Correcting temporal frequency distributions for taphonomic bias

Todd A. Surovell a, *, Judson Byrd Finley b, Geoffrey M. Smith a, P. Jeffrey Brantingham c, Robert Kelly a

a Department of Anthropology, University of Wyoming, Laramie, 1000 East University Avenue, WY 82071, USA
b Department of Earth Sciences, University of Memphis, 001 Johnson Hall, Memphis, TN 38152-3530, USA
c Department of Anthropology, University of California, Los Angeles, Los Angeles, CA 90095, USA

1. Introduction

In a previous paper, two of the present authors argued that the primary signal observable in most temporal frequency distributions of archaeological sites and dates is not human demography but instead taphonomic bias (Surovell and Brantingham, 2007). The crux of the argument is that the longer something is in existence, instead taphonomic bias (Surovell and Brantingham, 2007). The problem of taphonomic bias refers to the tendency for younger things to be over-represented relative to older things in the archaeological record due to the operation of destructive processes like erosion and weathering. Using a database of radiocarbon dated volcanic deposits from Bryson, R.U., Bryson, R.A., Ruter, A. [2006. A calibrated radiocarbon database of late Quaternary volcanic eruptions. Earth Discussions 1, 123–124], we develop an empirical model of taphonomic bias. In contrast to our prior study in which we modeled taphonomic bias as an exponential function wherein the likelihood of site loss remains constant through time, we argue that the probability of site destruction actually decreases with site age. We further demonstrate how this model can be used to correct temporal frequency distributions and extract demographic histories. We illustrate this approach using databases of radiocarbon dates from rockshelter and open-air sites in the Bighorn Basin, Wyoming, USA, and mammoths and humans in Siberia, Russia.

In this paper, we expand upon a prior study [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] that explored the problem of taphonomic bias. Taphonomic bias refers to the tendency for younger things to be over-represented relative to older things in the archaeological record due to the operation of destructive processes like erosion and weathering. Using a database of radiocarbon dated volcanic eruptions using a dataset compiled by Anderson et al., 1997; Hiscock, 2008: pp. 228–239; Johnson, 2006: pp. 136–137; Johnstone et al., 2006; Kirch, 1998: p. 288; Thorndycraft and Benito, 2006; Ugan and Byers, 2007, 2008). Furthermore, through the use of simple mathematical models, we showed how a constant rate of site loss (e.g., 0.01% of sites lost per year) can obscure any demographic signal that may have once been present (see also Johnson, 2006: pp. 136–137; Hiscock, 2008: pp. 228–239).

In this paper, we develop methods for correcting taphonomic bias, thereby allowing demographic signals to be extracted from archaeological temporal frequency distributions. We build on the idea proposed in our prior paper that demographic signals can be discerned from frequency distributions by examining the ratio of archaeological to geologic contexts (Surovell and Brantingham, 2007). If archaeology is common relative to geologic opportunities for sites to exist, we can infer that a strong demographic signal is apparent in the record. Conversely, if it is rare relative to geology, we can infer that population densities were low. We also demonstrate that taphonomic bias likely does not operate as a simple constant rate of site loss.

We begin by examining a temporal frequency distribution of radiocarbon dated volcanic eruptions using a dataset compiled by Bryson et al. (2006). We use this database as a proxy for the global temporal frequency distribution of sediments (i.e., geologic contexts) and to examine the true time-dependent nature of the taphonomic function. We then turn to two archaeological case studies from the Bighorn Basin, Wyoming, USA and Siberia, Russia to demonstrate how the volcanic dataset can be used to correct archaeological data.

* Corresponding author. Tel.: +1 307 766 3239; fax: +1 307 766 2473.
E-mail addresses: surovell@uwyo.edu (T.A. Surovell), judson.finley@western-geoarch.com (J. Byrd Finley), geoffrey_smith@hotmail.com (G.M. Smith), pjb@anthro.ucla.edu (P.J. Brantingham), RLKelly@uwyo.edu (R. Kelly).
2. A realistic model of taphonomic bias

Previously, we modeled taphonomic bias as an exponential function, making a simplifying assumption that a constant proportion of archaeological sites (and geologic contexts) are lost per year to destructive processes (Surovell and Brantingham, 2007). While this model is qualitatively similar to observed temporal frequency distributions and demonstrates how simple taphonomic processes can drive the emergence of distributions which increase curvilinearly toward the present, it has since become apparent that the model is not a quantitatively accurate description of how the world actually works. Although exponential functions can be fit to many temporal frequency distributions, we demonstrate that other models, such as power and inverse functions, often fit better. We extracted data from a handful of temporal frequency distributions involving non-archaeological data (Mandel, 1995; Bryson et al., 2006; Guthrie, 2006; Miller et al., 1999; Pierce et al., 2004).¹ We regressed three functions to each (Table 1), an exponential function and inverse and power functions, often fit better. We extracted data from a handful of temporal frequency distributions involving non-archaeological data (Mandel, 1995; Bryson et al., 2006; Guthrie, 2006; Miller et al., 1999; Pierce et al., 2004).¹ We regressed three functions to each (Table 1), suggesting that a model that assumes a constant rate of site loss through time is almost certainly incorrect.

To perform these analyses, radiocarbon dates were calibrated using OxCal v. 3.10 (Ramsey, 2005) using the IntCal04 calibration curve (Reimer et al., 2004). The calibrated age of each sample was estimated as the median of the 2σ continuous calibrated age range, and frequency distributions were created. The Genyornis data were digitized from Fig. 3 of Miller et al. (1999). Because each dataset exhibits a drop-off in frequency associated with the most recent time intervals sampled (attributed to extinction in the faunal data and research bias in the geologic cases), we excluded all data points younger than the peak in frequency, including the peak itself. We also trimmed all zero frequency data points, as inverse and power regressions cannot be performed if these are present. Non-linear regressions were performed in SPSS v. 15.0. For all datasets, inverse and power functions exhibited better fits to temporal frequency distributions than exponential functions (Table 1), suggesting that a model that assumes a constant rate of site loss through time is almost certainly incorrect.

To construct a more realistic empirical model of taphonomic bias, ideally one would seek a record of some phenomenon that is geographically widespread and characterized by a relatively constant rate of occurrence through time, such that long-term trends in frequency seen in the geologic record are entirely governed by the operation of taphonomic processes through time. We suggest that the terrestrial record of volcanism nicely fits these requirements, and we now turn to a database compiled by Bryson et al. (2006) of 2021 radiocarbon dates associated with volcanic deposits.² The database spans the past 40,000 years, and is global and terrestrial in nature (Fig. 1a). We assume that: (1) rates of volcanism were relatively stochastic over the late Quaternary; (2) the primary upward curvilinear pattern in this dataset from ca. 40,000 to 1000 BP is entirely attributable to taphonomic bias, an assumption also made by the original authors (Bryson et al., 2006), and therefore; (3) that this frequency distribution can be used as a proxy for the temporal frequency distribution of terrestrial sediments globally. The frequency of dates in the last millennium of this database drops dramatically, a phenomenon that we have noted is present in virtually every frequency distribution we have examined (Surovell and Brantingham, 2007). This drop-off cannot be attributed to taphonomic bias but instead is probably explained by research bias against recent deposits, or a persistent prejudice against the present in the paleo professions, if you will.

To examine the first and second assumptions, we turned to an independent record of Quaternary volcanism from the GISP2 ice core in Greenland (Zielinski et al., 1994, 1996), a record which, unlike terrestrial sedimentary contexts, is not subject to taphonomic bias due to continuous accumulation. In two related studies based on concentrations of SO₄²⁻ in glacial ice, Zielinski et al. have identified and dated major eruption events during the Holocene (8950 BP) and Late Pleistocene (100,000 BP). It should be noted that these datasets include only the most explosive volcanic eruptions of the late Quaternary, which include less than 2% of total explosive volcanism. For both datasets, we performed linear, exponential, inverse, and power regressions comparing age and frequency and found no significant relationship thereby supporting the hypothesis that volcanism has remained stochastic over the last 40,000 years (Fig. 1a).

In contrast, at face value, the Bryson et al. (2006) terrestrial record of volcanism would suggest a continual increase in volcanism over this time period (Fig. 1a), a difference we believe can be attributed only to taphonomic bias. In fact, taphonomic bias has so severely affected the terrestrial record of volcanism that there is no correlation between frequency data points in the terrestrial and GISP2 time series ($r = -0.206, n = 40, p = 0.202$). While this is not a good thing for those studying the frequency of volcanism from

1. Because archaeological frequency distributions should not only be affected by taphonomic bias but also human demography, to explore the true time-dependent nature of taphonomic bias, we limited this analysis to geologic and paleontological datasets, which should be less subject to demographic effects. While paleontological datasets are affected by demographic processes, animal population densities in the absence of human predation should be governed largely by changes in carrying capacity, and we assume that carrying capacity did not vary substantially for these populations.

2. For all functions, $n_t$ represents the frequency of some phenomenon at time $t$. The exponential function takes the form $n_t = n_0 e^{at}$, where $n_0$ is the frequency of dates at time $t$, and $a$ and $b$ (analogous to $k$ from Surovell and Brantingham, 2007) are the regression coefficients. A traditional inverse function takes the form $n_t = a + b/t$, while the function we used is $n_t = a + b/t + c$, where $a$, $b$, and $c$ are the regression coefficients. A traditional inverse function takes the form $n_t = a + b/t$. Instead we used $n_t = a/(t + b)$, where $a$ and $b$ are the regression coefficients. The addition of the extra coefficient in both functions causes them to intersect the y-axis at a finite value as would be expected in actual temporal frequency distributions.

3. While volcanoes do occur on every continent of the world, their spatial distribution is not random. We do not know to what extent this introduces a spatial bias into the Bryson et al. (2006) dataset.

---

**Table 1**

Non-linear regression of geologic and paleontological temporal frequency distributions.

<table>
<thead>
<tr>
<th>Volcanic eruption ¹⁴C dates, global (G)</th>
<th>1616</th>
<th>0.25</th>
<th>40–0.75</th>
<th>0.712</th>
<th>0.928</th>
<th>0.920</th>
<th>Bryson et al. (2006)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alluvium ¹⁴C dates, central plains, USA (G)</td>
<td>288</td>
<td>0.25</td>
<td>13–15</td>
<td>0.603</td>
<td>0.696</td>
<td>0.696</td>
<td>Mandel (1995)</td>
</tr>
<tr>
<td>Alluvial fan ¹³C dates, Wyoming, Idaho, USA (G)</td>
<td>50</td>
<td>1.0</td>
<td>13–2</td>
<td>0.819</td>
<td>0.963</td>
<td>0.964</td>
<td>Pierce et al. (2004)</td>
</tr>
<tr>
<td>Pleistocene fauna ¹³C dates, Alaska, USA (P)</td>
<td>296</td>
<td>1.0</td>
<td>49–15</td>
<td>0.458</td>
<td>0.368</td>
<td>0.570</td>
<td>Guthrie (2006)</td>
</tr>
<tr>
<td>Genyornis AAR dates, Australia (P)</td>
<td>191a</td>
<td>5.0</td>
<td>130–65</td>
<td>0.410</td>
<td>0.747</td>
<td>0.752</td>
<td>Miller et al. (1999)</td>
</tr>
</tbody>
</table>

Notes: typewriter abbreviations – G: geologic; P: paleontological.

* Sample size estimated because data were digitized from Miller et al. (1999: Fig. 3).

**b** Boldface indicates the function exhibiting the greatest degree of correlation.
terrestrial records, it is an ideal situation for examining the operation of taphonomic bias over the late Quaternary because pyroclastic sediments were strewn over the Earth at a relatively constant rate. Thus, their occurrence in the geologic record may serve as a global model of taphonomic bias.

While short-term fluctuations in the observed frequency distribution are likely a product of both changes in volcanic activity and chance in sampling, the long-term general trend should reflect taphonomic bias (Bryson et al., 2006; Surovell and Brantingham, 2007). Therefore, to construct a global model of taphonomic bias, we use a simple non-linear regression technique. We compiled counts of calibrated radiocarbon dates in each 250-year interval between 0 and 40,000 BP. As above, we discarded all data points \( n = 3 \) from the peak in dates (500–750 BP) to the present to eliminate time intervals we believe to be most seriously affected by research bias; we also trimmed all zero frequency data points \( n = 6 \). Using the median age and observed frequency of dates for each time interval, we regressed a power function to the frequency distribution. The regression model is highly significant \( r^2 = 0.928, p = 0.001 \) and explains more than 90% of the variance in date frequency (Fig. 1b). The best-fit model is:

\[
n_t = 5.726442 \times 10^6 (t + 2176.4)^{-1.3925309}
\]

where \( n_t \) is the number of radiocarbon dates surviving from time \( t \).

The annual likelihood of survival \( (1 - \lambda) \) can be calculated as:

\[
\lambda = \frac{1.3925309}{2176.4 + t}
\]

To illustrate how this model works, we consider a hypothetical example. If buried pyroclastic sediment can be used as an analogy for a buried archaeological site, this model can be used to examine the average global effects of taphonomic bias on the archaeological record. First, we examine the destruction of 10,000 archaeological sites produced in single year (Fig. 2a) by multiplying the regression model by a constant \( (k = 77.682) \). According to this model after approximately one year, six sites would be lost to destructive processes. After 1400 years half of these sites would be lost. After 10,000 years, only 900 or 9% of the original sites would remain. After 40,000 years, the limit of the regression model, approximately 160 or 1.6% of the original sites would still be available for archaeological research. It is not clear how far this curve can be extrapolated beyond 40,000 BP, the limit of the volcanic dataset, but if we do, this model would estimate that after approximately 1.6 million years, only one of the original 10,000 sites would remain.

Although this curve bears qualitative similarities to an exponential function, it differs in one important way, that the rate of site loss does not remain constant through time. From this curve, it is possible to estimate how the rate of taphonomic loss, or the probability of site destruction, changes as a function of site age. For any point on this curve, the taphonomic rate \( (\lambda) \) can be calculated as:

\[
\lambda = \frac{1.3925309}{2176.4 + t}
\]
3. Case 1: open-air and rockshelter sites in Wyoming

Although the empirical model of taphonomic bias derived from terrestrial records of volcanism seems intuitively reasonable with respect to actual rates of site loss and declining taphonomic rate with time, to show that it can be used to correct archaeological data requires demonstrating that a true demographic signal can be extracted from an actual archaeological case. The ideal test case would involve comparing a known human population history to a corrected archaeological frequency distribution, but we do not have this luxury. Instead we compare frequencies of radiocarbon dated archaeological components from open-air sites and rockshelters, assuming that because caves and rockshelters act as sediment traps (Colcutt, 1979; Farrand, 2001; Johnson, 2006: pp. 136–137; Waters, 1992: p. 242), they are largely insulated from taphonomic bias (but see Butzer, 2008).

Rockshelters in the Bighorn Basin, Wyoming are a good test case because local geomorphic conditions isolate most sites from erosion, making unconformities rare (Finley, 2008). Rockshelters commonly form in the Madison paleokarst (Sando, 1988) that is exposed in cliff faces well above drainage bottoms. Dominant sediment sources are roof fall, colluvium, and loess. Rockshelters that formed in valley bottoms with permanent streams are prone to massive erosion but still preserve highly stratified geologic sequences (Finley, 2007, 2008). Although not all rockshelters contain complete Late Quaternary depositional sequences the fact that erosional unconformities are rare justifies using rockshelter radiocarbon frequency distributions as a proxy for past demographic conditions. Therefore, we feel the rockshelter frequency distribution is a valid check on the frequency distribution from open-air sites after correction with the volcanic model.

We compiled a list of radiocarbon dates from archaeological sites in seven northern Wyoming counties that encompass the Bighorn Basin and adjacent mountain ranges. To develop this dataset, we first searched the Wyoming State Historic Preservation Office’s Wyoming Cultural Resource Information System on-line radiocarbon database (Wyoming State Historic Preservation Office, 2008) for dates from sites in the Bighorn Basin. This produced the bulk of the radiocarbon dates included in our sample. Second, we amassed both published and unpublished dates generated through decades of fieldwork conducted by researchers at the University of Wyoming (e.g., Frison, 1991) and other academic institutions (e.g., Finley, 2008). Finally, we solicited unpublished radiocarbon dates from federal and state agencies as well as local contracting firms involved in managing cultural resources in the Bighorn Basin. We made every effort to ensure that the radiocarbon dates included in this study are the result of cultural (i.e., human activity) and not natural processes (i.e., natural fires). This work produced 825 radiocarbon dates from archaeological sites in the Bighorn Basin including 270 dates from caves and rockshelters and 555 dates from open-air sites. Methods of calibration and construction of frequency distributions were identical to those previously described, except dates were aggregated in 500-year intervals to accommodate the smaller size.

In Fig. 3a, we present frequency distributions of uncorrected open-air and rockshelter radiocarbon dates. For the most recent time periods, the datasets are virtually identical, with both peaking between 1000 and 1500 BP. Both datasets also show indication of bias against recent deposits with dramatic drops in frequency associated with the two most recent time intervals between 0 and 1000 BP. For earlier time periods, the two frequency distributions diverge with open-air archaeological components occurring at lower frequencies than rockshelter components for every time interval but one. The only exception is from 12,500 to 13,000 BP where one date occurs in an open-air site, and none occur in rockshelters, a difference we do not consider to be significant. The relative scarcity of open-air components dating to earlier time periods suggests that taphonomic bias has disproportionately affected this portion of the record in the Bighorn Basin.

To correct for taphonomic bias, we perform a simple procedure. Two assumptions are necessary: (1) we assume that the volcanically derived model of taphonomic bias reflects the frequency distribution of sediments in open-air contexts in the Bighorn Basin; and (2) that there was no change in the relative frequency of use of cave and open-air contexts through time in this area. For each time interval of our archeological frequency distribution, we divide the number of observed radiocarbon ages by the value of the taphonomic model for the median year of that temporal interval. For example, there are 13 radiocarbon ages from open-air sites from 6000 to 6500 BP. Plugging the median year 6250 into Eq. (1) for the variable \( t \) yields a result of \( n_t = 19.56 \). This is the number of volcanic deposits surviving from 6250 BP given our model taphonomic bias. Dividing 13 by 19.56 gives an unstandardized corrected frequency of radiocarbon ages of 0.66. We repeated this procedure for all time intervals and standardized the values to the maximum observed corrected value such that the values of the corrected frequency distribution vary between zero and one.
In Fig. 3b, we compare the corrected open-air frequency distribution to the uncorrected frequency distribution of radiocarbon dates from caves and rockshelters. Except for the interval from 8500 to 10,500 BP, the frequency distributions are virtually identical. These curves, built from independent datasets, match in two important ways: (1) they show similarly timed periods of population expansion and contraction, and; and (2) the magnitude of those changes are nearly identical. While the demographic prehistory of Wyoming is not our primary concern in this paper, it is interesting to note that apparent population declines seen in the early Holocene (9500–7000 BP) and again toward later part of the middle Holocene (ca. 4000–2000 BP) both correspond with known periods of drought (Booth et al., 2005; Meltzer, 1999). There are two likely explanations for the divergence of these curves in the early Holocene. First, humans may have increased their use of caves and rockshelters during this time period, and/or it is possible that volcanic model of taphonomic bias is not an accurate proxy record for the age distribution of sediments in the Bighorn Basin.

To assess the second hypothesis, and as a crosscheck on the volcanic model, we constructed a database of quasi-local geologic radiocarbon dates from the Great Plains and adjacent regions. This dataset, compiled from four sources (Baker et al., 2000; Mandel, 1995; Martin, 1992; Meyer et al., 1995), includes 462 radiocarbon dates from alluvial sediments in Oklahoma, Nebraska, Kansas, Missouri, Idaho, and Wyoming. The frequency distribution of local geologic dates is shown in Fig. 4a. Like most frequency distributions from this time period, it increases in frequency toward the present, peaks at approximately 1,000 BP, and declines in the most recent time intervals. Comparing the frequencies of radiocarbon dates by 500-year interval in the local dataset to those of the Bryson et al. (2006) global volcanic dataset shows both distributions to be similar and highly correlated ($n = 38$, $r = 0.952$, $p < 0.001$) when the most recent time intervals (0–1000 BP), those affected by research bias, are excluded (Fig. 4b). That the frequency of radiocarbon dates in the volcanic dataset so closely matches the frequency distribution of dated alluvial sediments in the Great Plains provides additional confidence that the volcanic correction for taphonomic bias may have broad geographic applicability.

To correct the Bighorn Basin open-air date frequency distribution using the local geologic dataset, for each time interval, we
divided the number of archaeological dates by the number of alluvial dates and standardized the resulting frequencies to the maximum value. The resulting frequency distribution (Fig. 4c), not surprisingly, is similar to that produced using the volcanic correction (Fig. 3b). The only major difference occurs from 12,500–13,000 BP where open-air occupations appear very abundant relative to rockshelter occupations. This difference can be attributed to very small sample sizes for both datasets where only one date is present in each. Importantly, the major deviation in the two curves from 8500 to 10,500 BP remains suggesting that it can be attributed to an increased use of rockshelters by humans during this period. Importantly, the frequency of open-air occupations declines over this time period, but the frequency of sedimentary contexts in which they could occur does just the opposite. Therefore, this decline in open-air occupations cannot be explained by taphonomic processes.

One independent constraint on the demographic realism of these curves is the biological capacity for human population growth. If we assume that periods of population expansion evident in these frequency distributions are entirely attributable to intrinsic growth (i.e., no immigration), one can ask whether inferred rates of population growth are within the limits of human reproductive capacities. In short, even those time intervals where the population appears to expand very rapidly are characterized by relatively minimal rates of population growth. For example, from 2250 to 1250 BP there is an apparent 438–668% growth in the size of the population, using the rockshelter and open-air datasets, respectively, but these large jumps in population size correspond only to annual rates of population growth of 0.15 and 0.19% per year, rates that are less than those observed for recent foraging populations (e.g., Early and Headland, 1998; Hill and Hurtado, 1996; Howell, 1979).

4. Case 2: human and mammoth demography in Siberia

A second approach to assessing the validity of the volcanically derived model of taphonomic bias is to compare a prehistoric demographic phenomenon to an analogous modern case to see if similar dynamics are evident. Here we turn to the demography of humans and woolly mammoths in Siberia. As an independent check on the realism of inferred population dynamics of Siberian primates and proboscideans after correction with the volcanic model, we use well-documented aspects of human and elephant demography in Sub-Saharan Africa.

In two recent papers, Ugan and Byers (2007, 2008) have examined spatial and temporal patterning in radiocarbon dates of archaeological sites and proboscidean remains from Europe, Siberia, and North America with the primary intent of examining whether the extinction of mammoths in these regions can be attributed to anthropogenic causes. They begin with the reasonable hypothesis that “if human predation negatively impacted proboscidean populations on continental scales, then... increases in human populations should have been accompanied by corresponding declines in proboscideans” (Ugan and Byers, 2008: p. 70). Comparing temporal frequency distributions of humans and mammoths in these regions, they find that no such relationship exists, providing no support for the overkill hypothesis. To control for taphonomic bias and climate, Ugan and Byers (2007, 2008) regressed temporal frequencies with time and benthic marine δ18O as independent variables, with the residuals of the regression serving as a measure of population density. Using regression with time as a method to erase the effects of taphonomic bias might bring out short temporal scale population changes (Surovell and Brantingham, 2007), but long-term trends in demography can be completely removed. In this section, we demonstrate how a different conclusion about mammoth extinction in Siberia is reached when the volcanic correction is used to adjust for taphonomic bias.

Temporal frequency distributions for Siberian humans and mammoths were taken from Ugan and Byers (2007: supplementary information), published as cumulative probabilities in summed calibrated probability distributions in 500-year intervals from 12,000 to 44,500 BP. Shown in Fig. 5a, the frequency distribution for Siberian humans is like many from the Quaternary, generally increasing through time in a curvilinear fashion, as would be expected if taphonomic bias had affected this record. In contrast, frequencies of radiocarbon dates on mammoths generally do not increase toward the present but instead show considerable fluctuation with no clear relationship to time (Fig. 5b). Nonetheless, both of these records should have been affected by taphonomic bias, and therefore, if these radiocarbon frequency distributions can be used as proxies for population densities, they would suggest that mammoths and humans had very different demographic histories in the late Pleistocene of Siberia.

To correct for taphonomic bias, we again turn to our empirical volcanic model. We repeat the procedure for correction as described in the previous section, and present the results in Fig. 5c and d. After correction, instead of showing relatively continuous population growth over this time period, human populations appear to grow slowly or perhaps remain relatively stable in the time period preceding the Last Glacial Maximum (LGM) (Fig. 5c). A decline in human populations occurs at the LGM followed by population increase with climatic amelioration in the latest Pleistocene. In contrast, mammoth numbers appear to decline more or less continuously over this time period (Fig. 5d). Numerous short scale population fluctuations are evident in both records, but it is unclear to what extent these reflect actual population dynamics or are instead an artifact of sampling in these distributions.

Are the corrected frequency distributions an accurate representation of population dynamics of these two species? The population ecology of humans and proboscideans in Africa can provide some guidance. A handful of studies have found regular inverse relationships between human and elephant population densities in Africa, with the relationship either taking a log-linear or wedge-shaped form (Hoare and du Toit, 1999; Parker and Graham, 1989a,b; see also Barnes et al., 1991, 1995). In northwestern Zimbabwe for example, Hoare and du Toit (1999: p. 636) found that “elephant density is unrelated to human density until a threshold of human density is reached about 15.6 persons/km². After this threshold, resident elephants effectively disappear.” Since the relative probabilities in Fig. 5c and d are intended to be proxies for population density, we can use these data to test whether a similar relationship is evident among Siberian humans and mammoths.

We present a scatter plot of crude population densities of humans and elephants from the Hoare and du Toit (1999: Fig. 3) study in Fig. 6a and b, the inferred population densities of humans and mammoths in Upper Paleolithic Siberia based on the corrected and standardized relative probabilities for each population (Fig. 5c and d). We have removed the three most recent data points in the Siberian dataset due to the dramatic drop in the human frequency distribution after 13,500 BP, no doubt an effect of research bias. For comparison, we present the uncorrected values in Fig. 6c.

The modern African and corrected Siberian data are qualitatively similar. For the corrected Siberian data, there is a significant negative correlation between the inferred human and mammoth population densities in the late Pleistocene of Siberia (n = 63, r = -0.432, p < 0.001) suggesting simply that more people corresponded with fewer mammoths (Fig. 6b). Because wedge-shaped distributions are not well suited to traditional parametric methods of correlation (Beaver, 2004), we confirmed this result using...
a randomization technique in which points were divided into one of four quadrants defined by the median of the maximum and minimum observed value of each variable. Probability was assessed as the proportion of trials in which the minimum number of points assigned to a quadrant was less than or equal to the observed distribution \( n = 2 \) points in the upper right quadrant, and again a significant result was obtained \( \text{trials} = 10,000, \ p = 0.022 \). Following Johnson (2006: p. 218), who observed a similar trend among Australian dingoes and foxes, we also compared mean inferred human population density to maximum mammoth population density for five intervals of human population density in widths of 0.2 units of relative probability, and again found a highly significant negative relationship \( n = 5, r = -0.963, p = 0.009 \).

Regarding such wedge-shaped distributions, Beaver (2004: p. 133) states, “Triangular graphs can be interpreted as resulting from the dependent variable behaving as if limited by the independent variable, rather than being determined or predicted by it.” In other words, this distribution would suggest that mammoth population size was in fact limited by human population size. At times when human population densities were low, mammoth populations were unconstrained by people and fluctuated due to other factors, such as climatic variation. However, as human population densities grew mammoth numbers were increasingly constrained by people, a conclusion also reached by Hoare and du Toit (1999: p. 637) with respect to humans and African elephants, and by Johnson (2006: pp. 217–218) regarding Australian dingoes and foxes.

In contrast, the uncorrected Siberian data (Fig. 6c) suggest no correlation between mammoth and human population densities \( n = 63, r = 0.165, p = 0.196 \) (Ugan and Byers, 2007, 2008). The randomization routine described produced no significant relationship as well \( \text{trials} = 10,000, p = 0.735 \). While African elephants are not a perfect analog for Siberian woolly mammoths, it seems hard to believe that human mammoth hunting had no effect on populations of woolly mammoths (see Mithen, 1993), and there is no doubt that Upper Paleolithic humans in Siberia hunted woolly mammoth (Zenin et al., 2003).

It should be noted that there is another way to examine this problem, one that is not dependent on the use of the volcanic correction. As we noted in our prior paper (Surovell and Brantingham, 2007: p. 1876), ratio data should be immune to the effects of taphonomic bias. In other words, if we assume that both mammoth remains and archaeological sites are equally subject to taphonomic bias, when a ratio of the two is taken, taphonomic bias is cancelled out. Here is a simple example: assume that in the year 21,213 BP humans created a total of 1500 archaeological sites in Siberia, and 3000 mammoths died. In this year, the ratio of mammoths to archaeological sites is 2:1 or 2. If taphonomic bias has removed 98% of those sites and mammoths from existence, today we would be left with 30 archaeological sites and 60 mammoths, and the ratio remains unchanged. It does not matter what percent of sites are lost, as long as it is constant for both, because it factors out of the calculation. Therefore, the ratio of dates on mammoths to archaeological sites in this time series should be an indication of the relative abundance of each, and this method does not rely on the volcanic correction whatsoever.

To perform this analysis, we took the ratio of the cumulative probabilities of the Siberian mammoth and human summed calibrated probability distribution for every time interval from 44,500 to 13,500 BP (Fig. 7). Generally speaking, over this entire time period, mammoths become increasingly rare relative to humans. It should be noted, however, that this phenomenon could be explained multiple ways. Here are three possibilities: (1) human populations grew while mammoth populations remained constant; (2) mammoth populations declined while human populations remained constant; or (3) human populations grew while mammoth populations declined. While ratio data allow us to circumvent the problem of taphonomic bias with minimal
assumptions, they do not necessarily lead to clear explanations of the underlying demographic cause(s) of observed patterning. If the volcanic correction from above is valid, however, it would clearly support the third hypothesis because both growth in human and decline in mammoth populations is evident (Fig. 5c and d).

While our primary intent in this paper is not to investigate the cause of mammoth extinction in Siberia, given these findings, a few observations are warranted. We must bear in mind, however, that no matter how tempting it is to see causation in correlation, it remains possible that some independent underlying factor is responsible for both long-term population expansion in human populations and declines in mammoths, and climate change could be one such factor. Over the relevant time period, three things occur more or less simultaneously, climate warms, human populations grow, and mammoth populations decline, the correspondence of which is a long-recognized problem in assigning cause to the extinction of Siberian woolly mammoths. It should be noted, however, that the decline in mammoths relative to humans appears to begin long before the climatic amelioration of the latest Pleistocene (Fig. 7). Viewing the sum of the evidence that mammoths suffer extinction at the end of this time series at approximately 12,000 BP (Stuart et al., 2002, 2004), that the extinction of Siberian *Mammuthus* fits neatly within a pattern of local extirpation of proboscideans over four continents and 800,000 years (Surovell et al., 2005), and that extinctions of large animals are synchronous with major human population expansions globally (Martin and Steadman, 1999; Surovell, 2007), an anthropogenic cause cannot be ruled out.

5. Conclusion

We began with the simple premise that prehistoric demographic signals can be extracted from temporal frequency distributions if one controls for taphonomic bias. We suggest that a ratio of archaeological to geological contexts can provide such control. In short, archaeological abundance in and of itself is meaningless without reference to the geology in which archaeological finds occur. It is the ratio of archaeological occurrences to opportunities for geological preservation that is a meaningful indicator of demography. An abundance or paucity of archaeological remains from any given time period could indicate high or low population densities depending upon the frequency of sediments from the same time period.
We developed a global taphonomic correction for archaeological frequency distributions dating from 40,000 to ca. 1000 BP based on the Bryson et al. (2006) database of radiocarbon dates associated with volcanic deposits. In contrast with our prior work (Surovell and Brantingham, 2007), this model suggests that the taphonomic rate is not constant, but instead declines with site age. We demonstrated how this model can be used to correct archaeological (and palaeontological) frequency distributions in two case studies in which we use independent lines of evidence to verify the inferred demographic time series.

In the first example, after correction with both the volcanically derived model of taphonomic bias and a quasi-local database of geologic radiocarbon dates, the frequency distribution from open-air archaeological sites in the Bighorn Basin of Wyoming matches very well that of caves and rockshelters, a portion of the archaeological record which should be somewhat protected from the destructive processes that remove sites from the archaeological record. This provides us with some confidence that the primary signal in these frequency distributions is human demography and that the volcanic model of taphonomic bias may be used to correct archaeological frequency distributions in other areas. This conclusion is reinforced by the finding that a database of local geologic radiocarbon dates from Plains alluvial settings matches the volcanically derived model very well.

In the second example, we examined the population dynamics of Siberian Upper Palaeolithic hunter-gatherers and woolly mammoths using the regular negative correlation seen among population densities of modern humans and elephants in Sub-Saharan Africa for verification. We found that a negative wedge-shaped correlation between human and mammoth population densities emerged only after correction with the volcanic dataset. Using the ratio of radiocarbon dates on woolly mammoths to archaeological sites, we showed how a similar result is obtained without use of the volcanic correction at all.

While it may be tempting to see our new and improved model of taphonomic bias as a “magic bullet” that can be used to correct any frequency distribution, we caution against doing so. While it might eventually prove to have broad geographic and temporal applicability, this remains an open question. It only spans the last 40,000 years, and it is not clear to what extent and how far it can be extrapolated back in time. Also, temporal frequency distributions are the sum product of many factors, taphonomic bias and human demography being only two. Chance in sampling and interval selection in histogram creation can add considerable noise to frequency distributions, particularly in the tails and in distributions based on small sample sizes. Research bias can obscure demographic signals as well, a problem which appears to be particularly common in the last millennium, when many frequency distributions show a precipitous drop that is difficult to explain demographically. We have yet to develop methods for accounting for research bias, and for now, reconstruction of human demography for the last millennium in parts of world lacking historical records will be difficult. Research bias may operate to increase or decrease population signals in other times and places as well. We suggest that if future researchers turn to our volcanic model of taphonomic bias to extract demographic histories from frequency distributions that they attempt to confirm those trends with independent lines of evidence. The ideal approach would be to build local databases of geologic radiocarbon dates that can be used to correct for taphonomic bias, and to take into account local variation in sedimentation and erosion not captured by the global volcanic model. Despite these concerns, we are optimistic that the approach we have outlined in this paper brings us closer to seeing patterns in human prehistoric demography than we have ever been before.

Acknowledgments

Thanks are owed to the Wyoming SHPO, Paul Sanders at the Office of the Wyoming State Archaeologist, and the individuals at the state, federal, and private organizations that contributed radiocarbon dates to the project (there are too many to list individually). This work was partially funded by the National Science Foundation (Grant Nos. 0514863 and 0710868).

References


