[Radiocarbon, Vol 25, No. 2, 1983 P 493-500]

RADIOCARBON DATING ARCHAEOLOGIC AND ENVIRONMENTAL SAMPLES CONTAINING 10 TO 120 MILLIGRAMS OF CARBON

JOHN C SHEPPARD, J FRED HOPPER, and YVONNE WELTER

Department of Chemical Engineering
Washington State University
Pullman, Washington 99164

A conventional ^{14}C system with either a $15cm^3$ or a $100cm^3$ methane gas proportional counting tube, each constructed from high purity copper, has been used at the Washington State University ^{14}C laboratory for three years (Sheppard, Hopper, Westberg, 1981). The electronic components of this system included highly stable (John Fluke) power supplies, modified pre-amplifiers (Canberra), and NIM amplifiers, single channel analyzers, scalers, timers, etc, (ORTEC). Modules were selected for a highly stable low-noise system. The system was designed to minimize problems generated by ground-loops, electromagnetic noise pick-up, and line noise. It operates in a copper-lined basement room of a four-story brick and concrete building along with a second and older gas proportional system which has three $500cm^3$ copper counting tubes operated at 2 or 3 atmospheres of methane. Analysis of the older system's background data indicates that the background counting rate of 0.8 counts per minute is virtually independent of atmospheric pressure. It has not been possible to determine the pressure dependence of the small counting tubes.

During the past three years ca 40 samples were dated using the 15 and $100cm^3$ counting tubes and ca 10 samples, first prepared for the $500cm^3$ tubes, were counted in the larger and smaller counting tubes to calibrate the small tubes and to learn more about non-Poissonian errors believed to increase the actual standard deviation of a ^{14}C age measurement.

This analysis started with the assumption that the actual error associated with a ^{14}C age measurement is larger than the Poissonian error derived from the ^{14}C decay law and related counting statistics. Our experience with the updating of a marginal, lightly shielded system with noisy electronics, strongly suggested that non-statistical sources or error exist even for well-managed systems. Similarly, Currie (1973) found that such increased errors are likely, while Clark and Renfrew (1973) and Scott, Baxter, and Aitchison (1981) observed that the actual ^{14}C age error is likely to be twice the Poissonian

value. The following observations were developed from an analysis of paired age data and series of partial counts for the 15, 100, and 500 cm^3 counting tubes.

COMPARISONS OF THE SMALL AND LARGE COUNTING TUBE DATA

Ten pairs of ^{14}C age measurements for the 100cm^3 counting tube age were compared to those obtained with large 500cm^3 counting tubes. Application of the F test to these data yielded an F value of 0.02. Since 0.02 is much less than the 1% critical value of F(0.01, 1, 18) = 8.28, the null hypothesis could not be rejected, indicating that the mean ^{14}C ages determined by the 100 and 500cm^3 tubes were not significantly different. Use of the F test for ten pairs of ^{14}C age measurements for the 15 and 500cm^3 counting tubes yielded the same conclusion--the ^{14}C ages were not significantly different.

The two sets of paired ^{14}C age determinations were then subjected to regression analysis, with and without weighting. Comparison of the 15 and 100cm^3 age measurements with those for the 500cm^3 counting tubes yielded correlation coefficients >0.96 and 0.90 for the 100 and 15cm^3 counting tubes, respectively. Both analyses indicated potential biases in the small tube measurements because non-zero intercepts of -260 and 600 years for the 100 and 15cm^3 counting tubes, respectively, but the intercept errors of 270 and 440 years do not strongly support this hypothesis. Paired t tests also support these conclusions. ^{14}C ages for the 15 and 500cm^3 tubes were recalculated using running averages for the backgrounds and oxalate counting rates, as well as those just before and after the sample count. There was a slight improvement with the before-after approach which we believe is preferable because small samples are counted much less frequently for the large system. Further analysis of the oxalate counting data for the 100cm^3 tube suggests a non-Poissonian component of 1.5, slightly lower than the factor of two observed by Scott, Baxter, and Aitchison (1981).

To summarize, this analysis suggests that the small counting tube age determinations are in statistical agreement with the large tube data. However, there is some evidence for a bias and variability that require further study. During the comparison occasional cable noise was observed, several NIM modules (ORTEC, mainly) malfunctioned, and some problems with gas purity were encountered. Just how these episodes influenced the small tube data is not clear, but they were probably adversely affected, at least slightly.

FURTHER ANALYSIS OF THE ^{14}C COUNTING DATA FOR NON-POISSONIAN BEHAVIOR

It has been evident from the beginning of the operation of the WSU system that various interferences adversely influence results. For example, the first counting tube had small Teflon gaskets that outgassed a volatile impurity that completely quenched impulses. Noise generated by a teletype of an adjacent ^{14}C system was eliminated by replacement of an ORTEC pre-amplifier with one built by Canberra. It is not clear why ORTEC pre-amplifiers behaved this way. Later, a more subtle problem emerged when chi-square values began to increase. A search revealed that ORTEC scalers were dropping digits as the data were transferred from scalers to the line printer. Thus, some of our dates were probably biased, but by how much is unclear. We found that cables occasionally generate significant noise and must be replaced. All sources of error were corrected but, more important, they do occur at various times. When the background is ca 1 count every 6 minutes and sample count rates of 1 count every 50 minutes are being measured, it is clear that 1 count every 100 minutes is very undesirable. Most of the time the spurious count rate is 1 per 1000 minutes.

In addition to these readily corrected errors, slight variations of the high voltage and discrimination level may cause fluctuations in the observed counting rate; further, gas impurities may cause subtle changes in counting efficiency. These perturbations are generally classified as non-Poisson errors (sometimes referred to as dead-time distorted Poisson processes). Excellent discussions of such non-Poissonian behavior were given by Müller (1973; 1974), Currie (1971), Hooton and Parsons (1973), Hilaire (1973), and Pazdur (1976). In various ways, all have attempted to extract the non-Poissonian error component from the observed counting data.

In our analysis of counting data for the WSU ^{14}C systems we used the approaches of Hooton and Parsons (1973), Hilaire (1976), and Pazdur (1976). Most days, every ^{14}C laboratory generates a series of partial counts, usually 100 minutes long, that last from 1 to 7 days. Thus, anywhere from 14 to 100 partial counts are available for analysis. Hooton and Parsons (1973) and Hilaire (1973) employed the mean-square successive difference (MSSD) method, first developed by Hart (1942) and expanded by Bennett and Franklin (1954), to analyze counting data for indications of instrumental drift.

https://doi.org/10.1017/S0033822200005786 Published online by Cambridge University Press

We used the Hooton-Parsons method. The MSSD statistic ρ^2 is calculated as the mean of the squares of n-1 successive differences between n observations:

$$\rho^2 = \Sigma(x_{i+1} - x_i)^2/(n - 1) \quad i + 1, 2, -- n - 1. \tag{1}$$

Further, the average value of

$$\rho^2 \approx 2S^2 \tag{2}$$

where

$$S^2 = \frac{\Sigma(x_i - \bar{x})^2}{n-1} = \text{the sample variance of a series of particle counts.} \tag{3}$$

Also, the a value of ρ^2 is less affected by a gradual drift than s^2 but is more influenced than s^2 by rapid oscillations. Thus, the ratio

$$\rho^2/s^2 = \eta \tag{4}$$

is a measure of departures from a normal distribution. Drifting counting systems, eg, will have η's significantly less than 2, while η's significantly greater than 2 characterize oscillating systems.

Another way of looking at the MSSD method is that it is a time-dependent chi-square test, where

$$x^2 = \frac{n - 1}{\eta} \cdot \frac{\rho^2}{\bar{x}} \approx \frac{n - 1}{2} \cdot \frac{\rho^2}{\bar{x}} \tag{5}$$

and $\bar{x}$ = the mean counting rate for n partial counts. It is important to note here that n refers to the number of counting intervals of 100 minutes.

Many MSSD analyses were done for the 15, 100, and 500cm^3 gas proportional counting tubes and results are summarized in table 1.

TABLE 1. MSSD of counting data for the WSU radiocarbon systems

Counting tube	Volume	Electronic system	Number of counts	η(ave)
2	500	Old	19	1.93 ± 0.43
3	500	Old	14	1.97 ± 0.59
T4	100	New	15	1.92 ± 0.30
T2	15	New	15	1.91 ± 0.33

https://doi.org/10.1017/S0033822200005786 Published online by Cambridge University Press

Consider the data for the new system and small counting tubes. Since limiting values of η for n = 15 and a 95% confidence level are $1.21 < \eta < 2.79$, it can be assumed that the counting data are randomly distributed. None of the T2 or T4 data fell outside these bounds. Two of the MSSD analyses for counting tube 3 showed η's less than the lower bound for 14 counts. Subsequently, it was found that counting tube 3 had a leak. After a repair, the performance of counting tube 3 returned to normal. To summarize, application of the MSSD method to the WSU counting tubes indicated normal stable performance with the exceptions of two abnormal counts.

EXAMINATION OF THE LONG-TERM PERFORMANCE OF THE WSU COUNTING TUBES AND ELECTRONICS

Pazdur (1976) showed that a dead-time distorted distribution obeys the "negative binomial distribution" as developed by Greenwood and Yule (1920). Furthermore, Feller (1966) showed that such distributions belong to the class of compound Poisson distributions, which in the limit reduce to the ideal Poisson distribution. Pazdur (1976) also showed, for a series of counts with a partial counting time T, that the variance for a non-Poisson situation is

$$\sigma^2 = RT(1 + RT\delta^2) \tag{6}$$

where R is the average counting rate and δ^2 = rms fluctuation of the detection efficiency. To further clarify δ^2, let $R = R_o\varepsilon$, where $\varepsilon = \varepsilon(\nu_B, \nu_D)$. R_o is the true ^{14}C disintegration rate and ε is a detection efficiency with a variable component ν_B due to fluctuations of bias voltage and another component ν_D which varies with the discriminator level. Of course, this is grossly over-simplified but it emphasizes that such variations of counting efficiency do occur.

To estimate these fluctuations, let $V = \sigma^2/\bar{m}$, remembering that $\bar{m} = RT$. With substitution and rearrangement equation 6 becomes

$$V = 1 + RT\delta^2 \tag{7}$$

As Pazdur (1976) showed that values of V can be estimated from the ratio of the sample variance s^2 to the population or Poisson variance S_p^2, then

$$\hat{V} = S_s^2/S_p^2 \tag{8}$$

https://doi.org/10.1017/S0033822200005786 Published online by Cambridge University Press

For n partial counts of time T

$$\hat{V} = \frac{S_s^2 T}{R} \tag{9}$$

Thus, values of $\hat{V}$ should be calculated for as many long counts as available to obtain an average, $\bar{V}$. Since S_s^2 and S_p^2 are both estimators of the true variance of the ^{14}C counting rate, values of $\hat{V}$ should scatter around unity. It can be argued that the larger of S_s^2 and S_p^2 should be used in ^{14}C age error calculations to obtain a conservative estimate. Table 2 contains an analysis of the two small counting tubes and two large (500cm^3) tubes of the WSU systems.

TABLE 2. Pazdur analysis of WSU counting tubes

Detector/ volume(cm^3)	Pressure atm	Number of counts*	Average cpm,R	$\bar{V}$	δ^2
T2/15	3.0	41	0.27	1.01 ± 0.24	4.8×10^{-4}
T4/100	3.0	20	1.21	1.06 ± 0.17	5.3×10^{-4}
2/500	2.0	45	3.74	1.03 ± 0.45	7.0×10^{-5}
3/500	3.0	45	4.77	1.04 ± 0.44	1.0×10^{-4}

*The partial counting time for all these counts was 100 minutes with the number of partial counts ranging from 10 to 14 for the large counting tubes to 10 to 50 for the smaller tubes.

The data in table 2 reveal that $\bar{V}$ does fluctuate about unity as predicted. More important, values of δ^2 for all counting tubes are rather small, which means that variations of detector-efficiency-related fluctuations in bias and discriminator levels and other sources of random instrumental error are negligible for the "small" and "large" systems.

CONCLUSIONS

Analyses applied to the small and large systems at the WSU ^{14}C laboratory suggest that electronic components of a well-managed system are minor sources of non-Poisson error. MSSD analysis indicated that short-term drift and oscillation of the order of day(s) was not significant. Similarly, application of the Pazdur (1976) test, which covers weeks or months of instrument operation, showed that instrumental fluctuations were negligible. While these tests provide bases for the elimination of instrument-based errors, they do

https://doi.org/10.1017/S0033822200005786 Published online by Cambridge University Press

not eliminate potential intra-latorabory or systematic errors that include pulse quenching, isotopic fractionation, use or non-use of running averages for backgrounds and oxalate standards, and other unnamed sources of error.

It seems unreasonable to assume that all the ^{14}C age errors of a small tube system, or even a large one, are due solely to counting errors. Thus, we are left with extra-instrumental error sources as mentioned above. The recent data of Scott, Baxter, an Aitchison (1981) strongly support the existence of such errors for large and small systems.

If ^{14}C age errors are actually larger than those obtained by Poisson statistics, perhaps this problem should be recognized. Müller (1979) suggested that it is legitimate to multiply an error by a factor of two or three, provided this arbitrary augmentation is applied at the very end of the age calculation and not at some intermediate stage. If the Scott, Baxter, and Aitchison (1981) data are used as a criterion, this factor should be two. However, before such an augmentation is made, there must be universal agreement among ^{14}C laboratories on 1) whether to do it at all and 2) the actual magnitude. The latter requires considerable study and research.

REFERENCES

Bennett, CA and Franklin, NL, 1954, Statistical analysis in chemistry and the chemical industry; New York, John Wiley & Sons, p 677-684.

Clark, RM and Renfrew, C, 1973, The tree-ring calibration of radiocarbon and the chronology of ancient Egypt: Nature, v 243, p 266-270.

Currie, LA, 1971, On the interpretation of errors in counting experiments: Anal Letters, v 4, p 777-784.

---------- 1973, The evaluation of radiocarbon measurements and inherent statistical limitations in age resolution, in Rafter, TA and Grant-Taylor, T, eds, Internatl conf on radiocarbon dating, 8th Proc: Wellington, Royal Soc New Zealand, p 598-611.

Feller, W, 1966, An introduction to probability theory and its applications, v II: New York, John Wiley & Sons, p 56-57.

Greenwood, M and Yule, GU, 1920, An inquiry into the nature of frequency distributions representative of multiple happenings: Jour Royal Stat Soc, v 83, p 255-279.

Hart, BI, 1942, Significance levels of the ratio of the mean square successive difference to the variance: Anal Math Statistics, v 12, p 1-13.

Hilaire, M, 1973, Treatment of errors in low-activity measurements: Nuclear Instruments Methods, v 112, p 385-390.

Hooton, KAH and Parsons, ML, 1973, Equipment stability in x-ray fluorescence spectrometry and radioactive counting--a statistical approach: Anal Chemistry, v 45, p 2218-2227.

Müller, JW, 1973, Dead-time problems: Nuclear Instruments Methods: v 112, p 47-57.

---------- 1974, Some formulae for a dead-time-distorted Poisson process: Nuclear Instruments Methods: v 117, p 401-404.

---------- 1979, Some second thoughts on error statement: Nuclear Instruments Methods: v 163, p 241-251.

Pazdur, MF, 1976, Counting statistics in low-level radioactivity measurements with fluctuating counting efficiency: Internatl Jour Appl Isotopes, v 20, p 179-184.

Scott, EM, Baxter, MS, and Aitchison, TC, 1931, An Assessment of variability in radiocarbon dating, in Methods of low-level counting and spectrometry: IAEA, Vienna, p 371-391.

Sheppard, JC, Hopper, JF, and Westberg, HH, 1981, Radiocarbon dating of milligram-sized archaeological and environmental samples, in Methods of low-level counting and spectrometry: IAEA, Vienna, p 409-416.