Assessing precision in dating formulas
Kit W. Wesler
Department of Geosciences
Murray State University

Abstract

Historical archaeology has been proud of its ability to date sites accurately with pipestem and mean ceramic date formulas. The methods are based on fundamental concepts of archaeological analysis: horizon, tradition, and testing by repeated application. These principles allow extension of the methods to prehistoric cases, such as my own Mississippian handle and plate rim formulas for the Lower Ohio Valley. The formulas, however, produce single dates that may be misleading in their apparent precision. I use data sets from historic sites around the Chesapeake Bay, an eighteenth-nineteenth century synagogue site in Jamaica, and a Mississippian site in western Kentucky to assess the statistical precision of formula dates. I suggest that historical archaeologists need to report these dates with their standard deviations and not as single numbers.

A paper presented at the Symposium on Ohio Valley Urban and Historical Archaeology,

Chillicothe, Ohio, 8 March 2003.
Since the early 1960s, dating historic sites by mathematical formulas has been taken as both standard and reliable. The most widely applied formulas use ball clay pipestems and ceramics as the basic data for the calculation of dates. The techniques are so taken for granted by historical archaeologists, in fact, that I’ve used the model to propose formulas for calculating dates using late prehistoric ceramics in the lower Ohio Valley, and I promoted their utility and accuracy as a challenge to further testing. In short, the formulas have been widely accepted because they work.

Recently, though, I've been wondering exactly what we mean when we say, "They work." That is, how accurate are these formulas, and have they ever been rigorously investigated as statistical expressions?

Formula dating began in 1954, when J. C. Harrington published data showing that the bore diameter of white clay pipestems decreased through time. A few years later, Lewis Binford took Harrington's data and derived a simple formula that assumed a regular diminution of bore size, down to a theoretical zero in a year he calculated to be 1931. He proposed that the application of this formula to the average bore size of a sample of pipestems would give a date within the period of occupation of the site. He tested the formula on Harrington's data and also on several sites of his own, and pronounced it accurate.
In the early 1970s, Stanley South (1972) devised a similar formula for ceramics. South assigned manufacture date ranges to the types, and median dates of manufacture, and proposed that the average of the median dates (weighted by the number of fragments of each type) would produce a mean ceramic date that would be approximately the median date of the occupation of the site. South tested the formula on his sites and declared it good, and we’ve been applying it happily ever since.

There are some well-known caveats to using the formula. An obvious problem case would be a site that was occupied from 1700 to 1750, abandoned from 1750 to 1800, then reoccupied from 1800 to 1850: the mean ceramic date would, presumably, be about 1775, when in fact nobody lived there. South pointed out this problem, and warned against an uncritical use of the formula.

There are other potential perturbing factors, some of which colleagues have pointed out. Comparability of sampling methods comes immediately to mind. Also, we can expect that larger samples would yield more representative dates.

Use of ceramics increased through the eighteenth century, so that the average site occupied throughout this period would have many more pieces of pottery late in the century than early, and the ceramic collection would tend to be biased toward a later date. Or, what if a site were occupied by 10 people for 90 years, and 1000 people for the following ten years?
Obviously the mean ceramic or pipestem date would be late. What if another site had the opposite occupation pattern? The two sites would have the same period of occupation, but different mean ceramic dates.

So, the formulas have to be applied with a critical eye and full understanding of the data they are based on. But, even in the best of cases, how accurate are they?

Because we can calculate these dates to several decimal places, I think we get a misleading sense of precision. Does a calculated date of, say, 1698.89 mean September 18, 1698 at 10 p.m.? Obviously this would be to take the number far too literally.

It is interesting to go back to the original articles to see how accurate the designers of the formulas thought they were. Binford wrote that his pipestem dates "fall between the known estimated periods of occupation of the site." If the sites are occupied for a long period of time, and the formula produces a date that lands somewhere in the period, then by Binford's stated standard, the date is a success. This is not really very encouraging.

South (1972) tackled the accuracy question much more explicitly. He compared the calculated dates for his test sites against the known historical dates of the sites' occupation, and found that "the correlation... is seen to be quite high in most
instances" with an "average deviation of plus or minus four years" (his emphasis). He did note that his seventeenth century date was noticeably less accurate, and recommended the formula for eighteenth-century sites.

I submit that we have taken South's (1972) assessment of accuracy of ± four years uncritically, and we need to consider the statistical properties further.

I first began wondering about these issues more than twenty years ago, when I reported pipestem and mean ceramic dates from historic sites in the Chesapeake Bay (1982). Consistently, the pipestem dates were earlier than the mean ceramic dates from the same contexts, often by decades.

I analyzed previously excavated data from Icehouse Point, a site occupied by a very prominent family from about 1665 to 1765. The median occupation date would be about 1715. There were four provenience groups: Tests I(1-4) within the house foundation, a more general excavation around the foundation, Test IIC-D in an overbank refuse deposit, and a general surface collection. The mean ceramic dates are reasonably consistent, especially given the range of sample sizes, ranging from 1739 to 1751. The pipestem dates are consistently earlier than the ceramic dates, and also much more variable--although the two dates from the foundation are quite comparable to each other. But there’s a 40 to 50 year discrepancy between the foundation...
and the other groups’ pipestem dates. Also, the mean ceramic
dates are consistently decades later than the historic median
occupation date.

I explained this as reflecting the foundation’s inclusion
of a builder’s trench, which would have been filled before
domestic occupation of the site began accumulating ceramics,
thus biasing the foundation dates early. The explanation
apparently satisfied my dissertation committee, but I’ve had the
problem nagging the back of my mind ever since.

Then I had a group of data from Londontown. Test I was a
plow-mixed deposit, potentially including materials from ca.
1700 or earlier to the later 20th century. Test II Zone 4 is a
stratum that accumulated between about 1830 and 1968, and not
surprisingly, had only a couple of stray pipestems. Test II
Zone 5 dated between about 1760 and 1830. Test II Zone 5-bottom
was sealed by a shell pavement that I think was created about
1760. In each case, the mean ceramic date is fairly reasonable,
but the pipestem date is noticeably earlier. Only the Zone 5-
bottom pipestem date is at all similar to the mean ceramic date
or consistent with my interpretation of the deposit.

I can make the same comparison with a set of surface-
collected sites from both the Eastern and Western Shores of
Maryland. Of six sites, five show the same pattern of later
ceramic dates than pipestem dates, in one case by more than 100
years. The sixth was an apparent isolated feature with very few ceramics, and we might dismiss the reversal of dates as a vagary of sampling—but again, the two dates give results more than two decades apart.

I’ve always wondered: did I mishandle one or the other set of calculations in some way, or is there a systematic problem with the agreement of these two dating techniques?

Recently I had the opportunity to look at pipestems and ceramics from Neveh Shalom, Jamaica, with these issues in mind.

Neveh Shalom is a synagogue site in Spanish Town. The congregation purchased the site in 1704. The synagogue building was struck by lightning in 1844, and was repaired, and then was damaged by the earthquake of 1907. Other than over-the-wall garbage disposal, there has been no other occupation of the site. Based on these dates and an inspection of many of the ceramics, I take 1700 to 1850 as a round set of dates for the primary occupation represented by the archaeological assemblage. The median occupation date then should be roughly 1775.

The site was investigated in 1998 and 1999 by a team from the University of the West Indies and the Jamaica National Heritage Trust. The investigators cleared portions of the synagogue's foundations, and recovered a large set of artifacts. They found that the soils were thoroughly mixed, so that deposits representing short periods were lacking.
The ceramics include types ranging from early eighteenth-century slipwares to mid-nineteenth century whitewares. Creamwares and delftware were the most numerous, suggesting that the period of heaviest deposition of ceramics was the second half of the eighteenth century.

Pipestems range in bore diameter across the entire range of known sizes, ostensibly representing activities from the early seventeenth to late eighteenth centuries. The greatest frequencies of those measured fell into the 5 and 6/64ths size categories, generally most prevalent in late seventeenth to mid-eighteenth century sites.

Five UWI students studied artifacts and calculated dates from the 1999 excavations representing a block of 4 x 4 m grid squares on the south side of the foundation. Each student studied materials from one or more grid square, and calculated each square's samples independently. Therefore, we have a number of samples, recovered by identical methods, from the same area of the same site. How consistent are the results?

Pipestem dates represent 11 different units. Sample sizes range from 8 to 111, and the dates range from 1697 to 1743, a variation over nearly 50 years. The date calculated on the total collection is 1718. This date is almost 60 years earlier than the assumed historic median date of 1775.

The students calculated mean ceramic dates from six
samples, two of which came from the same grid square (one surface collection, one excavated level). The sample sizes range from 68 to 297, and the dates range from 1757 to 1780. The date calculated from the total collection is 1770.

The mean ceramic dates are fairly consistent. Hearteningly, the two dates from the same unit, E11, were 1763.1 and 1763.5, a very close agreement. The mean ceramic dates are also fairly close to the historic median of 1775, which is admittedly rather arbitrary.

On the face of it, the mean ceramic dates are much more accurate than the pipestem dates—meaning that they are closer to the interpreted historic median and are in better agreement with each other. The pipestem dates are consistently older by decades. The pattern is the same as my Chesapeake dates.

Why?

Well, Spanish Town was occupied by the British in 1655 and by the Spanish considerably earlier, so it is possible that the pipestem collection includes activities before the establishment of the synagogue. This would not explain the Chesapeake data.

There are also biases built into the techniques because of the differences in the materials represented. It’s well known that the pipestem formula is not applicable later than 1780. The smallest pipestem bore, 4/64th inch, gives a calculated date of 1778, so that the latest possible date produced by the
pipestem formula is 1778. Therefore, a site occupied into the nineteenth century cannot possibly produce a nineteenth-century pipestem date: the date would be biased towards the early end of the occupation.

There is a similar bias in the mean ceramic formula, but towards a later date. Nineteenth-century ceramic types are not as closely datable as their eighteenth-century counterparts. Ceramic types introduced circa 1820, whitewares and ironstones, are still in use, and South assigned an arbitrary median manufacture date of 1860. Yellow wares were manufactured from 1840 to 1940, for a median manufacturing date of 1890. If a site yields only a handful of these sherds, they will pull the mean ceramic date toward the later end of the occupation, in the opposite chronological direction from the pipestem date. These tendencies clearly apply to the period of Neveh Shalom.

Given all of these caveats, we can wave our hands and explain away the systematic difference between the pipestem dates and mean ceramic dates in a town that was occupied from the sixteenth century and a site that was occupied into the mid-nineteenth century. This is sort of what I did twenty years ago with my Chesapeake data. But some of my Chesapeake sites included deposits that appear not to post-date 1800. The hand-waving gets pretty frantic at this point.

I'm still not satisfied that we understand the notion of
accuracy in respect to these dating formulas. We have not considered the statistical properties of the data.

Neither Binford nor South assessed the accuracy of their calculations by means of descriptive statistics. Binford suggested offhandedly that the standard deviation of the pipestem date would provide "a rough estimate of the length of time over which the sample was accumulating," which he cannot have thought about much. A standard deviation is a mathematical expression of the variation within a set of numbers, and its relationship to the human behavior that produced an archaeological assemblage is by no means straightforward.

For the Icehouse Point data, standard deviations of the pipestem dates range from 33.7 to 49.1. The mean ceramic date standard deviations range from 16.5 to 45.1. At the upper end, these dates give us a period of uncertainty of nearly a century at one standard deviation.

For other data sets, the results are the same, or even greater. At Londontown, the standard deviations range as high as 54 years: that’s based on two pipestems, of course. The largest standard deviations represent small sample sizes or mixed deposits, but the statistics are eye-opening.

The results for my Chesapeake survey collection are also comparable. For Lower Marlboro, whose main occupation was circa 1750-1830, the mean ceramic date of 1781 looks pretty accurate
until you realize that there’s a 41-year plus or minus on a sample size of 1232. That is, for an occupation I date at 1750-1830, the mean ceramic date has a one-sigma range of 1740-1822. That looks almost like Binford’s idea that the standard deviation gives you an estimate of the occupation period. But that’s not what the mathematics say. The mathematics say that there is a two-thirds chance that the mean ceramic date is sometime within the historic period of occupation.

For the Neveh Shalom data, the standard deviations of the pipestem dates range from a low of 13.5, on a sample of 8, to a high of 48 years. In other words, there is about a two-thirds chance that the latter date, 1689, falls between 1641 and 1737. This represents a range of almost 100 years!

Despite the nearly 50 year variation among the eleven samples' dates, the ranges at one standard deviation overlap. To take the site total, which is the usual practice for calculating a pipestem date, the standard deviation is almost 40 years, so the calculated date at one standard deviation could actually be expressed as a range from 1679-1758.

Similarly, the standard deviations of the mean ceramic dates range from 24 to 39 years. The standard deviation of the totaled sample is 33 years. Thus, the calculated mean ceramic date for the site at one standard deviation could be expressed as 1737-1804.
For the total-collection dates, the standard deviations are more than thirty years on fairly large sample sizes. The two dates are nearly sixty years apart. The one-sigma ranges overlap, and the overlap range is about 1737 to 1757. Does this mean that the “true” median occupation date is within the period of overlap?

Should we be comforted by the fact that the MCD range of 1737-1804 encompasses the historic median date of 1775? Maybe, but I find the magnitude of the uncertainty factor, the standard deviation, very disquieting. Nor, contra Binford, do the standard deviations offer any correlation to the period of occupation of the site.

I can only conclude that the dates that most of us have assumed to provide reasonably accurate estimates of the median occupation dates of a site, in statistical perspective have a considerably larger range of uncertainty than we have realized.

Prehistorians went through a similar development of understanding about chronometric dating methods such as radiocarbon. We now express dates, not as single numbers, but as calibrated ranges, specifying the number of standard deviations incorporated. Thus, to see a date expressed as AD 1350±100, for example, or AD 1250-1450, is quite usual.

I propose that we must begin to report pipestem and mean ceramic dates in the same way, in order to express the
statistical confidence intervals inherent in the calculation. Thus, for example, I would report the dates of our study sample from Neveh Shalom as: pipestem date AD 1718±39, and mean ceramic date AD 1770±34, both at a confidence level of one standard deviation.

What this means for the interpretation of the period and intensity of human behavior on the site is another question entirely, and that's where statisticians turn the problem over to archaeologists.

Thank you.

Acknowledgments

I thank Philip Allsworth-Jones and Karen Spence, of UWI, for access to the Neveh Shalom data, and students Siran Mitchell, Kerry Ann Jones, Nicole Murphy, Kerry Ann Coombs, and Lisan Daley for compiling the data as a class project.
### Maryland dates summary

<table>
<thead>
<tr>
<th>Site</th>
<th>MCD</th>
<th>st. dev.</th>
<th>N =</th>
<th>pipe date</th>
<th>st. dev.</th>
<th>N =</th>
</tr>
</thead>
<tbody>
<tr>
<td>Icehouse Point</td>
<td>1746.46</td>
<td>22.18</td>
<td>388</td>
<td>1694.49</td>
<td>43.00</td>
<td>201</td>
</tr>
<tr>
<td>Tests I(1-4)</td>
<td>1739.90</td>
<td>33.25</td>
<td>31</td>
<td>1698.81</td>
<td>49.11</td>
<td>33</td>
</tr>
<tr>
<td>IIC-D</td>
<td>1744.42</td>
<td>16.49</td>
<td>458</td>
<td>1724.63</td>
<td>33.72</td>
<td>185</td>
</tr>
<tr>
<td>General surface</td>
<td>1751.83</td>
<td>45.15</td>
<td>99</td>
<td>1735.83</td>
<td>37.23</td>
<td>373</td>
</tr>
<tr>
<td>Londontown</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test I</td>
<td>1823.58</td>
<td>49.09</td>
<td>202</td>
<td>1736.72</td>
<td>24.51</td>
<td>20</td>
</tr>
<tr>
<td>Test II Z. 4</td>
<td>1848.20</td>
<td>30.61</td>
<td>25</td>
<td>1702.29</td>
<td>54.11</td>
<td>2</td>
</tr>
<tr>
<td>Test II Z. 5</td>
<td>1799.56</td>
<td>50.14</td>
<td>18</td>
<td>1744.03</td>
<td>20.64</td>
<td>11</td>
</tr>
<tr>
<td>Test II Z. 5-bot</td>
<td>1747.38</td>
<td>10.16</td>
<td>8</td>
<td>1733.80</td>
<td>20.22</td>
<td>17</td>
</tr>
<tr>
<td>Surveyed sites</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower Marlboro</td>
<td>1780.99</td>
<td>40.90</td>
<td>1232</td>
<td>1734.96</td>
<td>28.63</td>
<td>130</td>
</tr>
<tr>
<td>Doncaster</td>
<td>1733.38</td>
<td>18.68</td>
<td>174</td>
<td>1703.38</td>
<td>34.08</td>
<td>35</td>
</tr>
<tr>
<td>Carvel</td>
<td>1765.23</td>
<td>36.28</td>
<td>57</td>
<td>1705.02</td>
<td>23.56</td>
<td>14</td>
</tr>
<tr>
<td>Greenwood Creek I</td>
<td>1853.64</td>
<td>22.48</td>
<td>1562</td>
<td>1731.72</td>
<td>27.74</td>
<td>13</td>
</tr>
<tr>
<td>Tanyard Creek</td>
<td>1726.40</td>
<td>19.76</td>
<td>10</td>
<td>1699.35</td>
<td>32.99</td>
<td>13</td>
</tr>
<tr>
<td>Greenwood Creek II</td>
<td>1727.38</td>
<td>12.08</td>
<td>8</td>
<td>1750.12</td>
<td>27.05</td>
<td>8</td>
</tr>
</tbody>
</table>

### Neveh Shalom dates summary

<table>
<thead>
<tr>
<th>Unit</th>
<th>MCD</th>
<th>st. dev.</th>
<th>N =</th>
<th>pipe date</th>
<th>st. dev.</th>
<th>N =</th>
</tr>
</thead>
<tbody>
<tr>
<td>D10</td>
<td>1756.82</td>
<td>24.11</td>
<td>68</td>
<td>1734.17</td>
<td>29.13</td>
<td>24</td>
</tr>
<tr>
<td>D11</td>
<td>1771.10</td>
<td>39.27</td>
<td>249</td>
<td>1696.60</td>
<td>38.59</td>
<td>102</td>
</tr>
<tr>
<td>D12</td>
<td>1772.50</td>
<td>36.58</td>
<td>104</td>
<td>1723.10</td>
<td>31.96</td>
<td>46</td>
</tr>
<tr>
<td>D13</td>
<td></td>
<td></td>
<td></td>
<td>1737.50</td>
<td>22.47</td>
<td>39</td>
</tr>
<tr>
<td>E10</td>
<td></td>
<td></td>
<td></td>
<td>1734.17</td>
<td>14.88</td>
<td>12</td>
</tr>
<tr>
<td>E11</td>
<td>1763.50</td>
<td>29.20</td>
<td>199</td>
<td>1730.20</td>
<td>34.91</td>
<td>111</td>
</tr>
<tr>
<td>E11</td>
<td>1763.10</td>
<td>24.75</td>
<td>157</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E12</td>
<td>1780.20</td>
<td>33.95</td>
<td>297</td>
<td>1689.30</td>
<td>47.99</td>
<td>56</td>
</tr>
<tr>
<td>E13</td>
<td></td>
<td></td>
<td></td>
<td>1735.77</td>
<td>13.52</td>
<td>8</td>
</tr>
<tr>
<td>F11</td>
<td></td>
<td></td>
<td></td>
<td>1737.99</td>
<td>30.55</td>
<td>15</td>
</tr>
<tr>
<td>F12</td>
<td></td>
<td></td>
<td></td>
<td>1743.10</td>
<td>22.71</td>
<td>15</td>
</tr>
<tr>
<td>F14</td>
<td></td>
<td></td>
<td></td>
<td>1717.60</td>
<td>22.88</td>
<td>20</td>
</tr>
<tr>
<td>total</td>
<td>1770.30</td>
<td>33.64</td>
<td>1074</td>
<td>1717.79</td>
<td>38.98</td>
<td>448</td>
</tr>
</tbody>
</table>