Tutorial ^{14}C wiggle-match dating: an Excel-approach

Maarten Blaauw

^{14}C wiggle-match dating assumes linear accumulation of, e.g., a peat core, and using this assumption it assigns calendar ages to all dated levels of such a core. Then the ^{14}C dates of the core are compared with the ^{14}C ages of the calibration curve at the assigned calendar ages.

By changing the two parameters involved in the linear accumulation model (a straight line relating depth to calendar age; a is the accumulation rate in yr/cm, ß is the horizontal placement), the match of the dates to the calibration curve is changed.
The goodness-of-fit of the core can be measured in several ways; here we use Maximum Likelihood. With this method, first the probability distribution on the calendar axis of every ^{14}C date is calculated. Then, using a certain combination of the parameters a and ß, a calendar age is assigned to every dated level. The height of the probability distribution of every date at that calendar age is then found. The product of all heights is calculated. The highest product will give us the best match of the ^{14}C dates of the core with those of the calibration curve.

In this tutorial, for a very small dataset of only 4 dates (way too small, but if we would have used more dates, calculation times would be far too long), we will find the best match with the Maximum Likelihood approach. First we will go through the file, and then we will perform a small exercise.

This tutorial is based on Blaauw *et al.* (2003); the 4 dates were taken from Mauquoy *et al.* (2002). The user is also suggested to look at Bpeat (Blaauw and Christen, in press) for more advanced techniques.

-Blaauw, M., Heuvelink, G.B.M., Mauquoy, D., van der Plicht, J. and van Geel, B. 2003: A numerical approach to ^{14}C wiggle-match dating of organic deposits: best fits and confidence intervals. Quaternary Science Reviews 22: 1485-1500.

-Blaauw, M., and Christen, J.A. in press. Radiocarbon peat chronologies and environmental change. Applied Statistics.

-Mauquoy, D., van Geel, B., Blaauw, M. and van der Plicht, J. 2002: Evidence from northwest European bogs shows 'Little Ice Age' climatic changes driven by variations in solar activity. The Holocene 12, 1-6.

You can mail me at blaauwm@cimat.mx or drieteenmeeuw@hotmail.com.

Let's go through the file first. It is called **WMD.xls**, and it contains several worksheets. Excel is by no ways the perfect software to calculate all this, but at least it works.

1) Open the first, leftmost worksheet, called '**Calibration curve**'.
There are loads of columns here:

The first part (columns A-E) shows the INTCAL04 ^{14}C calibration curve. The data are interpolated linearly, so as to get a ^{14}C age for every calendar age. (For this tutorial, in order to keep the file small only a part of the calibration curve is included. If you wish to have the entire dataset, ask me.)

Column A: the calendar years in cal BP.
Column B: the corresponding ^{14}C ages of the calibration curves, linearly interpolated when necessary.
Column C: This is a column that cumulates all ^{14}C ages. Don't worry; this is only needed for getting Excel to calculate wiggle-matching solutions.
Column D: the 1 standard deviation (sd) error limits from the calibration curve, interpolated when necessary.
Column E: the 1 sd error limits, now accumulated, again for calculation purposes.

The second part (columns G-K) again contains the calibration curve data, but now only those calendar years where the calibration curve has data, so, without any linear interpolation. These data are needed for drawing the calibration curve in a chart (in the worksheet 'Wiggle-match dating'; see later).

Column G: the calendar years in cal BP.
Column H: the corresponding ^{14}C ages of the calibration curves.
Column I: the 1 sd error limits from the calibration curve.
Column J: the ^{14}C ages with the 1 sd's subtracted. These are then the lower limits of 1 the standard deviation error envelope of the calibration curve.
Column K: similar to column J, but now showing the upper limits for the 1 sd error envelope.

The third part (columns M & N) contains the calibration curve data again, but now ordered from the youngest to the oldest ^{14}C dates (for calculation purposes).

In column P, references are given for the calibration curves. The files can be downloaded from www.radiocarbon.org.

2)   Now let's move on to the next worksheet, called '**Wiggle-match dating**'. This is where you can actually see what happens to the dates when they are wiggle-matched to the calibration curve. Here you can also change the match by hand, by changing the parameters a and ß.

The graph shows the calibration curve (1 standard deviation error envelope; green lines) and the 4 dates of the sequence (blue; vertical lines show the 1 sd ^{14}C age error bars, horizontal lines show the deposition period of the samples). Again, 4 dates would normally not be enough to obtain a good, reliable wiggle-match; for time- and file-size saving purposes we will stick to 4 dates only for now.
Do you see how the placements of the dates already indicates which of the local optima of, say, the lowermost date, is the most probable one, because of the placements of its neighboring dates? With calibration of individual dates, you would not know which of the local optima would be the most probable.

In column A are depicted the levels (depths) where ^{14}C dates have been taken, column B gives the actual midpoints of these depths, column C gives the thickness of those samples (divided by half), column D gives the ^{14}C ages and column E gives the 1 sd errors of the ^{14}C dates as supplied by the AMS laboratory.

Then we get the most important column, column F, where calendar ages are assigned to the dated levels, with help of two parameters:

Calendar age = a × depth + ß

In fact, it is a bit more complicated than this, because using this formula, the anchor point (point around which the line turns) would be at depth = 0. We choose another, more convenient anchor point so that the line turns at the middle of the sequence:

Calendar age = a × (depth – average depth) + average calendar age + ß.

The average calendar age is the calendar age at which the calibration curve has the same ^{14}C age as the average ^{14}C age of all dates in the sequence.

The parameter a (the accumulation rate in yr/cm) is printed in cell K1; the parameter ß is given in cell L1; the average depth and average calendar age are calculated in the column where the calendar age for every dated level is calculated, column F. Just click on cell F2 for example, to see how the calendar age is calculated. Try to find out which cells and worksheets are referred to.

-The ROUND function gives us calendar ages rounded to a year (no decimals, so ROUND(*value*;0)).

-The $ sign fixes columns, lines or single cells, so that if the formula is copied to other cells, the same column, line or cell will be referred to.

- The AVERAGE function calculates the average of a range of cells, e.g., cells B2 to B5.

- (VLOOKUP((AVERAGE(D$2:D$5)),'Calibration curve'!M:N,2,TRUE) means: the average ^{14}C age of the sequence is looked up in the calibration curve (in the worksheet 'Calibration curve'), and the corresponding calendar age is returned.

In column G, using the accumulation rate (a) and the (half of the) thickness of the samples, the (half of the) deposition period of the samples is calculated. They make the horizontal bars on the blue dates in the chart.

There we are! You can change the parameters a and ß, and see what happens to the placements of the dates.

In fact, you should try thousand and thousands of possible combinations of a and ß to see how they change the fit of the dates with the calibration curve. Also, it would be nice if we could find out how good a fit is, and if we could tell whether one combination of parameters gives a better fit than the other. That is what the other sheets are for. They calculate the Maximum Likelihood, as explained shortly on page 1.

3) Let's proceed to the sheet '**Probability distributions**'. That's where for every ^{14}C date, the probability distribution of its calendar age is calculated. Every ^{14}C date has its mean (μ) and 1 standard deviation error (s), as provided by the AMS laboratory. We assume that this date has a Gaussian distribution on the ^{14}C axis:

$$p_x = \frac{1}{s\sqrt{2p}}\, e^{-[(x-m)^2/2s^2]}$$

x is a value. If x is far away from μ, p_x becomes small. See the Gaussian curve on the y-axis of the figure below. This example is a ^{14}C date of 2450 ± 50 BP. We can calculate the chance (the height of the probability distribution) that x is, say, 2450, or 2460, or 2500.

Now. We have the ^{14}C calibration curve, where of tree-rings of exactly known calendar age, the ^{14}C ages have been measured precisely. We could find the calendar age(s)! where the calibration curve is exactly 2450 ^{14}C BP, in order to find the calendar age(s) of our ^{14}C date of 2450±50 BP. But, this value of 2450 is a measurement only, so we need to find out how big the chance is that the actual calendar age of the sample is, say, 2900 cal BP, or 2899 cal BP, an so on. We calculate these chances by finding the ^{14}C age of the calibration curve that corresponds to a certain calendar age, and placing this value (x) in the above formula. In this way the height of the probability distribution of a ^{14}C date on the calendar axis is found. This is the same as calibrating an individual ^{14}C date.
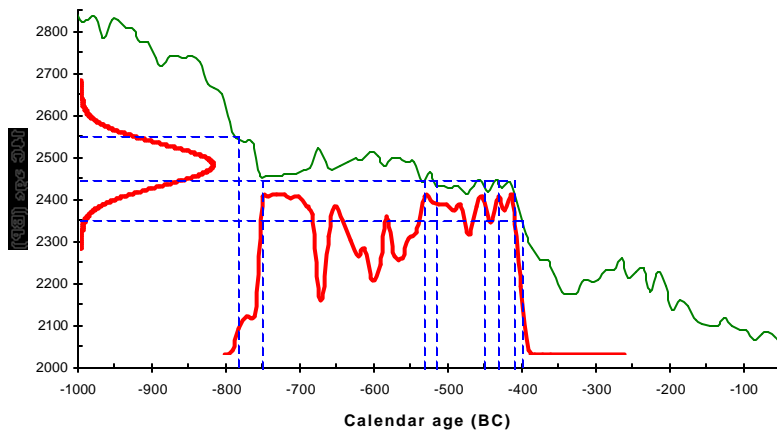
That's what done in the present sheet. Columns A, B and C give the calendar ages, ^{14}C ages and 1 standard deviation error bars of the calibration curve, and the following columns give the probability distributions of the four dates of our sequence. Go to cell D5 and see what it says. It's actually the formula above; it calculates the probability distribution for the date 410±30 ^{14}C BP at cal BP 0. The s is composed of the s of the ^{14}C date, and of the s of the calibration curve at the corresponding calendar year. The lines below line 5 contain the probability distributions for all following years until cal BP 25979. For calculation purposes, the values are cumulative. Sorry for that.

4)   Let's go calculate and visit the worksheet '**Maximum likelihood**'. This sheet is composed of three parts. In the leftmost part, calendar ages are assigned to dated depths, using a combination of parameters a and ß. We have seen this before in the worksheet 'Wiggle-match dating'. The average depth and the calendar age of the calibration curve corresponding to the average ^{14}C age of the sequence are calculated in cells E1 and E3, respectively. The parameters are taken from an automated list in columns E and F. This list starts with the chosen values 2 for a and –100 for ß. In the lines below, the parameters are changed according to some rules.

Click on cell E6 and find out how a changes, and do the same for cell F6 for ß. In order to reduce calculation time for this tutorial, a and ß change in very big steps (0.5 for a, 5 for ß). For precise calculations, the parameters would need to change in far smaller steps (e.g., 0.02 for a, 1 for ß).

Now the fits are calculated (columns H to K). Cell H5 says something complicated:
=(VLOOKUP((A5+(ROUND((\$E5*A\$2),0))),'Probability distributions'!\$A:\$G,4)-VLOOKUP((A5-(ROUND((\$E5*A\$2),0))-1),'Probability distributions'!\$A:\$G,4))/(ROUND(((2*\$E5*A\$2)+1),0)).

This formula finds the height of the probability distribution of the ^{14}C date; it searches for the calendar age of cell A5 in the previous worksheet, and finds the corresponding 'height'. But, as every sample has a certain thickness, it has accumulated during a certain period. So, the average height of the probability distribution over this deposition period is calculated.

Finally, the product of all heights of the probability distributions is calculated for every combination of parameters a and ß; this is depicted in column G. Those are the values we are interested in.

Suggested task: Calculate the goodness-of-fit for a range of combinations of the parameters a and ß, and by doing so, find the best combination of a and ß, so, the best match of the ^{14}C dates of the sequence to those of the calibration curve.

This can be done automatically in the sheet 'Maximum likelihood'. Select cells A9 to K9. In the lower right corner of the selection rectangle, there is a small square. Point your mouse there until the cursor changes into, probably, a small black cross. Click and hold the left mouse button, and move downwards until you arrive at, say, line 1111. This will give you just enough combinations of a and ß so that all combinations of (2<a<15) and (–100<ß<+100) are tried (with big steps though, as explained above). Let go of the mouse button and wait while the computer calculates for you.

Now there's a whole list of combinations of a and ß, and their corresponding assigned calendar ages, heights of the probability distribution, and the product of the latter.

Make (separate or composite) graphs of the probability distribution of the four dated levels. E.g., for the date at 91.5 cm depth, plot cells D5 to D1111 (x-axis) against cells G5 to G1111 (y-axis). You can do this by selecting both ranges (hold down ctrl to select separate ranges) and finding the option to make an XY scatter chart. Alternatively, you can choose to make an XY scatter chart, go through the options, and select the appropriate data ranges when asked for.

You will now see a scatter of points. The highest point is the one that gives the best fit. Find which combination of a and ß made this fit, and fill these values in the appropriate cells in the 'Wiggle-match dating' worksheet to see how the dates fit.