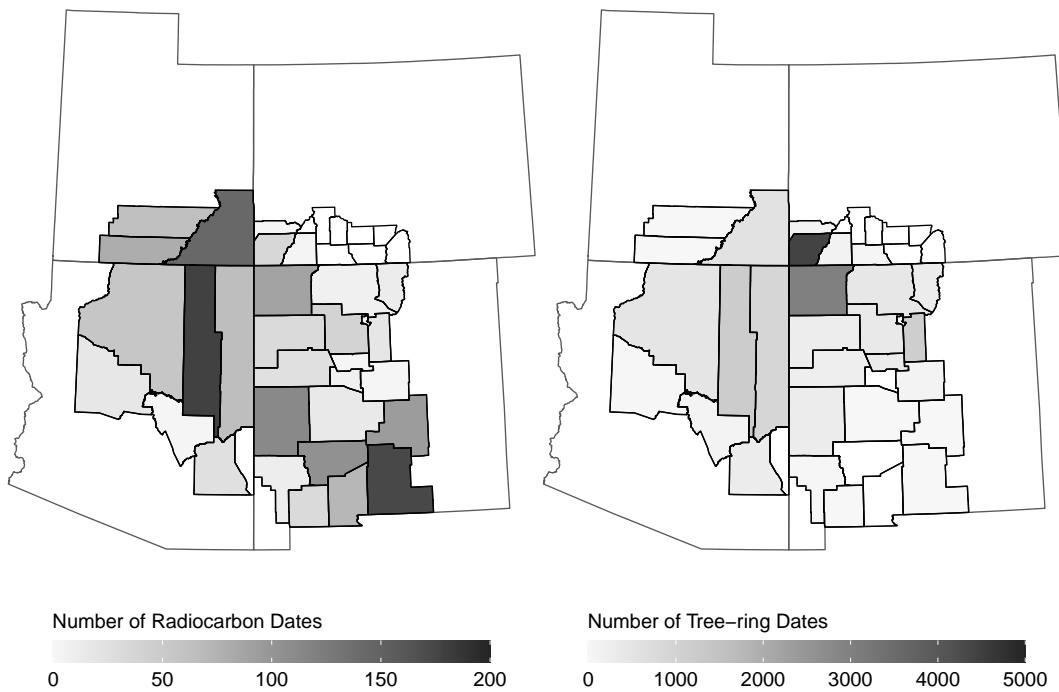
We cleaned the radiocarbon database by first removing all non-archaeological dates. To ensure our data are reasonably precise, we removed all radiocarbon dates with  $1\sigma$  errors greater than 300 years and more than 25% of the uncalibrated age. We removed dates with duplicate lab numbers whose radiocarbon ages,  $1\sigma$  errors, or county information did not match. We then trimmed these data to counties within the UUSW. We retained radiocarbon dates that likely derive from our period of interest by first calibrating all of the dates, then retaining dates whose 95% calibrated probability masses overlapped 1750–150 BP.

We draw on the tree-ring date database presented by Bocinsky et al. (2016). These data were compiled by Tim Kohler and Rebecca Higgins from many contributors, and the vast majority of dates were produced by the Laboratory of Tree Ring Research at the University of Arizona. The original dataset contained 32,863 cutting, near-cutting, and non-cutting dates from across the US Southwest. Here, we only use the dates that are most likely to be accurate — cutting dates ('B', 'G', 'L', 'c', or 'r' dates, with or without a '+') for which we know the year the tree died, and near-cutting dates ('v' or 'v+' dates) which are likely within 0–3 years of the true date of the outermost ring. The Bocinsky et al. (2016) data contain site locations; for this analysis, we first spatially cropped the database to our study area, and then translated those locations to the county level. We converted the original date determinations to BP to match calibrated radiocarbon year conventions, and retained dates within 1750–150 BP.

Finally, due to much of the radiocarbon data only being resolved to the county level, we had to approximate inclusion in the VEP, CMV and NRG study areas. The CMV study area conforms well to Montezuma county, Colorado; the NRG overlaps with Los Alamos, Rio Arriba, Sandoval, Santa Fe, and Taos counties in New Mexico. In the analyses below, dates were included as “within” each study area if they derive from those respective counties.

Our final dataset includes 1,531 radiocarbon dates from 482 unique sites, dating from 2260–90 BP (uncalibrated; 309 BC–1860 AD) and 15,077 tree-ring dates from 770 unique sites, dating from 1749–152 BP (201–1798 AD). Table 1 presents the counts of dates and sites of each date type in the CMV and NRG study areas.

Date densities across the study area were very uneven, with high counts of radiocarbon dates deriving from Navajo County, Arizona, and almost a third of the tree-ring dates in our study being from Montezuma County, Colorado. Below, we overcome some of these sampling issues by binning dates by site and occupation phase. Still, these disparities likely demonstrate more fundamental differences in research design across archaeological projects. There are very few radiocarbon dates from the two VEP study areas, indicating a clear bias towards using tree-ring dating where (and when) possible. Figure 2 presents the number of radiocarbon and tree-ring dates by county within our study area. Figure 3 presents the number of unique sites with radiocarbon and tree-ring dates by county within our study area. There are similar ratios of radiocarbon dates to sites with radiocarbon dates for the



**Figure 2:** The number of radiocarbon and tree-ring dates by county in this study.

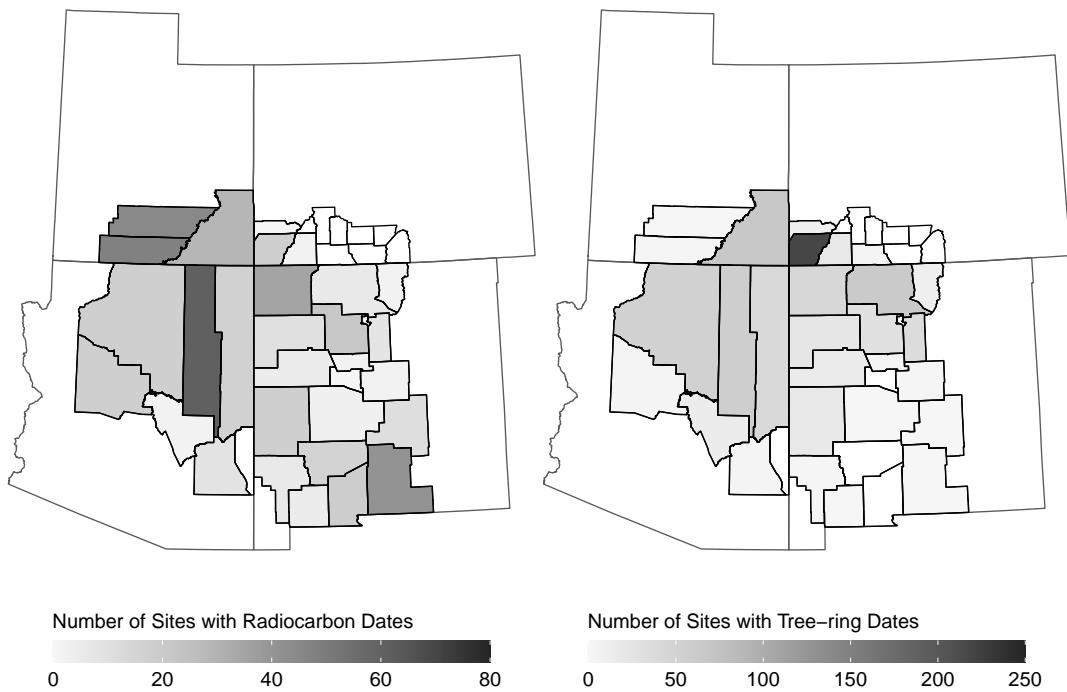
CMV and NRG (roughly 2:1), but sites in the CMV tend to have a third more tree-ring dates than those in the NRG (20:1 versus 15:1, respectively).

### Preparing the SPDs

To prepare the radiocarbon dataset, we used functions provided by the `rcarbon` package for R (Bevan et al., 2019). We calibrated the radiocarbon dates using the methods of Bronk Ramsey (2008), using the ‘intcal13’ calibration curve (Reimer et al., 2013). To control for the inconsistent number of dates among the archaeological sites represented in the database (a form of investigation bias), we aggregated radiocarbon dates known to be from the same phase of an archaeological site using the `rcarbon::binPrep` function using the single linkage agglomeration method with an *h*-value of 100 (see [Supplementary Material] for a discussion of various *h*-values). This creates clusters of dates from the same site whose mean values (uncalibrated) are within 100 years of each other, and splits site occupations into distinct phases separated by 100 years or more. We constructed summed probability distributions (SPDs) from the calibrated dates using the `rcarbon::spd` function. We averaged probability distributions of binned dates, and then constructed SPDs from the resulting site-phase probability distributions.<sup>1</sup>

We treated the tree-ring dates in as similar a way possible as the radiocarbon dates. We modeled faux-calibrated tree-ring date probabilities as discrete degenerate distributions ( $\Pr(X = k_0) = 1$ ) with  $k_0$  equal to the tree-ring date — in other words, a probability of 1 at the date value, and 0 elsewhere. This is akin to modeling the tree ring date as a normal distribution with a mean of the date

<sup>1</sup>Using this method, a site-phase with a single date and a site-phase with 10 dates would each contribute equally to the resulting SPD. This method of developing the SPD is therefore a model of the relative abundance of site occupation events through time rather than of dates through time. This method does not account for investigation biases deriving from the definition of “site” or the decision to use a particular dating technique at a site in the first place.



**Figure 3:** The number of sites with radiocarbon and tree-ring dates by county in this study.

value, and a standard deviation of zero. Thinking of tree ring dates this way allows us to use the same binning and SPD algorithms as for radiocarbon dates.<sup>2</sup> Here, we binned the tree-ring dates into site-phases using the `rcarbon::binPrep` function and constructing SPDs using `rcarbon::spd`, both with the same parameter selections as used for the radiocarbon samples. As above, we averaged probability distributions of binned dates, and then constructed SPDs from the site-phase distributions.

## Results

Figure 4 presents radiocarbon and tree-ring SPDs for the UUSW, CMV, and NRG regions, as well as VEP population reconstructions for the CMV and NRG regions. Across the UUSW, we see striking agreement in directionality and timing between the radiocarbon and tree-ring series (Figure 4A); both series trend upwards until the late thirteenth century, and then decline rapidly. The tree-ring series—the light orange line is the raw SPD, and the dark line is smoothed—shows more variation on annual to decadal timescales, and yet matches the broad trends of the radiocarbon SPD. Roughly 150-year periodic fluctuations in the tree-ring SPD, described as periods of “exploration” and “exploitation” by Bocinsky et al. (2016), are not present in the radiocarbon SPD, the result of a much smaller sample size of site-phases and lower temporal resolution of radiocarbon dating relative to tree-ring dating.

```
#>           used (Mb) gc trigger   (Mb) limit (Mb) max used   (Mb)
#> Ncells  716943  38.3  1232191   65.9        NA 1232191   65.9
#> Vcells 1639461  12.6  508769831 3881.7      716800 507690663 3873.4
```

<sup>2</sup>One might consider alternative probability distributions for non-cutting or near-cutting dates. The probability mass function in such a case would be zero for years prior to the date of the outer ring of the sample, and non-zero for subsequent years, presumably decaying quickly. For instance, we could imagine a study that derives empirical probability distributions for non-cutting dates based on cutting-dates from the same contexts.

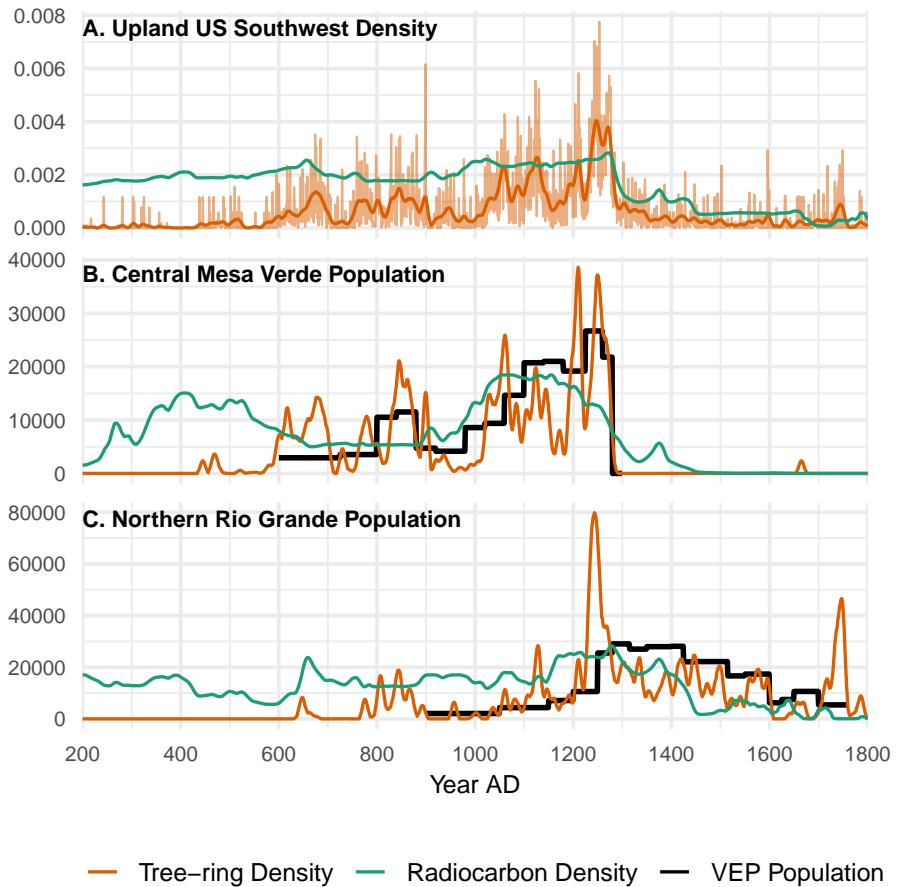
```
#>           used (Mb) gc trigger   (Mb) limit (Mb) max used   (Mb)
#> Ncells    717083 38.3     1232191   65.9        NA 1232191   65.9
#> Vcells   1639786 12.6    407015865 3105.3      716800 507690663 3873.4
```

Comparisons between the CMV and NRG regional SPDs and the VEP population estimates are also revealing. In Figures 4B and 4C, the regional SPDs have been scaled such that the areas under their densities (the orange and green lines) are equal to the area under the VEP population estimates (the black lines). For both regions, the radiocarbon and tree-ring SPDs capture the broad directionality of the VEP population estimates. In the CMV region (Figure 4B), the radiocarbon SPD has substantial probability mass prior to the VEP reconstruction, which is likely due to sampling bias in the radiocarbon record that is skewed in this region towards older (pre-Pueblo) sites. The radiocarbon SPD also does not capture the Pueblo I (AD 750–900) population peak (which may be too short-lived to be captured by such a small sample of radiocarbon dates), and seems to begin declining by the early thirteenth century AD. The tree-ring SPDs better capture the variability in population, though include a decline in the mid AD 1100s that is not in the VEP population estimates. In the NRG region (Figure 4C), both SPDs rise until about AD 1250, then decline thereafter. The high peak in the tree-ring record at AD 1250, far in excess of the VEP population estimate, is particularly interesting — it likely signals the net impact of the immigration of many thousands of people to the NRG region, and the resulting (and simultaneous) burst in construction. Such a sudden and synchronous demographic and architectural event is hard to capture in the relatively lower resolution and sparsely sampled radiocarbon record, which tends to be smeared across sub-centennial scales courtesy of the calibration process (Baillie, 1991).

These results yield key insights on the relative sensitivity of the radiocarbon record to capture long-term, coarse-grained demographic change across the Southwest. As stated in the introduction, our aim is to understand how well the radiocarbon record captures population declines across the Southwest, as we need to develop a measure of confidence in radiocarbon records from regions across and beyond the Southwest that lack other demographic proxies. Both the radiocarbon and tree-ring SPDs show declines post-AD 1300 across all regions. In both the CMV and NRG the decline in the radiocarbon SPD begins earlier than in the VEP population series. This may indicate that radiocarbon time series make the decline seem to begin earlier than it actually does, which could be caused by a ‘cliff’ in the radiocarbon calibration curve from AD 1200–1300 that is surrounded by two plateaus at AD 1100–1200 and AD 1400 (Reimer et al., 2013).

Comparisons of the radiocarbon and dendrochronological record with the architecture/ceramic record show similarities and differences that reinforce the need for multiple-proxy studies of palaeodemography when moving between broader and finer spatial scales. While the multi-proxy population reconstruction for the CMV region was similar to the radiocarbon and dendrochronological records, it showed a different trend for the NRG region (lower panel, Figure 4). There, the population reconstruction peaks from AD 1300–1400, whereas the radiocarbon record declines, at first quickly and then slowly. These different trends between the VEP population reconstructions and dendrochronological and radiocarbon proxies reflect some investigation bias in both the radiocarbon and dendrochronological records at smaller, intra-regional scales, which calls for caution in the use of radiocarbon records to assess possible paleo-migration scenarios at finer spatial scales. This result reinforces other work that has called for the use of multiple proxies as cross-checks on radiocarbon time series (Williams, 2012, p. @Crombe2014).

The central aim in comparing radiocarbon to dendrochronological records here was to assess the potential impacts of investigation bias in radiocarbon records, caused by researchers not taking radio-



**Figure 4:** Radiocarbon and tree-ring density distributions in the UUSW (A), CMV (B), and NRG (C) regions, and VEP population estimates for the CMV and NRG. The dark orange tree-ring density curve smooths the raw tree-ring density (light orange) using a 21-year center-aligned Gaussian kernel with a 5-year standard deviation, following Bocinsky et al. (2016). In Panel A, the radiocarbon density (y-axis) is scaled by a factor of 3 to facilitate comparison with the tree-ring density over the target time period. The radiocarbon and tree-ring densities (y-axis) in panels B and C are scaled such that the area under their densities is equal to the area under the VEP population estimates (i.e., the cumulative population).

carbon samples in favour of other types of chronometric techniques (e.g., diagnostic artefact counts, tree-ring dates). Specifically, we are interested in the potential impact of this kind of bias in the post AD 1300 decline of radiocarbon SPDs across and beyond the Southwest. Similarities in the decline of both radiocarbon and dendrochronological records across the UUSW suggest that, at this particular regional scale, radiocarbon records reflect long-term, coarse-grained trends in paleodemography. Indeed, not only do we see corresponding declines in the radiocarbon and dendrochronological records at this large scale, but these patterns are also consistent with previous work on demographic and social change in the American Southwest: the Coalescent Communities Database, a Southwest-wide database of room counts and occupation spans of habitation sites with 12 rooms or more, supports a “population decline that began between AD 1300 and 1350 and continued to European contact” (Phillips et al., 2018; following Hill et al., 2004: Figures 2,4).

## Conclusion

Understanding the geographic scale of post-AD 1300 human population declines in the American Southwest requires a reliance on radiocarbon-based reconstructions of palaeodemography, as radiocarbon is often the only proxy record available in many regions for developing continuous time series. However, radiocarbon records may include investigation bias caused by researchers not selecting radiocarbon samples in favour of Late Prehistoric diagnostic artefacts or other dating methods. In this paper we develop a new method for systematically comparing radiocarbon time series against multiple palaeodemographic proxies. We selected three regions—the Upland US Southwest, Central Mesa Verde, and the Northern Rio Grande—where we expected the most investigation bias to occur, because in these regions multiple higher resolution dating techniques are available. Results indicated that all three palaeodemographic proxies reveal post-AD 1300 population declines. While there is some variability at finer spatial and temporal scales that warrant the collection of other data that can be used to cross-check and validate the resolution of radiocarbon records, at coarse-grained, large spatial and temporal scales radiocarbon records provide a robust relative measure of long-term palaeodemographic trends. These results lend confidence to single proxy radiocarbon-based reconstructions of palaeodemography outside the Southwest that suggest post-AD 1300 population declines in many parts of North America.

Investigation bias is an inevitable issue that all researchers using radiocarbon records have to address. The results of this paper advance our confidence in radiocarbon records as a coarse-grained measure of the directionality of human paleodemographic trends. The methods developed can be used for other regions of the world that have dendrochronological records, such as Europe (Haneca et al., 2009; Hillam et al., 1990; Kuniholm, 2001), where comparisons with radiocarbon records could help to advance knowledge of Neolithic ‘booms and busts’ of populations (Shennan et al., 2013; Whitehouse et al., 2014) and to overcome limitations of the Bronze Age radiocarbon calibration plateau (Armit et al., 2013).

**Acknowledgements** We thank the guest editorial board for inviting us to contribute to this volume. We thank two anonymous reviewers for their helpful suggestions for improvement. We thank our colleagues in the PAGES People3000 working group, and PAGES for funding this working group. We thank the National Foundation for supporting this research. Lastly, we thank the generations of different stakeholders working across the Southwest for their collection, storage and publication of the data presented here.

**Data Accessibility** A compendium of data and code used in this analysis is available on Github at <https://github.com/people3k/robinson2020> and is archived with Zenodo at <https://doi.org/10.5281/zenodo.3647599>. Radiocarbon data used in this analysis are archived with the Digital Archaeological Record (tDAR) and are available at <https://doi.org/10.6067/XCV8455460>. Tree-ring data used here are archived with tDAR at <https://doi.org/10.6067/XCV82J6D7B>. Archaeological site locations are protected by US federal and state statutes. In order to protect these sensitive data, we pre-processed both the radiocarbon and tree-ring data; data included in the Github repository do not include site locations, and researchers will need to gain permission from the State Historic Preservation Office (SHPO) of each state in order to access the site location data.

**Ethics** No ethical issues are associated with this paper.

**Authors' contributions** E.R. and R.K.B designed the study, collected and analysed the data, and wrote the first draft. J.F. and D.B collected and organized data. R.L.K collected data. All authors revised/improved the manuscript.

**Competing interests** We have no competing interests.

**Funding** Funding for this research was provided by NSF grants BCS-1418858, BCS-1624061, BCS-1822033 and PAGES (Past Global Changes) funding to the Working Group [Paleoclimate and the Peopling of the Earth \(PEOPLE 3000\)](#).

## References

- Armit, I., Swindles, G.T., Becker, K., 2013. From dates to demography in later prehistoric Ireland? Experimental approaches to the meta-analysis of large  $^{14}\text{C}$  data-sets. *Journal of archaeological science* 40, 433–438. <https://doi.org/10.1016/j.jas.2012.08.039>
- Baillie, M.G., 1991. Suck-in and smear: Two related chronological problems for the 90s. *Journal of theoretical archaeology* 2, 12–16.
- Bamforth, D.B., Grund, B., 2012. Radiocarbon calibration curves, summed probability distributions, and early Paleoindian population trends in North America. *Journal of Archaeological Science* 39, 1768–1774. <https://doi.org/10.1016/j.jas.2012.01.017>
- Berry, M.S., 1982. Time, space and transition in anasazi prehistory. University of Utah Press, Salt Lake City, Utah.
- Bevan, A., Crema, E.R., 2020. rcarbon: Methods for calibrating and analysing radiocarbon dates.
- Bevan, A., Palmisano, A., Woodbridge, J., Fyfe, R., Roberts, C.N., Shennan, S., 2019. The changing face of the Mediterranean—Land cover, demography and environmental change: Introduction and overview. *The Holocene* 29, 703–707. <https://doi.org/10.1177/0959683619826688>
- Bluhm, L.E., Surovell, T.A., 2019. Validation of a global model of taphonomic bias using geologic radiocarbon ages. *Quaternary Research* 91, 325–328. <https://doi.org/10.1017/qua.2018.78>
- Bocinsky, R.K., Rush, J., Kintigh, K.W., Kohler, T.A., 2016. Exploration and exploitation in the macro-history of the pre-Hispanic Pueblo Southwest. *Science Advances* 2, e1501532. <https://doi.org/10.1126/sciadv.1501532>
- Boyer, J.L., Moore, J.L., Lakatos, S.A., Akins, N.J., Wilson, C.D., Blinman, E., 2010. Remodeling immigration: A Northern Rio Grande perspective on depopulation, migration, and donation-side models, in: Kohler, T., Varien, M., Wright, A. (Eds.), *Leaving Mesa Verde: Peril and Change in the Thirteenth Century Southwest*. The Amerind Foundation; University of Arizona Press, Tucson, Arizona, pp. 285–323.
- Bronk Ramsey, C., 2008. Radiocarbon dating: Revolutions in understanding. *Archaeometry* 50, 249–275. <https://doi.org/10.1111/j.1475-4754.2008.00394.x>
- Cameron, C.M., 2010. Advances in understanding the thirteenth-century depopulation of the northern Southwest, in: Kohler, T., Varien, M., Wright, A. (Eds.), *Leaving Mesa Verde: Peril and Change in the Thirteenth Century Southwest*. The Amerind Foundation; University of Arizona Press, Tucson, Arizona, pp. 346–363.
- Contreras, D.A., Meadows, J., 2014. Summed radiocarbon calibrations as a population proxy: A critical evaluation using a realistic simulation approach. *Journal of Archaeological Science* 52, 591–608. <https://doi.org/10.1016/j.jas.2014.05.030>
- Crema, E.R., Habu, J., Kobayashi, K., Madella, M., 2016. Summed probability distribution of  $^{14}\text{C}$  dates suggests regional divergences in the population dynamics of the Jomon period in eastern Japan. *PLoS One* 11, e0154809. <https://doi.org/10.1371/journal.pone.0154809>
- Crombé, P., Robinson, E., 2014.  $^{14}\text{C}$  dates as demographic proxies in Neolithisation models of north-western Europe: A critical assessment using Belgium and northeast France as a case-study. *Journal of Archaeological Science* 52, 558–566. <https://doi.org/10.1016/j.jas.2014.02.001>
- Freeman, J., Baggio, J.A., Robinson, E., Byers, D.A., Gayo, E., Finley, J.B., Meyer, J.A., Kelly, R.L., Anderies, J.M., 2018. Synchronization of energy consumption by human societies throughout

- the Holocene. *Proceedings of the National Academy of Sciences* 115, 9962–9967. <https://doi.org/10.1073/pnas.1802859115>
- Freeman, J., Byers, D.A., Robinson, E., Kelly, R.L., 2018. Culture process and the interpretation of radiocarbon data. *Radiocarbon* 60, 453–467. <https://doi.org/10.1017/RDC.2017.124>
- Goldberg, A., Mychajliw, A.M., Hadly, E.A., 2016. Post-invasion demography of prehistoric humans in South America. *Nature* 532, 232–235. <https://doi.org/10.1038/nature17176>
- Haneca, K., Čufar, K., Beeckman, H., 2009. Oaks, tree-rings and wooden cultural heritage: A review of the main characteristics and applications of oak dendrochronology in Europe. *Journal of Archaeological Science* 36, 1–11. <https://doi.org/10.1016/j.jas.2008.07.005>
- Hill, J.B., Clark, J.J., Doelle, W.H., Lyons, P.D., 2004. Prehistoric demography in the Southwest: Migration, coalescence, and Hohokam population decline. *American Antiquity* 69, 689–716. <https://doi.org/10.2307/4128444>
- Hillam, J., Groves, C.M., Brown, D.M., Baillie, M.G., Coles, J.M., Coles, B.J., 1990. Dendrochronology of the English Neolithic. *Antiquity* 64, 210–220. <https://doi.org/10.1017/S0003598X00077826>
- Kohler, T.A., Varien, M.D., Wright, A.M. (Eds.), 2010. *Leaving Mesa Verde: Peril and Change in the Thirteenth-Century Southwest*. The Amerind Foundation; University of Arizona Press, Tucson, Arizona.
- Kuniholm, P.I., 2001. Dendrochronology and other applications of tree-ring studies in archaeology, in: Brothwell, D.R., Pollard, A.M. (Eds.), *The Handbook of Archaeological Sciences*. John Wiley & Sons, Ltd., London.
- Marwick, B., 2019. rrtools: Creates a Reproducible Research Compendium.
- Marwick, B., 2017. Computational reproducibility in archaeological research: Basic principles and a case study of their implementation. *Journal of Archaeological Method and Theory* 24, 424–450. <https://doi.org/10.1007/s10816-015-9272-9>
- Michczynski, A., Michczynska, D.J., 2006. The effect of PDF peaks' height increase during calibration of radiocarbon date sets. *Geochronometria* 25, 1–4.
- Naudinot, N., Tomasso, A., Tozzi, C., Peresani, M., 2014. Changes in mobility patterns as a factor of  $^{14}\text{C}$  date density variation in the Late Epigravettian of Northern Italy and Southeastern France. *Journal of Archaeological Science* 52, 578–590. <https://doi.org/10.1016/j.jas.2014.05.021>
- Ortman, S.G., 2016. Uniform probability density analysis and population history in the northern Rio Grande. *Journal of Archaeological Method and Theory* 23, 95–126. <https://doi.org/10.1007/s10816-014-9227-6>
- Palmisano, A., Woodbridge, J., Roberts, C.N., Bevan, A., Fyfe, R., Shennan, S., Cheddadi, R., Greenberg, R., Kaniewski, D., Langgut, D., others, 2019. Holocene landscape dynamics and long-term population trends in the Levant. *The Holocene* 29, 708–727. <https://doi.org/10.1177/0959683619826642>
- Peros, M.C., Munoz, S.E., Gajewski, K., Viau, A.E., 2010. Prehistoric demography of North America inferred from radiocarbon data. *Journal of Archaeological Science* 37, 656–664. <https://doi.org/10.1016/j.jas.2009.10.029>
- Phillips, D.A., Wearing, H.J., Clark, J.J., 2018. Village growth, emerging infectious disease, and the end of the Neolithic Demographic Transition in the Southwest United States and Northwest Mexico. *American Antiquity* 83, 263–280. <https://doi.org/10.1017/aaq.2018.3>

- Reimer, P.J., Bard, E., Bayliss, A., Beck, J.W., Blackwell, P.G., Ramsey, C.B., Buck, C.E., Cheng, H., Edwards, R.L., Friedrich, M., others, 2013. IntCal13 and Marine13 radiocarbon age calibration curves 0–50,000 years cal BP. Radiocarbon 55, 1869–1887. [https://doi.org/10.2458/azu\\_js\\_rc.55.16947](https://doi.org/10.2458/azu_js_rc.55.16947)
- Rick, J.W., 1987. Dates as data: An examination of the Peruvian preceramic radiocarbon record. American Antiquity 52, 55–73. <https://doi.org/10.2307/281060>
- Robinson, E., Nicholson, C., Kelly, R.L., 2019. The importance of spatial data to open-access national archaeological databases and the development of paleodemography research. Advances in Archaeological Practice 7, 395–408. <https://doi.org/10.1017/aap.2019.29>
- Russell, T., Steele, J., 2009. A geo-referenced radiocarbon database for Early Iron Age sites in sub-Saharan Africa: Initial analysis. Southern African Humanities 21, 327–344.
- Schwindt, D.M., Bocinsky, R.K., Ortman, S.G., Glowacki, D.M., Varien, M.D., Kohler, T.A., 2016. The social consequences of climate change in the central Mesa Verde region. American Antiquity 81, 74–96.
- Shennan, S., Downey, S.S., Timpson, A., Edinborough, K., Colledge, S., Kerig, T., Manning, K., Thomas, M.G., 2013. Regional population collapse followed initial agriculture booms in mid-holocene europe. Nature communications 4, 1–8.
- Surovell, T.A., Brantingham, P.J., 2007. A note on the use of temporal frequency distributions in studies of prehistoric demography. Journal of Archaeological Science 34, 1868–1877.
- Surovell, T.A., Finley, J.B., Smith, G.M., Brantingham, P.J., Kelly, R., 2009. Correcting temporal frequency distributions for taphonomic bias. Journal of archaeological Science 36, 1715–1724.
- Timpson, A., Colledge, S., Crema, E., Edinborough, K., Kerig, T., Manning, K., Thomas, M.G., Shennan, S., 2014. Reconstructing regional population fluctuations in the european neolithic using radiocarbon dates: A new case-study using an improved method. Journal of Archaeological Science 52, 549–557.
- Wang, C., Lu, H., Zhang, J., Gu, Z., He, K., 2014. Prehistoric demographic fluctuations in china inferred from radiocarbon data and their linkage with climate change over the past 50,000 years. Quaternary Science Reviews 98, 45–59.
- Whitehouse, N.J., Schulting, R.J., McClatchie, M., Barratt, P., McLaughlin, T.R., Bogaard, A., Colledge, S., Marchant, R., Gaffrey, J., Bunting, M.J., 2014. Neolithic agriculture on the European western frontier: The boom and bust of early farming in Ireland. Journal of Archaeological Science 51, 181–205.
- Williams, A.N., 2012. The use of summed radiocarbon probability distributions in archaeology: A review of methods. Journal of Archaeological Science 39, 578–589.

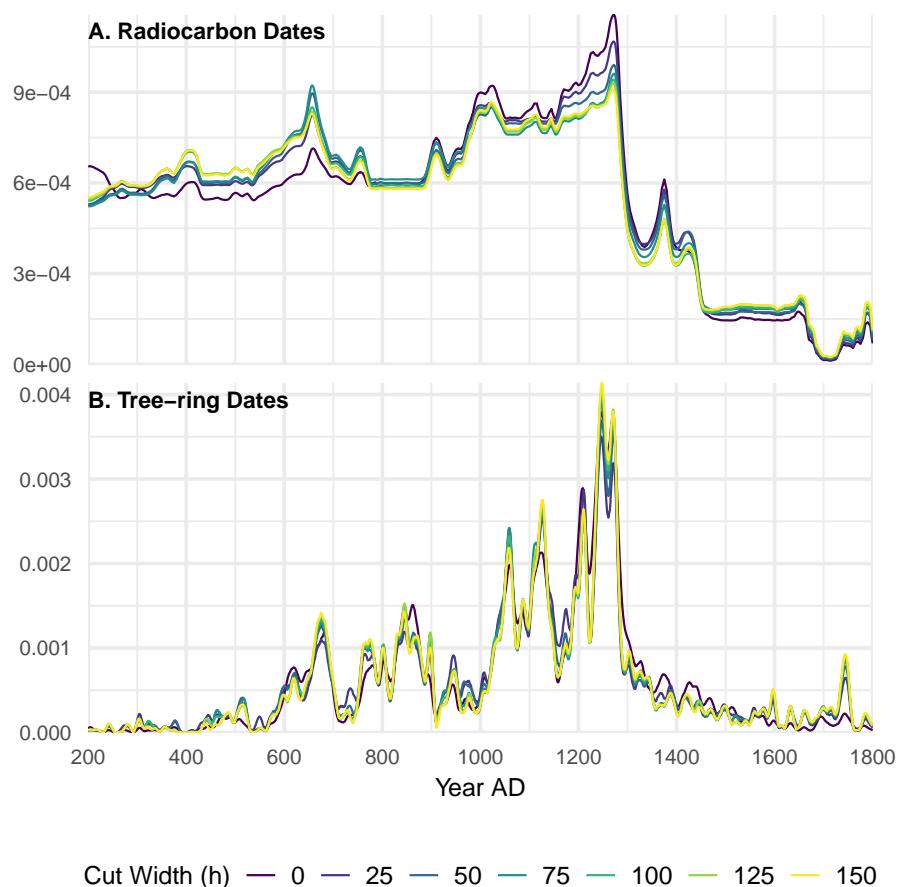
## Supplementary Material: Cut-width sensitivity for temporal binning

Uneven sampling across archaeological contexts is a persistent problem when constructing summed probability distributions from radiocarbon dates, and perhaps even a larger issue for the more ubiquitous tree-ring dates in regions like the Upland US Southwest. In SPDs of radiocarbon dates, it is conventional to combine dates from single sites prior to creating the SPDs. Dates from separate components or phases of multi-component sites should be counted separately, however.

To control for inconsistent sample size across the archaeological sites represented in the database, we aggregated dates known to be from the same phase of same archaeological site using the `rcarbon::binPrep` function using the “single linkage” agglomeration method with a cut width (or  $h$ -value) of 100 (Bevan and Crema, 2020). This clusters dates from the same site that are within 100 years of each other, and splits site occupations into distinct phases separated by 100 years or more. Binned calibrated dates are then summed, and the resulting probability mass function is divided by the number of dates in the site-phase.

Because the wider archaeological SPD research community still lacks a standard protocol for determining  $h$ , we perform a sensitivity analysis to demonstrate that the resulting SPDs for radiocarbon and tree-ring dates stabilize at a cut width of 100 or so. We calculated SPDs at 25-year intervals from 0 (only duplicate dates are binned) to 150 years (effectively binning by site).

As can be seen in Figure 5, binning in general has little impact on the resulting SPDs, but cut widths above  $h = 100$  years are virtually indistinguishable. The shape of the radiocarbon SPD (panel A) stabilizes above  $h = 0$  and its values stabilize by  $h = 100$ . The tree-ring SPD arguably stabilizes in both shape and value above  $h = 75$  or so, but in the analysis we use  $h = 100$  to keep our methods consistent between the radiocarbon and tree-ring SPDs.



**Figure 5:** Radiocarbon (A) and tree-ring (B) density distributions in the UUSW, given varying cut widths for binning dates within sites. The density distributions for the UUSW are normalized. Years are given in AD.

## Colophon

This report is part of a research compendium developed using the `rrtools` package for reproducible research (Marwick, 2019, 2017). This report was generated on 2020-04-17 08:31:33 using the following computational environment and dependencies:

**R version 3.6.3 (2020-02-29)**

**Platform:** x86\_64-apple-darwin15.6.0 (64-bit)

**locale:** en\_US.UTF-8||en\_US.UTF-8||en\_US.UTF-8||C||en\_US.UTF-8||en\_US.UTF-8

**attached base packages:** *stats, graphics, grDevices, utils, datasets, methods* and *base*

**other attached packages:** *ggplot2(v.3.3.0), robinson2020(v.1.0.0.9000)* and *magrittr(v.1.5)*

**loaded via a namespace (and not attached):** *Rcpp(v.1.0.4), compiler(v.3.6.3), pillar(v.1.4.3), tools(v.3.6.3), digest(v.0.6.25), evaluate(v.0.14), lifecycle(v.0.2.0), tibble(v.3.0.0), gtable(v.0.3.0), pkgconfig(v.2.0.3), rlang(v.0.4.5), cli(v.2.0.2), yaml(v.2.2.1), xfun(v.0.13), withr(v.2.1.2), stringr(v.1.4.0), dplyr(v.0.8.5), knitr(v.1.28), vctrs(v.0.2.4), rprojroot(v.1.3-2), grid(v.3.6.3), tidyselect(v.1.0.0), glue(v.1.4.0), here(v.0.1), R6(v.2.4.1), fansi(v.0.4.1), rmarkdown(v.2.1), bookdown(v.0.18), pander(v.0.6.3), purrr(v.0.3.3), backports(v.1.1.6), scales(v.1.1.0), ellipsis(v.0.3.0), htmltools(v.0.4.0), assertthat(v.0.2.1), colorspace(v.1.4-1), stringi(v.1.4.6), munsell(v.0.5.0) and crayon(v.1.3.4)*

### Affiliation: