

## STATISTICAL QUALITY CONTROL GRAPHS IN RADIOCARBON DATING

ROY SWITSUR

Cambridge University, Godwin Laboratory, Free School Lane, Cambridge, England

**ABSTRACT.** I describe here the establishment and use of statistical control graphs based on the analysis of variance for monitoring the stability of operation of radiocarbon dating counting systems.

### INTRODUCTION

Statistical Quality Control Graphs, pioneered and developed in the 1920s and 1930s, have found increasingly wide application in industry as well as in many research and analytical laboratories. A small number of radiocarbon dating researchers have used such graphs for many years as a basis for monitoring their radioactivity measurements. The recent increased interest in quality control and assurance by the radiocarbon community could bring about their more extensive use in the future. I describe here, briefly, the technique for the preparation and utilization of control graphs in conventional radiocarbon dating.

It is important for the radiocarbon research scientist to have confidence that his/her gas or liquid scintillation spectrometers are operating under stable and reproducible conditions. The random nature of the disintegration of the  $^{14}\text{C}$  isotope inevitably causes some dispersion of the data, but the extent of this may be predicted. For precise and accurate age determination, it is important to know that the system is under statistical control or when non-random or periodic changes are occurring. The term "in control" in the technical context of quality assurance, describes a process whereby stable system of chance appears to be in operation. The control graph technique has an appealing apparent simplicity, yet its power lies in its ability to aid in the separation of specific causes of variation. This enables us to diagnose a possible fault in the apparatus or system and hence could help in solving some of the problems encountered. Used sensitively, the technique may anticipate incipient problems, but it would be wrong to regard it as a panacea for all difficulties.

### BACKGROUND

The control graphs method is an application of the analysis of variance. The scientist chooses subgroups of sample data that are more likely to be similar to one another than to the overall readings. From these we can measure the variance. Checks are made for any significant difference between the means of the subgroups. We commonly use two measures of dispersion of the data; one derives from the variation of the mean range within the subgroups and the other is the root mean square deviation. The former practice stems largely from the earlier use of the graphs in industry when comprehensive tables were compiled for converting the range of subgroups of different sizes to an equivalent 'standard deviation'. This pre-computer method arose from the simple and rapid mental calculation of the range, whereas calculating the standard deviation without mechanical aids was difficult and liable to error. Now that computer spreadsheets, such as Lotus 1-2-3, are available, the computations have become trivial and we may use either of the measures of the data spread. The range method is still widely used because of its simplicity, but for subgroups of samples larger than ca 15, using the standard deviation is preferred.

In the control graphs technique, we assume that the variable plotted on the graph is normally distributed and may be described by two statistics, the mean value and the dispersion as given by

the standard deviation. It is usual to plot both the sample means and the range (or sample standard deviation) as complementary control graphs. This is because either a change in population mean or altered variability about the mean could indicate that the counting system is no longer in statistical control. These are known as the *Means Control Graph* and the *Range Control Graph* (or *Standard Deviation Control Graph*).

On the graph is drawn a central horizontal line representing the *Average* value of the variable to be plotted and on either side there are two further parallel *Limit* lines spaced according to the appropriate control conditions. We can infer from these whether samples are random from the same distribution or whether changes of distribution have occurred. Some variation is inevitable due to chance causes inherent in the system, but if these are the only causes of variability, then the system is under statistical control. On the Means Control Graph, the limit lines are symmetrical about the central line, except in the ultra-low count-rate Poisson range, but they are not equally spaced on the Range Control Graph or the Standard Deviation Control Graph. These additional lines are known as the *Inner* and *Outer Control Lines*, or perhaps more appropriately, the *Warning Limit* and the *Action Limit*. For a sample size,  $n$ , the means are approximately normally distributed with a Grand Mean value of  $\mu$  and variance of  $\sigma^2/n$ . Hence, 95% of the sample means will lie in the interval between the warning limits, the positions of which may be calculated from Equation 1:

$$\mu \pm 1.96 \frac{\sigma}{\sqrt{n}} . \quad (1)$$

Similarly, 99.8% of the sample means will lie within the interval between the action limits, the positions of which are calculated from Equation 2:

$$\mu \pm 3.09 \frac{\sigma}{\sqrt{n}} . \quad (2)$$

In the above, the grand mean is given by the average value of the subgroup sample means and the standard deviation is estimated by taking the square root of the variance between the samples' values. We can also estimate the standard deviation from the mean range of the subgroup samples,  $\bar{w}$ , assuming that the subgroup size is  $< 12$ .  $\sigma$  may be obtained from  $a_n \bar{w}$ , where  $a_n$  is given in Table 1.

TABLE 1  
Estimates of standard deviations from the mean range of samples  
for the Means Control Graph

n	2	3	4	5	6	7	8	9
$a_n$	0.8862	0.5908	0.4857	0.4299	0.3946	0.3698	0.3512	0.3367

The Standard Deviations Control Graph is similar to the Means Control Graph in that 95% of the samples will lie within the Inner Limits and 99.8% will lie within the Outer Limits. These limits, however, are not symmetrically arranged around the mean standard deviation. For sample size,  $n$ , and mean standard deviation,  $\sigma$ , the four limits, *Upper and Lower Warning Limits* and *Upper and Lower Action Limits*, may be calculated by multiplying  $\sigma$  by the limit factor,  $F_\chi\sigma$ , (where  $\chi$  is 1,2,3,4...) from Table 2 for the appropriate value of  $n$ .

TABLE 2  
Factors for the Warning and Action Limits  
for the Standard Deviation Control Graph

Limit = $F_\chi\sigma$				
Subgroup size	Upper Action factor F1	Upper Warning factor F2	Lower Warning factor F3	Lower Action factor F4
2	4.65	3.17	0.04	0.00
3	5.06	3.68	0.30	0.06
4	5.31	3.98	0.59	0.20
5	5.48	4.20	0.85	0.37
6	5.62	4.36	1.06	0.54
7	5.73	4.49	1.25	0.69
8	5.82	4.61	1.41	0.83
9	5.90	4.70	1.55	0.96
10	5.97	4.79	1.67	1.08
11	6.04	4.86	1.78	1.20
12	6.09	4.92	1.88	1.30

A concept central to the control graph technique is that of the rational subgroup. The success of the technique depends in no small measure on the researcher's discrimination in selecting the subgroups. Each should be made as homogeneous as possible by taking contiguous readings, but allowing some time to elapse before collecting the next subgroup data. This procedure allows for the maximum opportunity for variation between subgroups. The ideal subgroup size for many industrial processes is probably four, but this may not be the case in radiocarbon dating. It is important that the distribution of the means should be normal so that a reasonably large value for the number of readings in the subgroup sample would be an advantage. In this way, the distribution of the averages around the grand average becomes more compact as the subgroup size increases and the standard deviation is inversely proportional to the square root of the number of values in the subgroup. So, it is an advantage to make the subgroups as large as reasonably possible, for then the graphs become more sensitive to small variations in the average. Similarly, the larger the subgroup, the relatively narrower are the limits on the Standard Deviation Control Graph. Here again, smaller variations are easier to detect. Obviously, these statements are applicable only if the variations occur among the subgroups and not within them.

We can give no general rule for the frequency with which the subgroups need to be measured; this depends on the particular spectrometer or counter system. If the system is found to be relatively trouble-free, then the frequency could be as low as 5% of the unknown samples determined. With problematic systems, the frequency has to be increased to check that the counting data are reliable.

## ESTABLISHING CONTROL GRAPHS

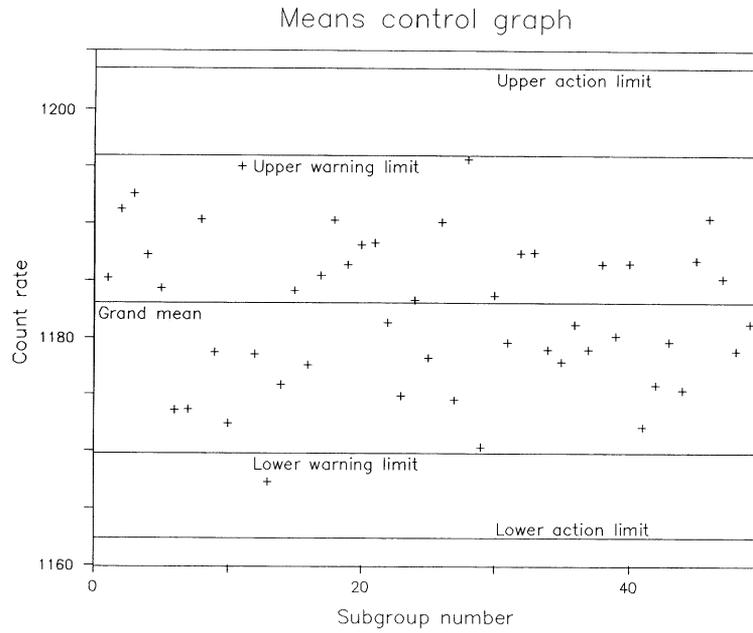
It may be useful to outline briefly a practical procedure for establishing control graphs. This would be equally suitable for monitoring a background standard such as carbon dioxide for gas counting equipment or analytical quality 'dead' benzene for a liquid scintillation spectrometer, or any other standard that may be required. In liquid scintillation spectrometry, for example, it is convenient to monitor sealed high-activity standards rather frequently since they require only very short counting times.

It is obvious that at the beginning of the process of establishing the control graphs, values of both the mean and the standard deviation of the specimen activity are unknown. We need to estimate at least approximate values for these before we can draw or plot any graph. We will update and subsequently redraw the control graphs as the values for the mean and standard deviation become better known. To start the graphs, we make a subgroup of, say five, determinations of the specimen count rate. From these we calculate the mean count rate and the mean range. (The range is given by the difference between the highest and lowest value in the subgroup.) This should be repeated to obtain 20 determinations of the means and 20 ranges. Next, we calculate the grand mean (GM) of the 20 count-rate means and the mean range (MR) of the 20 ranges. These are the basic data required for setting up the control graphs.

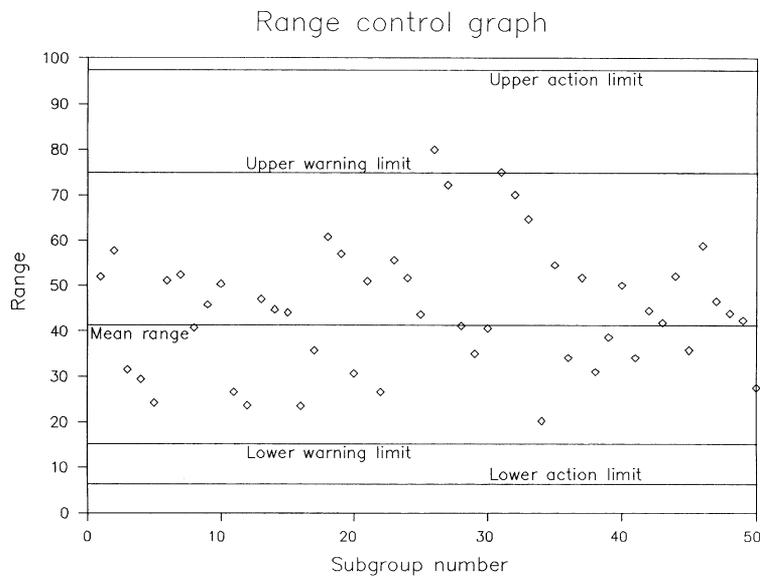
For the Means Control Graph (MCG), a graph of the variation of count rate with time, a line is drawn representing the GM. Two further horizontal lines are drawn on either side of the GM at a distance of  $0.377MR$ , corresponding to  $1.96\sigma$ ; these are the warning limits. The action limit lines, corresponding to  $3.09\sigma$ , are drawn at a distance of  $0.94MR$  on either side of the GM. We plot the 20 points representing the means of the subgroups to complete the graph. Figure 1A gives an example of a means control graph.

For the Range Control Graph (RCG), a graph of variation of range with time, a horizontal line is drawn representing MR. The theory indicates that the limit lines on this are not equally spaced as for the MCG. The upper action and warning limits are drawn at  $2.36MR$  and  $1.81MR$  and the lower warning and action limits are drawn at  $0.37MR$  and  $0.16MR$ , respectively. We plotted the 20 ranges of the sample subgroups count rates to produce the initial RCG. Figure 1B gives an example of an RCG.

We add further points to both of these graphs as the determinations become available. After some time, when the system has settled down and we obtain more data, we often find that the initial values of the GM and MRs need to be revised and the graphs replotted. We emphasize that when this procedure is carried out conscientiously for each counting system in a laboratory, much effort is needed for monitoring the various reference samples. In liquid scintillation counting, it may be necessary to monitor every dedicated counting vial in its own right. Each scientist must make provision for all these steps in his/her program, in terms of both time and budget, and although neither of these may be trivial, they are a requirement of the strict data analysis. For miniature or micro-gas counters, where the count rates are extremely low, the above procedure is prohibitive in counting time; an alternative technique for quality control should be adopted, using graphics and measurement of the time intervals between individual pulses. These intervals are compared with statistically calculated time intervals. The procedure, however, would still be quite lengthy.



**Fig 1A. Means Control Graph - showing system 'in control'**



**Fig 1B. Range Control Graph - showing system 'in control'. Graphs 1A and 1B complement each other.**

## INDICATIONS

Facility in using statistical control graphs, as with many techniques, increases with familiarity. We give here an outline of some of the obvious applications.

It is important to recognize at the outset that control graphs are not for retrospective use; they rapidly grow obsolete and do not need to be archived. Their real value lies in a current record which provides a basis for decision-making on immediate problems and their resolution during activity counting. It is convenient to enter the measurements into a spreadsheet, where the calculations are made and updated and the control graphs displayed on the monitor screen.

In what follows, we assume that the activity population is normally distributed so that most of the sample values fall within the warning levels on the graphs. The probability is that 1 value in 20 would lie outside the upper or lower inner limits and that 1 in 500 would fall outside the upper or lower outer limits of the means control graph. Figure 1A shows a typical set of readings with the appropriate control limits. Even in the event of points falling within the inner limits, these might still be masking possible interference within the subgroups and the variation between individual runs could be large. Thus, we need the range (or standard deviation) control graph to monitor the validity of this and check that there is no irregular interference. Figure 1B shows such a range control graph. It will be noticed that there is a tendency for some large range values to occur even though the system is in overall statistical control.

Lack of statistical control by the counting system may be revealed by a consistent change or variation in the value of the average with time. Alternatively, lack of control could be signaled by a change in the dispersion between subgroups; the change in the average and dispersion might occur simultaneously. To make these signs meaningful, we should plot our observations on the control graphs in the correct chronological sequence. A shift in the population average moves the mean sample average and the standard deviation graphs one way, but changes in the dispersion affect them differently. Changes of averages may be sustained over a long period of time or may be frequent and irregular. They may also be gradual and systematic. If the average value changes without change of dispersion, this indicates a change in bias.

If a measured value falls outside the inner limits but inside the outer limits, it may be taken as a warning only because we expect it to occur with 1 sample in 20. The next and subsequent samples are likely to be within the limits. The chances are only 1 in 500 that a point on the mean graph should fall outside either of the outer limits. This would be a signal to take action. The immediate action is to check the result with data from a further subgroup. If this value falls within the limits, then the outlier may have been within expectations or spurious and the system remains in statistical control. However, if the check sample result also falls beyond the outer limits, this exceeds probability expectations and the equipment should be examined at once for faults.

If values on the range or standard deviation control graph fall consistently outside the limits, spurious interference from nearby operating electrical equipment, such as a Tesla coil or a workshop lathe, might be to blame. A quiet supply phase should be sought.

A sustained decrease in the values of the means control graph could be due to a drift in the counter high voltage, movement of pulse-height discriminator levels or changes in amplification. It would be useful in this case to compare the control graph with one from an adjacent counting channel for the same period. If this shows an increase in values corresponding to the decrease observed on the first control graph, then it indicates that the counts are moving from one channel to the next and tends to confirm the possible drift of the high voltage or amplifiers. Figures 2A and 2B give an example of what would be seen on the graphs when a system becomes out of statistical control due to this effect. In Figure 2A, the values of the means graph fall below the GM established by the first 20 samples and progressively diminish with time without sign of recovery. Figure 2B shows that in an adjacent channel, the simultaneous count-rate values tend

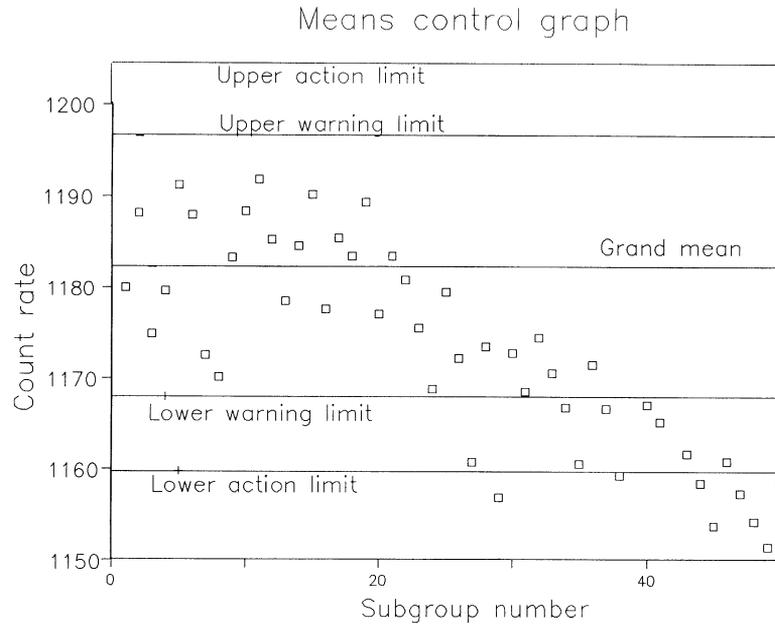


Fig 2A. Means Control Graph - showing lack of statistical control. Note the progressive diminution of count rate.

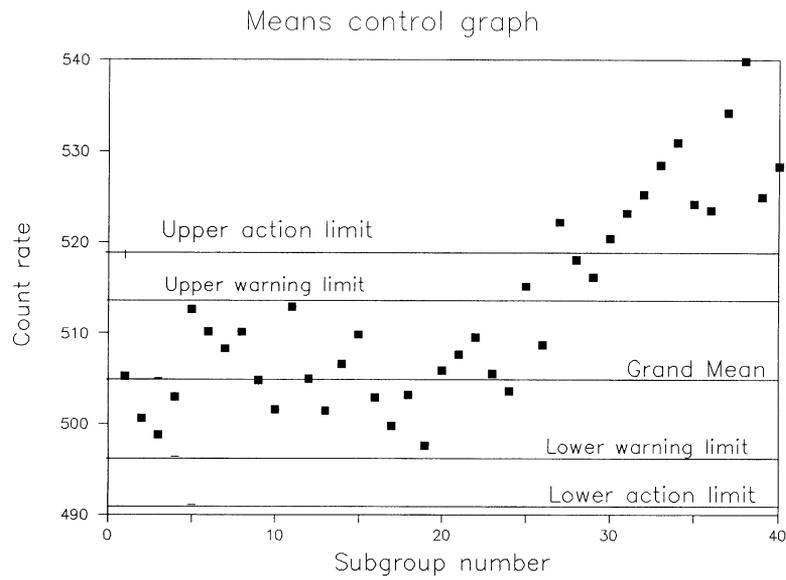


Fig 2B. Means Control Graph - showing lack of statistical control. Note the increase of count rate. Counting channel adjacent to that in Fig 2A.

to increase (the count rates are different because of unequal channel width) in a complementary manner. The values plotted here are fictitious; in a real situation, action would be taken to rectify the fault long before the state indicated by the later values on the graphs could occur.

In liquid scintillation counting, a steady decrease in the means control graph of a standard activity such as NBS oxalic acid, especially if this is accompanied by a similar decrease in an adjacent counting channel, probably does not indicate a problem with the electronics. This could easily be due to loss of activity from the evaporation of sample liquid from the counting vial. This effect might be expected if the vial had been used repeatedly for long periods. Thus, it is prudent to prepare fresh standards frequently as well as to monitor total vial masses. Every gas counter and vial constitutes an individual counting system that requires monitoring.

The values of the subgroup readings sometimes fall consecutively on the same side of the GM, as in Figures 2A and 2B. We have investigated statistically such runs, in long and short sequences, but for our purpose, we need note only a few results. For example, a change in counter background should be suspected only when seven or more successive readings lie on the same side of the GM as plotted. Values may tend to fall to one side of the mean although they may not be consecutive. In these cases the following rules may be of interest. If 10 values out of 11 consecutive points lie to the same side, suspect a change in the GM. This rule may be extended to include 12 points out of 14 consecutive ones, 14 out of 17, and 16 out of 20. This behavior suggests that consideration should be given to revising the value of the GM and the limit levels on the control graph. However, it should be borne in mind that the incidence of these sorts of runs is more probable than that of a point falling beyond the action limits.

#### CONCLUSION

The above are some of the more common indications given by statistical control graphs. Their consistent use in monitoring the current operation of counting systems will give increased confidence that the equipment is functioning under statistical control. The graphs will also help in the overall quality assurance in radiocarbon dating. Although important, statistical control is only one aspect of quality assurance. Other vital factors include standards, purification and preparation techniques, and the more subjective aspects such as sample integrity and experiment design. These are discussed in more detail by Long *et al* (1990) and Switsur (1990).

#### REFERENCES

- Long, A and Kalin, RM 1990 A suggested quality assurance protocol for radiocarbon dating laboratories. *Radiocarbon*, this issue.
- Switsur, VR 1990 A consideration of some basic ideas for quality assurance in radiocarbon dating. *Radiocarbon*, this issue.

*Editor's note:* This paper was not actually presented at the Workshop but was considered appropriate for inclusion in these Proceedings.