Using Lomb-Scargle frequency analysis to interpret radiocarbon dated Holocene fluvial units in the UK

Jessica Moloney¹, Tom Coulthard², Jim Freer² and Michael Rogerson¹

¹Department of Geography, Environment and Earth Science, University of Hull, UK
²Department of Geography, University of Bristol, UK

INTRODUCTION

There are a large number of fluvial archives available for the UK that are used to interpret past flood frequency. A database of fluvial archives collected in the UK has been compiled (Macklin et al., 2012). This database has been analysed using probability distribution functions (PDFs) to identify periods of flooding in the fluvial record (Lewin and Macklin, 2003; Johnstone et al., 2006). There are limitations with this approach, which have been discussed by Chiverrell et al. (2011).

This study re-analyses the database and uses a new method of time-series frequency analysis: Lomb-Scargle. Lomb-Scargle frequency analysis can be used with unevenly spaced data (Scargle, 1989) to show significantly significant cycles, whilst incorporating the ‘noise’ of the data. The results and interpretations from both methods are compared.

METHODOLOGY

‘Change after’ dates (samples found at the top of a sedimentary boundary marking the youngest age of the lithological change; Macklin et al., 2006) were selected from the database (245 data out of 776). Uncalibrated ages with associated error were calibrated using IntCal13 (Reimer et al., 2013) on OxCal (version 4.2, Bronk Ramsey, 2009). Summed probabilities from calibrated ages were plotted against time BP to create a PDF for ‘change after’ dates and all available data, based on a methodology used in Macklin et al. (2012). A cumulative probability frequency (CPF) plot was generated by dividing the summed probabilities of the ‘change after’ data by the summed probabilities of all available data.

Lomb-Scargle frequency analysis was carried out on Matlab using a standard script provided by Trauth (2010).

RESULTS

Probability distribution function and cumulative probability function

Figure 1 (left image): Probability distribution function per year of calibrated radiocarbon ages. Black = all fluvial archive data, blue = ‘change after’ data. Figure 1 could be interpreted as showing a general increase in flooding since the mid to late-Holocene until ~10000BP, followed by a decrease in flooding. Periods of flooding could be inferred by continuous peaks, such as 28000BP-30000BP, 32000BP-40000BP, 46000BP-50000BP, 59000BP-62000BP and 112000BP-118000BP.

Figure 2 (right image): cumulative probability function of summed probabilities of ‘change after’ dates divided by summed probabilities of the whole dataset.

Lomb-Scargle frequency analysis

Figure 3: output 1 from the Lomb-Scargle analysis has produced a plot of the wavelength of cycles against power. The largest peak in power has a wavelength of ~9390. There are two smaller peaks with wavelengths of ~3000 and ~4520.

Figure 4: output 2 from the Lomb-Scargle analysis has produced a plot of the wavelength of cycles against significance (p-value) for the null hypothesis. P-value < 0.05 (signified by the dashed line) are considered significant for this data. The three peaks in power shown in figure 3 correspond to p-values < 0.05. The cycles presented in this data are significant.

DISCUSSION

Probability distribution function and cumulative probability function

The number of data used to construct the ‘change after’ curve exceeds the number of samples necessary to create statistically significant results (Michczyńska and Pazdur, 2006).

From figure 1, both curves show an increase in probability until ~10000BP, then the probability decreases until 0BP. There are fluctuations throughout the Holocene within these two general patterns. The increase in probability could be the result of increased flooding or sample preservation that causes sample number bias in the late-Holocene.

Peaks in PDF curves have been attributed to increases in periods of flooding (Macklin et al., 2006; Macklin et al., 2012). However, the PDF approach is limited by uncertainties within the data and the methodology. The shape of the PDF is affected by the radiocarbon dating process, the calibration curve and the distribution of samples per time.

Lomb-Scargle frequency analysis

Lomb-Scargle frequency analysis was applied to the ‘change after’ data because this method is suitable for unevenly spaced data (Scargle, 1989). Figures 3 and 4 suggest that cycles within the ‘change after’ data occur with wavelengths of ~3000, ~4520 and ~9390 years.

The Lomb-Scargle frequency analysis identifies a signal within the ‘change after’ data and this confirms that the data is not random. A signal is present within the data but may not be consistently recorded in all of the samples in the ‘change after’ dataset. There is noise within the data, which disrupts the signal. There must be a sufficient number of data over a long enough time period for a reliable signal to be identified within the data.

Other uncertainties are still present within this method of frequency analysis. For example, peaks could be related to the radiocarbon dating process or to the paleo-flood data.

The next steps in this research will include:

Carrying out Lomb-Scargle frequency analysis on the data per depositional environment in order to identify if the signal is sensitive to the depositional site
Identify the source(s) of the signal through comparison to other climate proxies that are radiocarbon dated
Conclude if the signal present is strong enough within the current data available to make a confident interpretation

CONCLUSION

The large number of fluvial archives in the UK makes it a valuable source of data.

There is no robust methodology that has been applied to UK fluvial archives to date that take into account the uncertainties discussed above and in Chiverrell et al. (2011).

This study presents a method of analysing the statistical significance of cycles and wavelengths using a technique that is suited for unevenly sampled data.

More analysis is needed on the fluvial archives from the UK to test the sensitivity of the data to identify the signal present.