ETNOLOGISKA STUDIER

32

A Medicine-man's Implements and Plants in a Tiahuanacoid Tomb in Highland Bolivia

BY

S. HENRY WASSÉN

CONTRIBUTORS

Wolmar E. Bondeson, Helge Hjalmarsson, Carl-Herman Hjortsjö, Bo Holmstedt, Eskil Hultin, Jan-Erik Lindgren, Thomas Liljemark, and Richard Evans Schultes

GÖTEBORGS ETNOGRAFISKA MUSEUM

A Medicine-man's Implements and Plants in a Tiahuanacoid Tomb in Highland Bolivia

BY

S. HENRY WASSÉN

CONTRIBUTORS

Wolmar E. Bondeson, Helge Hjalmarsson, Carl-Herman Hjortsjö, Bo Holmstedt, Eskil Hultin, Jan-Erik Lindgren, Thomas Liljemark, and Richard Evans Schultes

GÖTEBORG

1972

THE ACCURACY OF THE RADIOCARBON DATING

ESKIL HULTIN

Institute of Biochemistry, University of Stockholm, Stockholm, Sweden

Systematic and random errors in radiocarbon dating have decreased substantially during the last decade as a consequence of improved purification of samples from contaminations, advanced apparatus technology, establishment of a long tree-ring chronology, and investigations on differential carbon isotope assimilation. The achievements of the various radiocarbon dating laboratories have been impressive.

Several international conferences on radiocarbon dating have been of importance not only for the exchange of stimulating ideas but also for the presentation and discussion of some archaeologists' observations of discrepancies between the radiocarbon age and the age established otherwise regarding many objects of vegetable or animal origin.[1-3]

TIME VARIATIONS IN ^{14}C PRODUCTION OVER THE MILLENNIA

One example of such controversies which have been largely reconciled concerns the age of various samples from ancient Egypt, for which the age determined from radiocarbon measurements could appear almost a thousand years less than the age determined from historical counting. The differences between the historical age and the age calculated from radiocarbon determinations were largely explained when measurements of the radioactivity remaining in tree-ring dated old wood[4] indicated that the concentration of the radiocarbon isotope ^{14}C in the air had not been constant during the last seven thousand years for which measurements are available. It was tacitly assumed earlier that the formation rate of ^{14}C was stable.[5,6]

Consequently, if we want to assay the true age of an object, it is not sufficient to measure the amount of radioactivity remaining now, and to calculate the apparent radiocarbon age, assuming that the same relative amount of the radioactive isotope ^{14}C is always originally taken up. Rather, we need in addition some kind of estimate of the relative amount of radiocarbon present in the atmosphere when the plant or animal in question lived.

DIFFERENCES BETWEEN LABORATORIES

An example of another kind of discrepancy between age, estimated from radiocarbon determination, and true age, was brought up by Rowe who found two different time scales for Peruvian archaeological objects.[7] According to one time scale, pottery was introduced on the Peruvian coast somewhere between 1200 and 1450 B.C., but according to the other time scale this happened earlier than 1900 B.C. The standard deviation for a determination was about 100 years, but the difference between the two series of age determinations was about 700 years. Thus, the two time scales appeared to be irreconcilable. However, the differences can now be understood, as will be explained in the following introductory example.

Three different laboratories have published results from the determination of radiocarbon age of tree-ring dated wood, covering also the period mentioned.[8–11] Calibration curves similar to those in Figure 1 can be calculated from these data. In the time range mentioned for the introduction of pottery on the Peruvian coast, the three calibration curves are about 140 years apart (Appendix 1).

Under such circumstances, if samples which are not known to be of the same age are sent to two laboratories, and the answers are 700 years apart, one would not be justified to conclude that the samples are of significantly different age, although one may guess that a sufficiently high level of significance for such a conclusion may be reached if some additional age determinations are done (Appendix 2). However, if only a few samples were available, and if the accuracy within laboratories (*i.e.*, disregarding the differences between laboratories, times, etc.) were only slightly less than it is now in most instances, the results would probably not differ from the 10% significance point, and one would not feel encouraged from a statistical point of view to suspect that the samples were of different age.

Consequently, if the calibration curves mentioned are applied to Rowe's two quotations of the time for the introduction of pottery on the Peruvian

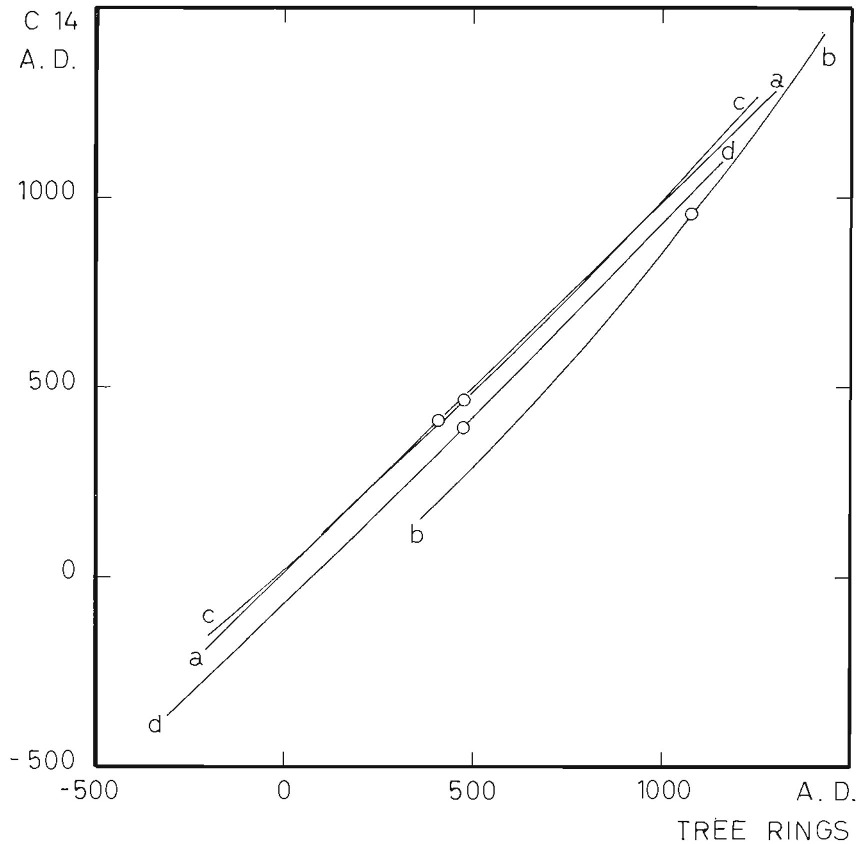

Fig. 1. Standard curves for the determination of true age from radiocarbon age,[40] calculated from analyses of tree-ring dated wood at four laboratories: a in Tucson,[10] b in Groeningen,[41] c in Philadelphia,[11] and d in San Diego[8] (Appendix 5). The publications used for the calculations include the year "0",[4a, 8b, 16a] and this imaginary year is therefore included in this Figure and the corresponding calculations, however, under protest that chronologers admit no year zero between 1 B.C. and A.D. 1.[42] Drawing by the author.

coast, 1200–1450 B.C., and before 1900 B.C., one would conclude that the introduction occurred about 2100 B.C.; the 95% confidence interval would be about ± 500 years.

Radiocarbon age is calculated from the number of disintegrations of the carbon isotope ^{14}C counted during a time interval of suitable length. The standard deviation of the number of counts equals the square root of the number of disintegrations actually counted, and this square root is used for the calculation of the standard deviation of the radiocarbon age determination.

However, it is well-known from official analytical laboratories that collaborative investigations on accuracy of analyses may reveal significant influences from such sources of variation as differences between seemingly equivalent procedures, differences between occasions (different year or month or week when the analyses are done), differences between analysts, and other differences between laboratories.

It is obvious that the information about standard deviation supplied by laboratories for radiocarbon dating represents a minimum, valid only for ideal conditions which the laboratories seek to attain but which no laboratory can reach entirely. The incompleteness of the information supplied about the accuracy of a radiocarbon age determination can be misleading if it is not properly understood and can then cause confusion and distrust.

So seldom are two measurements of the same sample published that one has the impression that most laboratories are not aware of the possibility of calculating the overall random accuracy within a laboratory from pairs of analyses.[12a, 12b] Such measurements must be done independently on two samples of the object whose age shall be determined; mere use of two counters is not sufficient. One collaborative work was actually done as early as in 1960 by Willis, Tauber, and Münnich,[13] however, without the statistical planning and evaluation which is a matter of routine between official analytical laboratories.

It may be mentioned for comparison that even such seemingly simple analyses as moisture determinations in an industrial process have appeared so doubtful that a statistically planned investigation was needed, as has been described in a manual of laboratory procedures.[43] How much more are statistical methods needed for evaluation of the actual accuracy of radiocarbon age determinations!

The short calibration curves just mentioned, and those presented in Figure 1, are not the first ones which demonstrate that results from various laboratories differ. Neustupný,[14] in an article on the accuracy of radiocarbon dating, read at the Twelfth Nobel Symposium in 1969, discussed several such observations and expressly pointed out that "the measurements of the tree-rings performed in Tucson deviate systematically from those carried out in San Diego and Pennsylvania".

Earlier calibration curves have been published, by Suess[8, 9] and by Stuiver and Suess.[15] These calibration curves have many wriggles with a time period about 100 years. At the Twelfth Nobel Symposium Suess presented extensive calibration curves.[16,16a] The ensuing discussion is of great interest. Suess stated that "in addition to the statistical error there is probably also a systematic error that may be different for different laboratories". Damon[16b] mentioned that he had plotted both Suess' results and his own results, and that he had found a difference of 180 years between the results from the two laboratories (Appendix 3).[16c] Ralph,[16b] from a corresponding plotting, had also found that a difference exists between the results of Suess' laboratory and his, and Ralph also distrusted the way in which Suess had drawn wriggles in his calibration curve, and mentioned two time periods, about A.D. 1200 and about 4200 B.C., in which the results from his laboratory were contradictory to Suess' results as to the direction of the wriggles.

Vogel expressly refuted Suess' calibration curve, pointing out that "much less than one-third of the points lie further than 1 σ away from the line, which implies that some of the small wriggles—I am not talking about the large ones—are probably not real". (Appendix 4)[16c]

Analyst changes which Suess mentioned as occurring in his laboratory, *i.e.*, "one has technicians and undergraduate students doing the routine preparations"[16b] are a possible source of variation of the kind rather often considered in statistical investigations (variation between analysts and between occasions). However, within Suess' laboratory, the standard deviation of the measurements from a second degree polynomial within the intervals of interest for this article (Appendix 1 and Figure 1) clearly does not give any unfavorable comparison with the work in other laboratories.

Nevertheless, although Suess stated that "it is extremely difficult to recognize these wriggles with certainty,[16b] he drew his calibration curve by "cosmic schwung"[16c] and defended the result as if he were not awere of the great weight of Vogel's critisism: the standard deviation cannot be less than the minimum standard deviation calculated from the radioactivity counting only.

It may be added here that although there is no statistically acceptable proof at present for the existence of the rapid wrigglings with a period of about a century, a main question is whether or not sufficiently sudden changes occur in the production of the isotope ^{14}C, and especially, if an excess or a deficiency disappears sufficiently rapidly as soon as the production of carbon-14 returns to a value which is normal, considering the changes over the millennia. High-altitude cosmic-ray neutron intensity variations and production of radiocarbon have been studied by Soberman.[17] Alleged short-time variations in carbon-14 production have been claimed to show some degree of negative correlation to the variation in sun spot numbers.[18,19] It appears from recent investigations that exchange times for carbon dioxide are sufficiently short for the appearance of wriggles with a period of about a century.[20,21] Consequently, the hypothesis about short-time variations, originally suggested by de Vries[22] may become a challenging project for a statistically planned collaboration between the radiocarbon dating laboratories.

However, the recent publication of Suess' calibration curve in the prestigious journal Scientific American by Renfrew[23] together with the erroneous statement that the carbon-14 determinations at University of Arizona, University of Pennsylvania, and University of California at San Diego in general agree fairly well with one another, makes it necessary to urge those who intend to submit samples for radiocarbon dating to see their statistician.

THE CARBON DIOXIDE CYCLE

Two causes for deviations between the radiocarbon age and the true age of an object have so far been discussed: long range variations in the production of the isotope carbon-14, and variations between the laboratories. These variations can be remedied with properly interspaced calibration experiments for which tree-ring dated wood fortunately is available.[4,16b]

Still another possible complication may be discussed briefly, although the magnitude of its effect is not known very well. As early as in 1952, Wickman found from carbon isotope analyses of 105 herbarium specimens that characteristic differences occur between the relative abundances of carbon isotopes in plants grown in different biotopes.[24] He concluded that these differences are related to the varying intensity of the local carbon dioxide cycle. A similar investigation of 104 species was published in 1971 by Smith and Epstein.[25] They discuss photosyntetic fractionation mechanisms. Neither they, nor Wickman, found any relationship between phylogeny and the specific isotope fractionation. Isotope fractionation studies have also been done on marine animals.[26] Wickman's hypothesis seems supported by investigations on variations in carbon dioxide concentration and carbon isotope relationships in various types of localities, *e.g.*, dense forests and open fields,[27–30] and the hypothesis also seems supported by an investigation of the amount of carbon dioxide generated by the soil.[31]

Consequently, one would hope that research on the influence of the biotope on carbon-13 relationships in plants, and research on carbon-14 concentrations in soil and in carbon dioxide released from soil will increase our information about the influence of the biotope and the local carbon dioxide cycle on the apparent radiocarbon age. This may enable the calculation of correction terms, both for species, and for biotopes.

A few words need to be said about the accuracy of the assay of the stable isotopes also. An inspection of plots of the specific isotope fractionation $\delta^{13}C$ in Sequoia wood[8] and in bristlecone pine wood[9] versus age indicates that even if a few outlying results could be eliminated with statistical methods,[12c,32,33] the overall standard deviation is much larger than the accuracy claimed for the measurements. Also in this case, fully duplicate, randomized analyses would have given additional information. Furthermore, the inspection leads to a guess that further research may show significant differences in the specific isotope fractionation for bristlecone pine wood and for Sequoia wood, at least if one uses tree-rings of Sequoia wood which have grown before the tree had reached a height sufficiently above the main carbon dioxide recycling layer.

AGE OF THE BOLIVIAN CAVE BURIAL FIND

A piece of a skull, two samples of *Ilex Guayusa* leaves, and one sample of ground *Ilex Guayusa* leaves, the latter extracted with methanol for the purpose of chemical analysis, and dried, were sent to *Laboratoriet för radioaktiv datering* in Stockholm. The samples were processed according to current methods.[34–37] As is customary in radiocarbon dating, the accuracy of the results was based on the radioactivity counts; however, standard deviations less than 100 years were rounded off to 100. The results from the age determinations were as follows (radiocarbon age and standard deviation): the skull A.D. 755 ± 100, the two leaf samples A.D. 375 ± 100, and 355 ± 200, and the sample of ground, extracted leaves A.D. 1120 ± 100.

The apparatus were recently rebuilt, and a calibration curve is not yet available, and the number of duplicate determinations now available is not sufficient for accuracy calculations. Consequently, calibration curves calculated from results published from other laboratories have to be used, and differences between laboratories become an additional source of variation. The calibration curves for the time interval of interest were calculated with an electronic desk-top computer which uses recorded programs, by fitting a second degree polynomial.[38] In some cases the radiocarbon age had to be calculated according to an equation published elsewhere[8a] from other data supplied. The calibration curves are shown in Figure 1.

The standard deviation between laboratories for A.D. 500 was calculated from the corresponding regression line points. Values for Student's t distribution to be used with combined standard deviations with different degrees of freedom were calculated according to Cochran and Cox.[39]

In order to get the time to which the skull dates back, we add the correction term obtained from Figure 1, 77 years, to the result from the radiocarbon determination. A combined standard deviation, 141 years, was obtained as the square root of the sum of squares of the standard deviations for variations within and for variations between laboratories. The size of the confidence interval is the product of the standard deviation and the t value. The 95% confidence interval arrived at for the skull is from about A.D. 470 to about A.D. 1190; similarly the 50% confidence interval is from about A.D. 710 to about A.D. 950.

The leaf sample which was ground and extracted with methanol deviates in radiocarbon age significantly from the other leaf samples. It was found in a bag strikingly similar to another bag from which one of the other leaf samples was taken. At present this discrepancy cannot be explained, and the result from this sample is regarded as outlying.

The time to which the other *Ilex Guayusa* leaves date back was obtained by adding the correction term to the weighted mean value of the result from the radiocarbon age determination. The standard deviation for this mean value, 80 years, was combined with the standard deviation for differences between laboratories, and the result, 134 years, used for the calculation of confidence intervals. The 95% confidence interval arrived at for the leaves is from about A.D. 100 to about A.D. 800, and the 50% confidence interval arrived at is from about A.D. 330 to about A.D. 560.

When the laboratory has got its calibration curve ready for the time period of interest here, the confidence intervals can be recalculated and are expected to be considerably less broad than those which can be calculated now. Such calculations are described elsewhere.[40]

Finally, the question remains whether or not the skull and the leaves are contemporary. Because the radiocarbon age of both kinds of objects was determined by the same laboratory, we use the standard deviations stated by the laboratory for the comparison. The age difference is $755-371=384$ years. The standard deviation for the difference[12d] is $(100^2+80^2)^{0.5}=128$ years; $t=3.0$; $P=0.002$, *i.e.*, the probability that the skull and the leaves are contemporary is about 0.2%. However, if the leaves were collected in a forest with a high carbon dioxide recycling, the leaves may date back to a time later than the one calculated here, and if so, the difference in age would be less. With the information at present available, the 5% confidence interval for the age difference is about ±250 years.

APPENDIX

Appendix 1. The calculation of regression line points from published radiocarbon age measurements: range of time, number of measurements included, mean value of the age (A.D.) of tree-ring dated wood, the corresponding estimate of radiocarbon age, variance ratio F for curvilinearity (an asterisk indicates significance at the 5% level), estimate of radiocarbon age for wood from 2100 B.C., standard deviation from quadratic regression.

Philadelphia. 2985 B.C. to 1350 B.C.; $n=41$; $\bar{x}=-2080$; $f(\bar{x})=-1672$; $F=0.5$; $f(-2100)=-1686$; $s_{y \cdot x}=115$.

San Diego. 2500 B.C. to 900 B.C.; $n=42$; $\bar{x}=-1972$; $f(\bar{x})=-1713$; $F=5.8^*$; $f(-2100)=-1912$; $s_{y \cdot x}=70$.

Tucson. 3165 B.C. to 1349 B.C.; $n=20$; $\bar{x}=-2015$; $f(\bar{x})=-1471$; $F=2.9$; $f(-2100)=-1533$; $s_{y \cdot x}=82$.

Appendix 2. Simplified calculation: $d.f.=2$; $s=140$; $s_{\bar{x}_1-\bar{x}_2}=198$; $t=700/198=$ $=3.5$; $0.05<P<0.10$.

Appendix 3. It appears that it did not occur to the discussers that this should be considered a comparison of distance between regression lines, here simplified to distance between mean values. Thus, if the standard deviation for a measurement is 100 years, the standard deviation for the difference between mean values obtained from series of 80 and 240 measurements[16c] is $100 \cdot \sqrt{(80+240)/(80\times240)}=13$; $t=14$, and the difference 180 years is indeed highly significant. Thus, it is not sufficient, as Vogel has done,[16c] to smooth Suess' line for use as a correction curve in archaeology.

Appendix 4. A calculation of the standard deviation of Suess' experimental points from his calibration curve between A.D. 310 B.C. and A.D. 1160 as measured from his graph[16a] gave 34 years. A comparison with the standard deviation for the measurements, shown in the graph, 38 years, gave $F=38^2/34^2=1.3$, *i.e.*, no significant difference. However, Suess' calibration curve is, by his "cosmic schwung",[16c] drawn at such a distance from the measurement point of wood which grew about A.D. 43 that this has a substantial effect on the standard deviation; and the situation is similar for wood which grew about A.D. 867. If these two points are disregarded because the calibration curve appears to be drawn fancifully here, the standard deviation of the points from the curve corresponds to 29 years, and $F=38^2/29^2=1.7$ $(P=0.01)$. This indicates significantly that the calibration curve is not correct.

Appendix 5. For explanation, se Appendix 1. The quadratic regression equation is also given; the point for A.D. 500 is calculated.

a. 206 B.C. to A.D. 1252; $\hat{Y}=11.7+0.9798X-0.0000035X^2$; $n=11$; $\bar{x}=410$; $f(\bar{x})=413$; $F=0.0$; $f(500)=501$; $s_{y\cdot x}=97$.

b. A.D. 359 to A.D. 1436; $\hat{Y}=-166.9+0.7920X+0.000227X^2$; $n=8$; $\bar{x}=1081$; $f(\bar{x})=954$; $F=2.4$; $f(500)=286$; $s_{y\cdot x}=41$.

c. 200 B.C. to A.D. 1200; $\hat{Y}=23.2+0.8907X+0.000082X^2$; $n=39$; $\bar{x}=473$; $f(\bar{x})=463$; $F=2.8$; $f(500)=489$; $s_{y\cdot x}=48$.

d. 310 B.C. to A.D. 1160. $\hat{Y}=-72.4+0.9593X+0.000040X^2$; $n=44$; $\bar{x}=478$; $f(\bar{x})=395$; $F=0.7$; $f(500)=417$; $s_{y\cdot x}=52$.

REFERENCES

1. Proceedings of the Sixth International Conference, Radiocarbon and Tritium Dating, held at Washington State University, Pullman, Washington. Peek Publications, Palo Alto, Calif. 1967.
2. International Atomic Anergy Agency. Proceedings Series. Radioactive Dating and Methods of Low-level Counting. Proceedings of a Symposium Organized by the International Atomic Energy Agency in Co-operation with the Joint Commission on Applied Radioactivity (ICSU) and held in Monaco, 2–10 March 1967. Vienna 1967.
3. OLSSON, I. U. (Ed.) Radiocarbon Variations and Absolute Chronology. Proceedings of the Twelfth Nobel Symposium held at the Institute of Physics at Uppsala University. Almqvist & Wiksell, Stockholm 1970.
4. FERGUSON, C. W. In Olsson, I. U. (Ed.) *l.c.*, p. 237. 4a. p. 239.
5. SUESS, H. E. Z. Physik 202 (1967) 1.
6. SÄVE-SÖDERBERGH, T. and OLSSON, I. U. In Olsson, I. U. (Ed.) *l.c.*, p. 35.
7. ROWE, J. H. In Radiocarbon and Tritium Dating, *l.c.*, p. 187.
8. SUESS, H. E. J. Geophys. Res. 70 (1965) 5937. 8a. p. 5940. 8b. p. 5943.
9. SUESS, H. E. In Radioactive Dating and Methods of Low-level counting, *l.c.*, p. 143.
10. DAMON, P. E., LONG, A. and GREY, D. C. In Olsson, I. U. (Ed.) *l.c.*, p. 615.
11. RALPH, E. K. and MICHAEL, H. N. In Olsson, I. U. (Ed.) *l.c.*, 619.
12. SNEDECOR, G. W. and COCHRAN, W. G. Statistical Methods. Sixth Edition. Ames, Iowa 1967. 12a. p. 91. 12b. p. 258. 12c. p. 157. 12d. p. 101.
13. WILLIS, E. H., TAUBER, H. and MÜNNICH, K. O. Amer. J. Sci. Radiocarbon Suppl. 2 (1960) 1.
14. NEUSTUPNÝ, E. In Olsson, I. U. (Ed.) *l.c.*, p. 23.
15. STUIVER, M. and SUESS, H. E. Radiocarbon 8 (1966) 534.
16. SUESS, H. E. In Olsson, I. U. (Ed.) *l.c.*, p. 303. 16a. Plate I. 16b. p. 309. 16c. p. 310.
17. SOBERMAN, R. K. Phys. Rev. 102 (1956) 1399.
18. STUIVER, M. Science 149 (1965) 533.
19. HOUTERMANS, J., SUESS, H. E. and MUNK, W. In Radioactive Dating and Methods of Low-level Counting, *l.c.*, p. 57.
20. NYDAL, R. J. Geophys. Res. 73 (1968) 3617.
21. WALTON, A., ERGIN, M. and HARKNESS, D. D. J. Geophys. Res. 75 (1970) 3089.
22. VRIES, H. DE. Proc. Kon. Ned. Akad. Wetensch. Ser. B. 1961 (1958) 94.
23. RENFREW, C. Sci. Amer. 225 (1971) No. 4, 63.
24. WICKMAN, F. Geochim. Cosmochim. Acta 2 (1952) 243.
25. SMITH, B. N. and EPSTEIN, S. Plant Physiol. 47 (1971) 380.
26. WEBER, J. N. and WOODHEAD, P. J. Chem. Geol. 6 (1970) 93.

27. Keeling, C. D. Geochim. Cosmochim. Acta 13 (1958) 322.
28. Keeling, C. D. Geochim. Cosmochim. Acta 24 (1961) 277.
29. Keeling, C. D. Geochim. Cosmochim. Acta 24 (1958) 299.
30. Allen, L. H. Jr. Agr. Meteor. 8 (1971) 5.
31. Galimov, E. M. Geochem. Int. 3 (1966) 889.
32. Grubbs, F. E. Ann. Math. Statist. 21 (1950) 27.
33. Dixon, W. J. and Massey, F. J. Introduction to Statistical Analysis. McGraw-Hill, New York 1957, p. 275.
34. Gejvall, N.-G. and Engstrand, L. Personal communication.
35. Vries, H. de, Barendsen, G. W. and Waterbock, H. T. Science 127 (1958) 129.
36. Oakley, K. P. Science 140 (1963) 488.
37. Sellstedt, H., Engstrand, L. and Gejvall, N.-G. Nature (London) 212 (1966) 572.
38. Hultin, E., Liljeqvist, G., Lundblad, G., Paléus, S. and Ståhl, G. Acta Chem. Scand. 23 (1969) 3426.
39. Cochran, W. G. and Cox, G. M. Experimental Designs. Wiley, New York 1950, p. 92.
40. Brownlee, K. A. Statistical Theory and Methodology in Science and Engineering. Second Edition. Wiley, New York 1965, p. 346.
41. Lerman, J. C., Mook, W. G. and Vogel, J. C. In Olsson, I. U. (Ed.) *l.c.*, p. 275.
42. Encyclopaedia Britannica. Chicago etc. 1969. Vol. 5, 728.
43. Gore, W. L. Statistical Methods for Chemical Experimentation. Interscience. New York 1952.