Drawing the line: Bayesian modelling and absolute chronology in the case-study of the Minoan eruption at Thera

Tiziano Fantuzzi and Federico Antolini

tiziano.fantuzzi@gmail.com – federico-antolini@uiowa.edu

Abstract: In the light of recent developments on the accuracy of the radiocarbon chronology of the LM IA Theran Eruption, this paper intends to clarify the difference between radiocarbon chronology and the statistical interpretation of radiocarbon results, suggesting a prudent application of statistical analysis to both radiocarbon dates and chronology problems. The authors consider the estimation of the Minoan eruption at Thera by means of a Bayesian model as an example of overconfidence on a statistic-based methodology. The same analysis will be applied to a simulated set of non-coeval radiocarbon dates with the purpose of showing that the result is adverse to the initial assumptions on the dates. On one hand, this occurs because the model fails to recognise the uncertainties due to the shape of the calibration curve in the interval of the considered dates. On the other hand, the outcome of a χ^2 test is improperly used to ascertain the contemporaneity of radiocarbon dates. The availability of simple statistical tools does not prevent their indiscriminate application and can lead to inaccurate conclusions. The authors suggest that an approach to the chronological problem should consider the environmental context of samples as carefully as their radiocarbon determinations.

Keywords: Theran eruption – Late Bronze Age Aegean chronology – Radiocarbon chronology – Bayesian modelling – Hypothesis testing

The absolute chronology of the LM IA eruption at Thera has been debated over more than three decades.[1] The textual-archaeological synchronisms between Egypt and Late Minoan Crete (mainly Cypriot White Slip I ware and pumice lumps found in LBA contexts in Egypt and Levant) have been used by the supporters of the so-called 'Traditional' or 'Low chronology'[2] to build a whole chronological framework. Consequently, the date of the Theran eruption has been variously attributed to 1530-1500 BC (corresponding to mature LM IA) on purely archaeological/historical grounds. However, since the late 70s the whole 'traditional' reconstruction of archaeologically attested synchronisms has been questioned by analyses of the radiocarbon dates (henceforth RDs) obtained from samples buried by the ejecta of the eruption.

By now, it is generally accepted that the radiocarbon dates of the final phases of occupation at Akrotiri[3] should imply a shift of some 50 to 120 years away from the traditional chronology.[4] In fact, the vast majority of the RDs obtained from samples collected in contexts sealed within the volcanic destruction level (henceforth VDL) shows results that are consistent with an eruption date in the XVII cent. BC.[5]

However, when individually calibrated (Figure 1), 24 out of 28 results in the Akrotiri VDL dataset can also be consistent with a date in the XVI cent. BC (consistent with the archaeological low-chronology), at a confidence interval of 95.4%.[6]

To overcome this impasse, proponents of the high chronology have tried to combine the radiocarbon dates from Akrotiri (and other sites) with the use of Bayesian modelling programs, and in particular OxCal's R_Combine and Tau_Boundaries. A Bayesian model for chronology estimates the probability that an event X occurred in a certain period *t* based on dating measurements RDs of related objects (posterior probability). This is done by combining the probability of finding measured RDs for objects belonging to *t* (likelihood) and a prior probability, i.e. information on the objects not derived from the measurement.[7] R_Combine then calculates the only possible chronological interval where all the RDs fit together,

---

[1] See Kemp – Merrillees 1980.

[2] Warren – Hankey 1989; Wiener 2001; 2003; 2006; 2009; 2014.

[3] Bronk-Ramsey *et al*. 2004; Manning 1999; Manning *et al*. 2006.

[4] Bietak 2013; Bronk-Ramsey *et al*. 2004; Höflmayer 2012; 2015; Manning 1999; Manning *et al*. 2006; 2014; Warren – Hankey 1989; Warren 2006; 2010; Wiener 2001; 2003; 2006; 2009.

[5] Bronk-Ramsey *et al*. 2004; Höflmayer 2012; 2015; Manning 1999; Manning *et al*. 2006; 2014.

[6] Fantuzzi 2007; 2009; Wiener 2003; 2009; 2014.

[7] Bronk-Ramsey 2009.

assuming that they all represent the radiocarbon date of the eruption event, i.e. that they are all contemporaneous, while Tau_Boundaries take into account the possibility that some of the dates may be earlier.[8] The absolute date of the eruption has been consequently suggested to be 3345±8 BP which, in calibrated terms, would date the eruption between 1643 and 1621 BC (Figure 2). This value is very close to the upper limit of the oscillating portion of the calibration curve, and 50 to 120 years earlier than what shown by the archaeological interrelations.[9]

However, the accuracy of this radiocarbon chronology relies on the nature of the prior information supplied. In the specific case of R_Combine this presumes that: 1) all the sampled organisms were contemporaneous, 2) they all died in a moment close to the final eruption, 3) there is no intrusive material in the tested samples, and 4) no sample in the dataset presents alterations or reservoir effects. The use of R_Combine requires that all these conditions are met *a priori*. Thus, it is not possible to use R_Combine to verify the reliability of those conditions without incurring a serious vicious circle. As for conditions 1) and 2), such programs can be useful to work out whether the radiocarbon determinations can represent the same event given their contemporaneity, but not to show if they represent a single event, let alone the real age of the event.

Therefore, 1) radiocarbon dates of different real ages may misleadingly be considered contemporaneous, and 2) the results in calibrated terms would be consequently altered, as shown by the following example.

Let us take into consideration a set of 40 radiocarbon dates (Figures 3-4) from four known ages (1660, 1620, 1600 and 1530 BC), all with a standard deviation of ±30 ^{14}C years (1 sigma), that overlap each other at least at 3 sigma (respectively, 3367±30, 3340±30, 3290±30 and 3287±30 BP).

Let us suppose we ignore the known ages and want to use R_Combine to verify if they can represent the same event. The dataset passes the χ^2 test, and R_Combine allows the combination of RDs producing as a result the interval 3321±4.7 BP. In calibrated terms, that would set the final date between 1626 and 1566 BC. The program is not meant to ascertain if the dates represent the same radiocarbon year-event, as this is the basic assumption. Under this assumption, our dataset passes the test, and the final (calibrated) result is in any case shifted by some decades: the known final age (1530 BC) is not even within the 95.4% confidence interval (Figure 4).

The problems in combining radiocarbon results for dating single events are well known since the 80s.[10] From the mathematical point of view, the potential problems within the accuracy of Bayesian high-precision radiocarbon dating have already been described in detail by Peter Steier and Werner Rom,[11] and more recently also by Bernhard Weninger *et al.*[12] The authors of a recent handbook on radiocarbon dating acknowledge the statistical validity of such protocols for RDs combination, but warn that:

> *Attempts to employ ^{14}C determinations without appropriate regard for the full range of inherent and effective variability in ^{14}C-inferred ages can result in unrealistic chronological expectations and a spurious precision.*[13]

Generally, obvious though it may sound, statistical methods should be approached cautiously when applied to archaeological problems. While mathematical and statistical analyses as well as computer techniques can provide invaluable benefits and result in sounding advancement of the discipline, an over-confidence in such methods represents a realistic threat to the discipline itself. The introduction of new scientific methods always comes with the risk of a rise of fascinating trends for data analysis, which can misguide the actual objective of research. The problem of misuse or abuse of statistical analysis in archaeology has been known for a long time,[14] and it might still be a problem nowadays.

The contemporaneity of radiocarbon age determinations, which is the prior statement in the evaluation of the Thera eruption event described above, is often the object of statistical tests. Therefore, it may be worth to clearly recall what a statistical test means. Hypothesis testing verifies the plausibility of an initial hypothesis (null hypothesis) in contrast with an alternative hypothesis. This consists in proving whether or not the null hypothesis is improbable and in quantifying its improbability. So, whichever testing technique is used, hypothesis testing is not meant to prove the null hypothesis. A typical application of hypothesis testing to archaeology is the comparison of age determinations, in which the null hypothesis is the contemporaneity of two or more samples.[15] The result of the testing cannot provide an argument in favour of contemporaneity, but only against it. This should always be kept in mind as too often, and definitely not only in archaeology, the result of a test is misinterpreted as a neat and definite response between the null and the alternative hypothesis. On top of that, the availability of an apparently clear and easy-to-obtain result generates overconfidence in the statistical tool and leads to a perfunctory application of the same.

---

[8] Höflmayer 2012; 2015.

[9] Bietak 2013; Bronk-Ramsey *et al.* 2004; Höflmayer 2012; 2015; Manning 1999; Manning *et al.* 2006; 2014; Warren 2006; 2010; Wiener 2001; 2003; 2006; 2009.

[10] Weninger 1986; Taylor 1987.

[11] Steier – Rom 2000.

[12] Weninger *et al.* 2011.

[13] Taylor – Bar-Yosef 2014, 159-160, italics in the original.

[14] Thomas 1978; Ammerman 1992.

[15] Ward – Wilson 1978.

Moreover, several aspects of the radiocarbon evaluation cannot be captured by relatively simple statistical analyses. Even excluding *a priori* possible stratigraphic/contextual and treatment/counting errors by the laboratory, the variability in measured ^{14}C ages may be caused by a long series of effects, such as:

1. The seasonal variability in ^{14}C absorption by plants, depending on the growing season, which may cause alterations from 8 to 32 radiocarbon years;

2. The local variability of the ^{14}C atmospheric content which is not recognized by the calibration curve (which is a smoothed, approximated band for the whole northern hemisphere);

3. Reservoir effects deriving from a) deep sea water upwelling and degassing, b) volcanic ventings, causing the absorption of old carbon from depauperated CO_2 by the sampled plants.

In particular, the presence of sources of volcanic CO_2 on Thera has been proved beyond any doubt,[16] although it is still unclear to what extent it might have affected the grains and olives found at Akrotiri. In any case, the radiocarbon dates for the Minoan eruption fall into a part of the calibration curve where a difference of only 20 radiocarbon years would be enough to shift it from the low to the high chronology or *vice versa*. All the above observations show that using the RDs from the VDL as an argument for high chronology (i.e. an eruption date no later than 1627-1600 BC) is at best over-optimistic, and represents a potentially misleading example for archaeologists dealing with similar situations.

Bibliography

Ammerman, A. J. 1992. Taking stock of quantitative archaeology, *Annual Review of Anthropology* 21, 231-255.

Bietak, M. W. 2013. Antagonisms in Historical and Radiocarbon Chronology. In A. J. Shortland – C. Bronk-Ramsey (eds.), *Radiocarbon and the chronologies of ancient Egypt*, 76-109. Oxford.

Bronk-Ramsey, C. 2009. Bayesian Analysis of Radiocarbon Dates. *Radiocarbon* 51.1, 337-360.

Bronk-Ramsey, C. – Manning, S. W. – Galimberti, M. 2004. Dating the Volcanic Eruption at Thera. *Radiocarbon* 46.1, 325-344.

Fantuzzi, T. 2007. The Debate on Aegean High and Low Chronologies: an Overview through Egypt. *RdA* 31, 53-65.

Fantuzzi, T. 2009. The Absolute Chronology of the Egyptian S.I.P.- New Kingdom transition and its Implications for Late Minoan Crete. *CretAnt* 10.2, 477-500.

Höflmayer, F. 2012. The Date of the Minoan Santorini Eruption: Quantifying the “Offset”. *Radiocarbon* 54, 435-448.

Höflmayer, F. 2015. Carbone-14 Comparé: Middle Bronze Age I (IIA) Chronology, Tell el-Dab'a and Radiocarbon Data. In J. Mynářová – P. Onderka – P. Pavúk (eds.), *There and Back Again - the Crossroads II*, 265–95. Prague.

Kemp, B. – Merrillees, R. 1980. *Minoan Pottery in Second Millennium Egypt*. Mainz am Rhein.

Manning, S. W. 1999. *A Test of Time*. Oxford.

Manning, S. W. – Bronk-Ramsey, C. – Kutschera, W. – Higham, T. – Kromer, B. – Steier, P. – Wild, E. M. 2006. Chronology for the Aegean Late Bronze Age 1700-1400 B.C. *Science* 312, 565-569.

Manning, S. W. – Höflmayer, F. – Moeller, N. – Dee, M. W. – Bronk-Ramsey, C. – Fleitmann, D. – Higham, T. F. G. – Kutschera, W. – Wild, E. M. 2014. Dating the Thera (Santorini) Eruption: Coherent Archaeological and Scientific Evidence Supporting a High Chronology. *Antiquity* 88, 1164–1179.

McCoy, F. W. – Heiken, G. 2000. The Late-Bronze Age explosive eruption of Thera (Santorini), Greece: regional and local effects. In F. W. McCoy – G. Heiken (eds.), *Volcanic Hazards and Disasters in Human Antiquity*, 43–70. Geological Society of America, Special Paper 345. Boulder, CO.

Reimer, P. J. – Bard, E. – Bayliss, A. – Warren Beck, J. – Blackwell, P. J. – Bronk-Ramsey, C. – Buck, C. E. – Cheng, H. – Lawrence Edwards, R. – Friedrich, M. – Grootes, P. M. – Guilderson, T. P. – Haflidason, H. – Hajdas, I. – Hatté, C. – Heaton, T. J. – Hoffmann, D. L. – Hogg, A. G. – Hughen, K. A. – Felix-Kaiser, K. – Kromer, B. – Manning, S. W. – Niu, M. – Reimer, R. W. – Richards, D. A. – Marian Scott, E. – Southon, J. R. – Staff, R. A. – Turney, C. S. M. – Van der Plicht, J. 2013. Intcal13 and Marine13 Radiocarbon Age Calibration Curves 0-50,000 Years Cal BP. *Radiocarbon* 55.4, 1869-1887.

Steier, P. – Rom, W. 2000. The Use of Bayesian Statistics for ^{14}C Dates of Chronologically Ordered Samples: A Critical Analysis. *Radiocarbon* 42.2, 183-198.

Taylor, R. E. 1987. *Radiocarbon Dating. An archaeological perspective*. Orlando.

Taylor, R. E. – Bar-Yosef, O. 2014. *Radiocarbon Dating. An archaeological perspective* (2nd edition). Walnut Creek.

Thomas, D. H. 1978. The awful truth about statistics in archaeology. *American Antiquity* 43.2, 231-244.

Ward, G. K. – Wilson, S. R. 1978. Procedures for comparing and combining radiocarbon age determinations: a critique. *Archaeometry* 20.1, 19-31.

Warren, P. M. – Hankey, V. 1989. *Aegean Bronze Age Chronology*. Bristol.

Warren, P. M. 2006. The date of the Thera eruption in relation to Aegean-Egyptian interconnections and the Egyptian historical chronology. In E. Czerny – I. Hein – H. Hunger – D. Melman – A. Schwab (eds.), *Timelines. Studies in Honour of Manfred Bietak*, 305-321. Leuven – Paris – Dudley.

Warren, P. M. 2010. The absolute chronology of the Aegean circa 2000 B.C.-1400 B.C. A summary. In *Die Bedeutung der minoischen und mykenischen Glyptik VI*, 383-394. Mainz.

Weninger, B. – Edinborough, K. – Clare, L. – Jöris, O. 2011. Concepts of Probability in Radiocarbon Analysis. *Documenta Praehistorica* 38, 1-20.

Weninger, B. 1986. High-precision calibration of archaeological radiocarbon dates. *Acta Interdisciplinaria Archaeologica* 4, 11-53.

Wiener, M. H. 2001. The White Slip I of Tell el Dab’a and Thera. Critical Challenge for the Aegean Long Chronology. In V. Karageorghis (ed.), *The White Slip Ware of Late Bronze Age Cyprus*, 195-244. Wien.

---

[16] McCoy – Heiken 2000.

Wiener, M. H. 2003. Time Out: The Current Impasse in Bronze Age Archaeological Dating. In K. P. Foster – R. Laffineur (eds.), *Metron. Measuring the Aegean Bronze Age,* 363-399. Aegaeum 24. New Haven.

Wiener, M. H. 2006. Chronology Going Forward (With a Query About 1525/4 BC). In E. Czerny – I. Hein – H. Hunger – D. Melman – A. Schwab (eds.), *Timelines. Studies in Honour of Manfred Bietak,* 317-328. Leuven – Paris – Dudley.

Wiener, M. H. 2009. The State of the Debate about the Date of the Theran Eruption. In D. A Warburton (ed.), *Time's Up: Dating the Minoan Eruption of Santorini*. 197-206. Aarhus.

Wiener, M. H. 2014. Radiocarbon Dating of the Theran Eruption. *Open Journal of Archaeometry* 2, 5265.

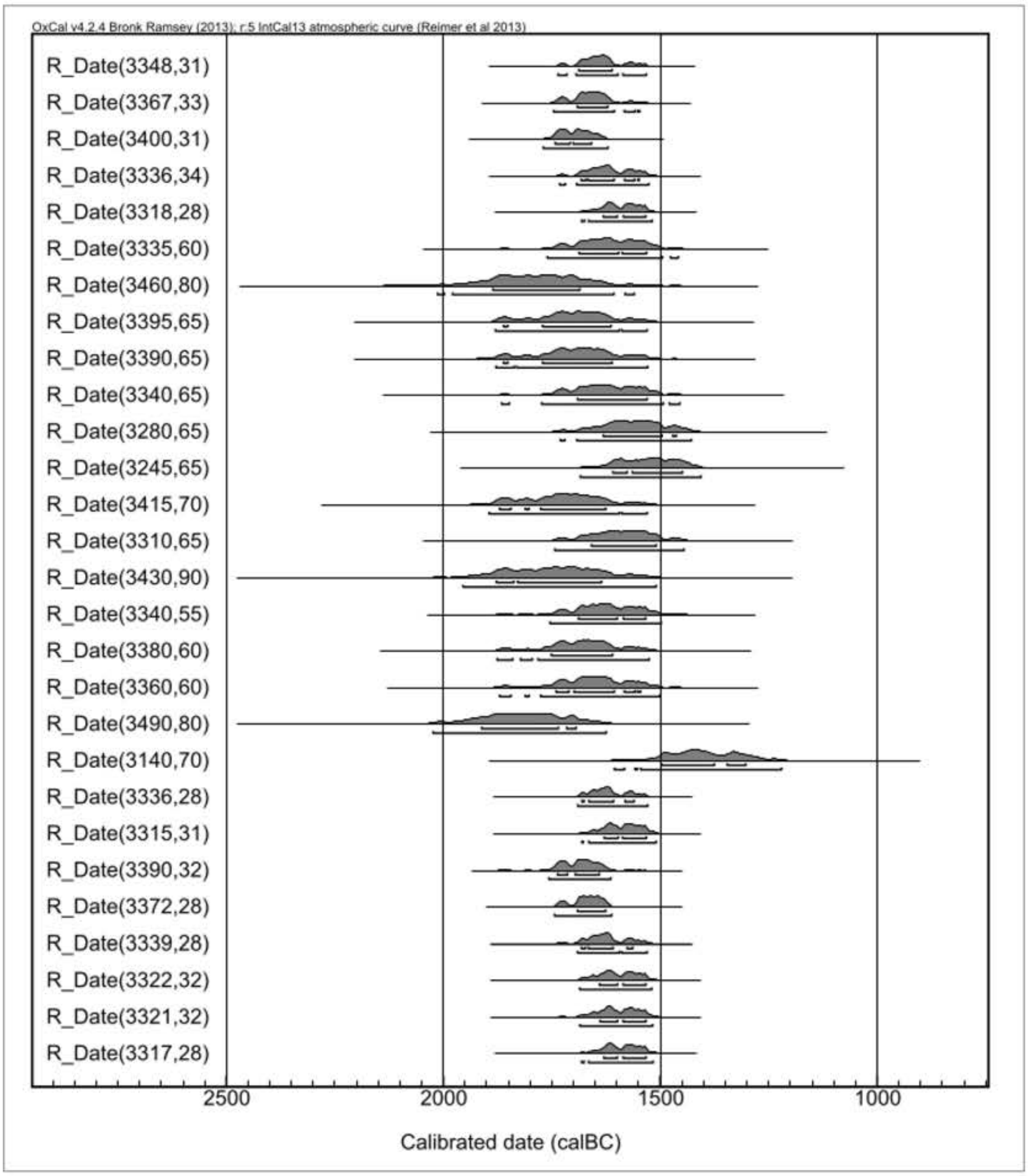

Figure 1. Uncombined calibrated results of the Akrotiri dataset (data from Manning *et al*. 2006).

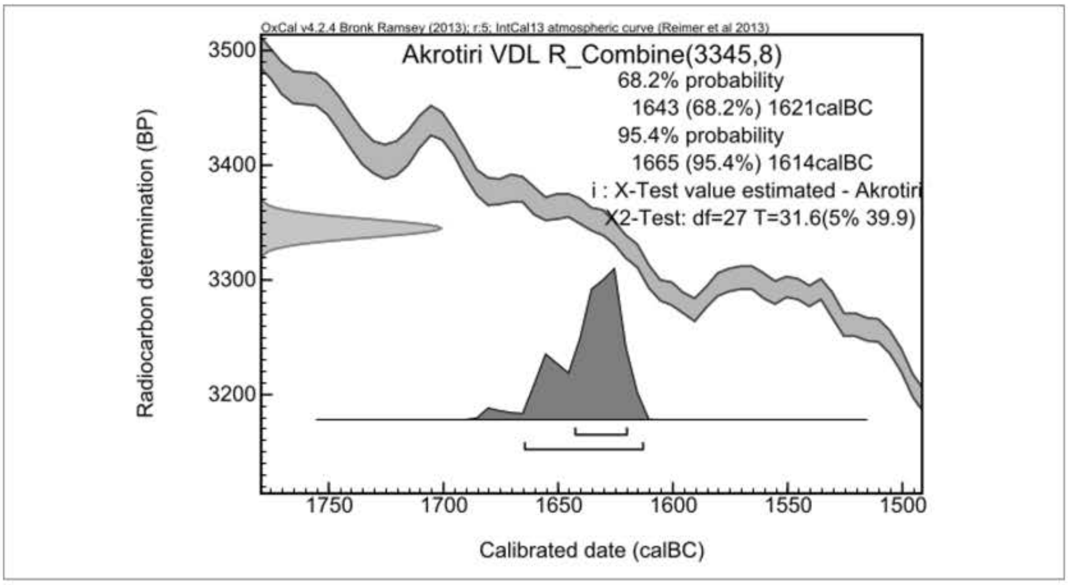

Figure 2. R_Combine result of the Akrotiri dataset (data from Manning *et al*. 2006).

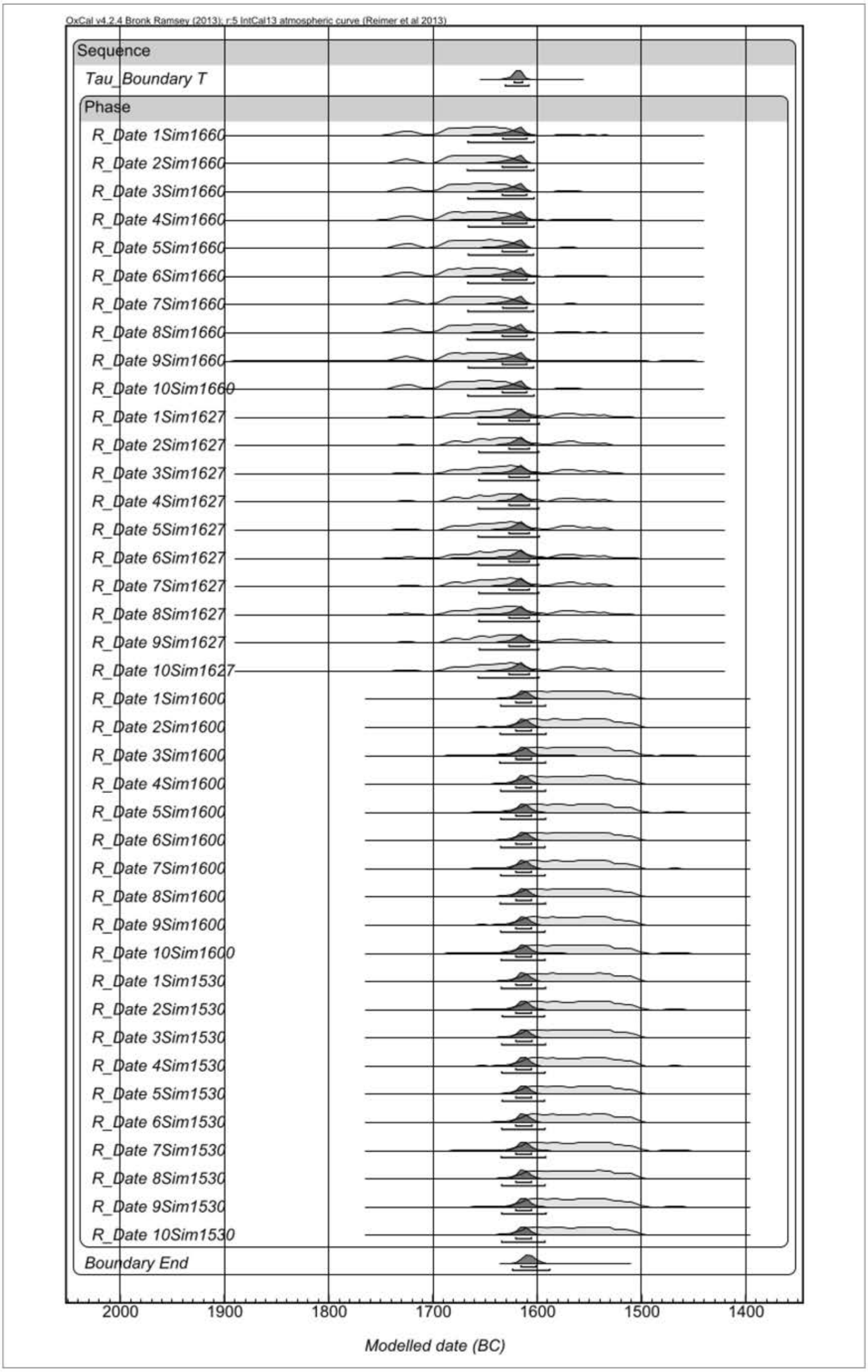

Figure 3. Uncombined (light grey) vs. Tau_boundaries (dark grey) results of the simulated dataset of ten dates of four known ages (1660, 1620, 1600 and 1530 BC, respectively 3367±30, 3340±30, 3290±30 and 3287 ±30 BP). Although the program is designed to take into account the possibility that some of the dates may be significantly earlier than the final event, the final ('Boundary End') result is in any case shifted towards a misleadingly earlier date.

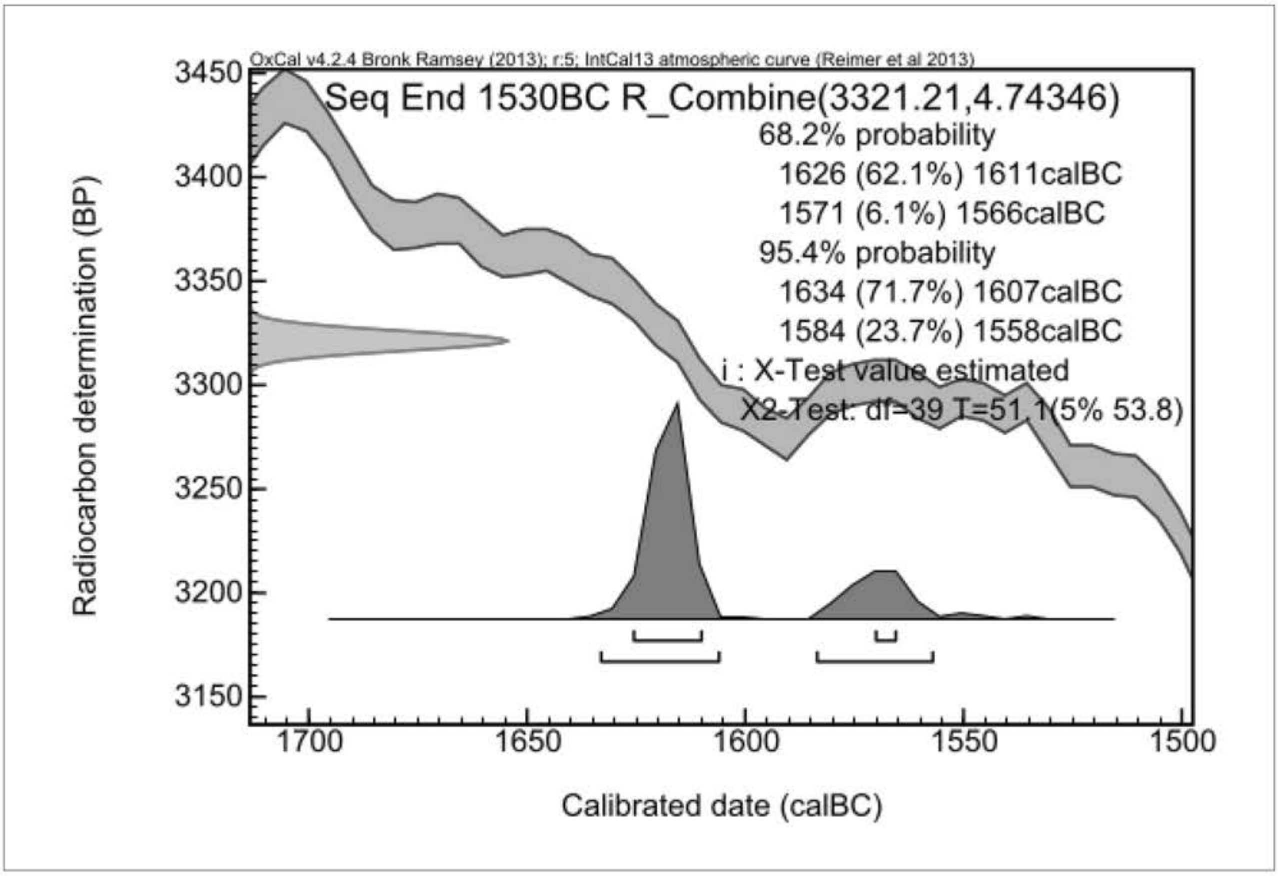

Figure 4. R_Combine results of the simulated dataset of 10 dates of four known ages (1660, 1620, 1600 and 1530 BC, respectively 3367±30, 3340±30, 3290±30 and 3279±30 BP). The results pass the Chi-square test but the 'known' real end age (1530 BC) is not even within the 95,4% confidence calibration interval.