Radiocarbon dating of the Theran eruption

Malcolm H. Wiener, Jason W. Earle
The Institute for Aegean Prehistory, Greenwich, CT, USA

Abstract

Radiocarbon dates from material found in the Theran Volcanic Destruction Level are frequently said to require an earlier date for the event than the date range provided by the traditional chronology based on Near Eastern and Egyptian texts, archaeoastronomy, datable deposits of pumice from the eruption and numerous archaeological interconnections between Egypt, Cyprus and the Aegean. This paper considers a number of problems inherent in the proposed $^{14}$C dating.

Problems of radiocarbon measurement and calibration

Inter-laboratory variation in measurements of $^{14}$C in divided samples, while recently significantly reduced in general, continues to exist. Unresolved disparities in measurements of trees of known dendrochronological date occur among logs from the same site (e.g., Gordion in Turkey for the years 1580, 1570 and 1560 BC); within measurements from trees in the same geographic area (e.g., two forests near the Rhine River at 1550 BC); and between the trees from Germany which form the backbone of the calibration curve and the trees from Gordion in a number of periods (e.g., c. 1325 and 1225 BC, 1080-990 BC, and especially between 850 and 750 BC, when the average offset is 28.8±26.3 radiocarbon years) (Kromer et al., 2010). When radiocarbon measurements of trees of the same known date differ markedly, either at least one of the measurements is erroneous, whether because of differences in pretreatment or for some other reason, or the difference is deemed acceptable in light of the nature of the enterprise, in which case the result is of limited relevance to questions requiring narrow ranges to resolve, as in the case of the Theran measurements described below.

Intra-year seasonal variation in the production of $^{14}$C also affects radiocarbon measurements. The difference between the summer high and winter low today generally ranges between eight and thirty-two radiocarbon years, and may exceed this range. In Egypt, where the growing season of seeds is principally in January-February in contrast to the May-June growing season of the Northern European oaks of the calibration curve, radiocarbon measurements of seeds of known date collected between AD 1700 and 1900 disclosed an offset of 19.5±5 calibrated years earlier than calibration curve dates for this period (Dee, 2010). On Thera, the growing season of seeds is earlier than the calibration curve trees by only a matter of weeks, but the difference is still of potential relevance with respect to a claimed precision of ±7.5 radiocarbon years for the mean error of 28 measurements of 16 seeds and a twig from the destruction level, where a mean error range of ±20 would cause an overlap with the more recent peak of the calibration curve at 1530-25 BC. Further, seasonal variation prior to the impact of industrialisation on the atmosphere may have been greater (Keenan, 2004; Housley et al., 1999; Levin et al., 1992; Levin and Hesshaimer, 2000). In addition, calibration of radiocarbon measurements of seeds (with brief growing seasons within a single year) using decadal tree segments (with years of greatly varying ring thickness) creates an area of uncertainty.

In a study published in Radiocarbon in 1998 the uncertainty was estimated to be in the range of five to eight radiocarbon years (Stuiver et al., 1998). The same study estimated the uncertainty resulting from the interaction of the 11-year sunspot cycle with the secular calibration curve at a standard deviation of eight $^{14}$C years. Accordingly the narrowness of the mean of the error range stated above would be suspect even apart from the likely and potentially more significant impact of the $CO_2$ reservoir effects described below.

Reservoir effects

Reservoir effects of various kinds which add $^{14}$C-deficient carbon from sources in the earth or water to the $^{14}$C content of the atmosphere present major problems with regard to Aegean radiocarbon measurements. Periodic upwelling of...
deep seawater releases 14C-deficient carbon into the atmosphere where it may be absorbed by plants, thus skewing measurements from coastal sites toward older dates (Rapp and Hill, 2006). Several studies suggest the possibility of upwelling of 14C-deficient carbon from the Aegean, either via the periodic exchange of water with the Black Sea which is rich in 14C-deficient carbon (Keenan, 2002), or the release of 14C-deficient carbon from underwater volcanic vents, one of which is located 7 km north-north-east of Thera (Carey et al., 2013; for vents near Melos see Pain, 1999). The critical sources of 14C-deficient carbon in various areas include volcanic and non-volcanic gas vents, geothermal fields, and general soil degassing (Frezzotti et al., 2009). In Italy the area of terrestrial CO2 emissions stretches from Tuscany to Sicily, and from the Tyrrhenian Sea to the Apennines (Frezzotti et al., 2009). Similar emissions have been observed in areas of southern Italy, and on the volcanic island of Stromboli (Rogie, 1996; Chiodini et al., 1999, 2004; Rogie et al., 2000; Cardellini et al., 2003; Gambardella et al., 2004; Minissale et al., 1997; Carapezza et al., 2009; Brandherr, 2008; Nijboer and van der Plicht, 2008). Reported radiocarbon dates from Italian sites whose historical contexts are clear are often 70-250 years too early (Turfa, 2006). A similar phenomenon is reported with regard to Iceland, where volcanic/geothermal effects are thought to be the cause of radiocarbon dates 100-200 years earlier than presumed historical dates for the earliest European occupation levels (Sveinbjörndóttir and Heinemeier, 2011). At Sulphur Banks on Hawaii, three living tree ferns and one Ohia leaf growing 1 to 5 miles from a volcanic vent gave radiocarbon ages between 81 and 303 years (Chatters et al., 1969). In addition, groundwater that has been in contact with soil degassing or with limestone, a notorious source of 14C-deficient carbon, if absorbed by plants or consumed by living creatures, sometimes results in measurements which provide dates that are older than true dates (Mörner and Etiope, 2002; Fischer and Heinemeier, 2003; Rapp and Hill, 2006). While in a few reported cases measurements of contemporary plants growing immediately adjacent to a current point source of volcanic carbon have produced radiocarbon ages far too early (Bruns et al., 1980; Pasquier-Cardin et al., 1999), many measurements of short-lived samples from areas of gas emissions have provided dates in the range of 70-300 years earlier than the dates appropriate for the historical contexts of the samples.

Statistical and calibration issues

The combination and calibration of radiocarbon measurements, based on Gaussian and Bayesian statistics and conventions employed in the various calibration programs, pose many issues. AMS laboratories employ different methods of analyses, some unpublished. The method chosen can have a significant impact on the reported findings and conclusions. The commonly used standard error of the mean can be too small or too large by a factor of two, but on average is too small; moreover, radiocarbon intercomparisons assume a Gaussian bell-curve distribution of the results even though this is known not to be the case (Palonen et al., 2010). Boaretto et al. (2005) noted that 14C measurements generally do not conform to the ideal standard distribution; rather the actual distribution under extensive replication is wider than the quoted errors, even after the exclusion of outliers. The quoted error denotes at best the variability within the same measurement run. Sometimes even the meaning of basic terms differs between laboratories. The OxCal, Calib and Demokritos Laboratory (Athens) programs differ, for example, in their use of probability. In the OxCal program 80% probability means 80% of 100, whereas in the Calib program it means 80% out of the two-sigma standard deviation of 95.4%, and the Demokritos program is unique, generally giving broader ranges (in our view appropriately so), but is susceptible of reaching results different from those of the other programs from the same data (Y. Maniatis, personal communication of 6 April 2010). Other differences exist as well. Another radiocarbon laboratory, Beta Analytic, employs the Pretoria Calibration Procedure which smooths actual calibration curve measurements employing a spline (Talma and Vogel, 1993), whereas INTCAL04 and INTCAL09 utilise a Random Walk Model, but smooth the results by incorporating data from earlier and later decadal measurements.

Most importantly, all calibration programs narrow stated error ranges in direct response to the number of measurements included – the more measurements, the narrower the range – without great regard to the consistency or inconsistency (even when substantial) of the measurements themselves, provided the measurements meet the chi-square test. Note that the OxCal 4.1 programme, described in Ramsey (2009), deals with disparate measurements of samples from the same event horizon by giving significantly less weight to measurements on the boundaries as distinguished from those in the center, while acknowledging that the definition of outliers is somewhat arbitrary. Differing statistical treatment of outliers can produce significantly different dates, as illustrated in Ramsey (2009). The chi-square test itself, however, is logically dependent on the existence of Ward and Wilson Case I (Ward and Wilson, 1978) which specifies that the method is appropriate only when the measurements are known to represent the same event (e.g., measurements from the bones of a single corpse). The test has often been utilised when the evidence is ambiguous in this regard, however. For example, the Ward and Wilson Case I analysis has been applied to seeds from the Volcanic Destruction Level at Thera (Manning et al., 2006). Here, however, it is not clear whether some of the seeds were collected before the preliminary eruption which caused the populace to flee, or between that point and the final phase, during which some returned to try to rebuild or rescue items of value, nor whether all the seeds measured came from the same field or fields of varying distance from preexisting point, fault, or field of gas emissions capable of producing distorting reservoir effects. In any event, the Theran measurements produced radiocarbon age ranges whose central dates were as much as 350 years apart; dates for one species (grass pea) 215 years apart, and dates for another (barley) 97 years apart (Manning et al., 2006). However, because 28 measurements (some of which preceded modern pretreatment and high-precision dating methods) were available, the stated mean error range was only ±7.5 radiocarbon years. In a recent issue of Radiocarbon, Russell et al. (2011) in their discussion of errors associated with 14C analysis procedures highlight inappropriate errors placed on the age measurements, observing that “this has to be a realistic estimate of the error and should not be based solely on counting statistics. Even apart from the likelihood of at least some reservoir effects on measurements of material from Thera, the narrowness of the stated error range reflects the statistical conventions employed rather than the underlying reality of the volcanic nature of Thera and the surrounding sea. The claimed precision is essential, however, to the proposition that radiocarbon dates strongly support Manning’s proposed Aegean Long Chronology because, as Manning (2005) has previously noted, any age in the 3320s BP is potentially compatible with historical dating in light of the oscillating calibration curve (Figure 2). With an error band of ±25 rather than ±7.5, the stated radiocarbon age of 3344.9 BP would be potentially compatible with a date for the eruption of c. 1325 BC, even without any correction for any potential seasonal/Regional offset or more significantly, for the reservoir effects of 14C-deficient carbon (Wiener, 2009a; 2009b).

Ramsey has noted (personal communication of 30 December 2009) that there are two potential methods within the OxCal program of addressing the sources of uncertainty noted. The first is to insert a range which the investigator – e.g., the archaeologist submitting the samples – regards as a minimal acceptable error range, based on his/her understanding of the circumstances concerning the samples.
submitted for radiocarbon measurement. In practice this is never done, inasmuch as archaeologists seldom understand the complex statistical issues involved. The second method is to introduce an asymmetric adjustment to the error range toward older dates to allow for estimated local reservoir (and/or regional/seasonal) effects. Such putative effects are also impossible to quantify, however. Considerable sophistication in relevant sciences and statistics, as well as a degree of courage, would be required of any archaeologist choosing such a course. It would also be possible to employ the earliest plausible date indicated by the well-founded historical date range for the Theran eruption as a Bayesian boundary, but it is precisely this historical dating which the radiocarbon determinations in question are said to challenge.

A recent attempt by Friedrich et al. (2006) to date a branch from an olive tree found in the volcanic tephra on Thera c. 7 km from Akrotiri presents its own problems. First, there is no way of knowing whether the olive branch was living at the time of the eruption. Aegean olive trees are often covered in long-dead branches since their owners hesitate to remove them for fear of damaging the tree (O. Rackham, personal communication of 11 May 2008; H. Blitzer, personal communication of 23 July 2008). Second, the claim that the investigators have been able via X-ray tomography to count year rings on the branch is severely challenged by reports from dendrological laboratories that rings of olive trees are impossible to count with any confidence. In blind tests no two laboratories could agree on the number of rings on olive tree branches recently collected on Thera. Moreover, olive trees do not form anything approaching annual rings even if agreement could be achieved on the number (Humbel, 2009; Cherubini et al., 2003), which vitiates also the wiggle-match to the calibration curve claimed by Friedrich et al. (2006). The claim by Friedrich et al. (2006) that they have allowed for uncertainties in the ring count of up to 25% in this regard seems disingenuous, inasmuch as it refers to an allowance of ±3 rings in an asserted 12-ring segment, given that no such precision is achievable, even apart from the fact that olive trees do not make annual rings. The fact that the 4 segments of the branch studied by Friedrich et al. (2006) gave radiocarbon measurements which descended in order is also not germane, since this would be true in any event – e.g., if the 4 segments each contained a 1% reservoir of 14C-deficient carbon resulting in measurements c. 80 years too old, the measurements of the four segments would still descend in order.

Figure 1. Location of Thera, Greece.

Figure 2. Oscillating calibration curve of radiocarbon dates.

Other archaeometric evidence for the dating of the eruption

Four hundred and fifteen samples of pumice and tephra from the Minoan eruption of Thera have been found a minimum of 15 sites in Egypt, the Levant, Cyprus and the Aegean, all of them from New Kingdom contexts, whereas the 27 examples of pumice from earlier contexts all come from non-Theran eruptions. The investigators note that if the Aegean Long Chronology were correct, it would indeed be most peculiar a phenomenon that pumice from the Minoan Santorini eruption were abundantly available along the shores of the Eastern Mediterranean, yet for some reason had been left unnoticed and unused by the local inhabitants for 100–150 years (Steinhauser et al., 2010), to which we may add that it would also be peculiar if so valuable a substance with 14 known uses in antiquity, and especially useful in metallurgy (Wiener and Allen, 1998), was ignored for a century before traders brought it to all the Egyptian, Levantine, Cypriot and Aegean sites where it is attested, as suggested by Manning (1999) (cf. also Bietak, 2004). At one of the sites, Tell el ‘Ajjul near Gaza, Theran pumice, Egyptian New Kingdom pottery, and Cypriot White Slip I pottery similar to a bowl recovered from the Theran Volcanic Destruction Level are accompanied by four radiocarbon dates from seeds which center on c. 1525 BC., the historically and archaeologically appropriate date for the eruption, while only touching the Aegean Long Chronology eruption date range at the limit of the two-sigma probability distribution (Figure 3).
Conclusions

Radiocarbon determinations for the date of the Theran eruption may center somewhat earlier than the well-established historical date range due to: i) factors specific to the area of the eruption; and ii) uncertainties inherent in the process of calibration and the assumptions on which the estimates of probability and precision are based. Specific factors include i) the oscillating nature of the calibration curve between c. 1615 and 1525 BC; ii) the particular reservoir effects likely present in measurements of samples from volcanic Thera; and iii) seasonal/regional differences between Aegean seeds/trees and northern oaks in cases where an asserted error range of less than ±20 14C years is critical to the conclusion stated. Inherent problems include i) the irregular character of the calibration curve and the intrinsic uncertainties of calibration in general; ii) the statistical paradigms employed, which narrow error bands in response to the number of measurements made, however disparate, provided they meet the chi-square test, the appropriateness of which rests on the assumption that the requirements of Ward and Wilson Case I have been met; and iii) the general failure to convey to archaeologists who submit samples for dating that the meaning of the term probability in statistical usage differs from the meaning in general usage, where the term implies that all sources of uncertainty and all contradictory evidence have been considered. (The terms likelihood and confidence display similar differences between statistical and standard English meaning). Archaeologists submitting samples typically are unaware of the existence of such issues, notwithstanding that articles dealing separately with many of the problems discussed have appeared over the years in radiocarbon, statistical or general archaeological journals. Radiocarbon laboratories accordingly have a professional obligation to inform their clients clearly of the uncertainties and limitations of radiocarbon dating, both with regard to the particular samples submitted and in general. The duty is magnified when proposing a date range for an event as important as the date of the Theran eruption, which affects the chronology and history of the entire Eastern Mediterranean world.


Article

CO2 emission from some vents in central Italy. J Geophys Res 105:8435-45.
Steinhauser G, Sterba JH, Oren E, Foster M, Blichler M, 2010. Provenancing of archaeo-
logical pumice finds from North Sinai. Naturwissenschaften 97:403-10.
Wiener MH, 2003b. Time out: the current impasse in Bronze Age archaeological dat-
Wiener MH, 2007. Times change: the current state of the debate in Old World chronolo-