Periodicity in extinction rates

ANATOLY D. ERLYKIN[1], DAVID A.T. HARPER[2], TERRY SLOAN[3] and ARNOLD W. WOLFENDALE[4]

[1]Lebedev Institute, Russian Academy of Sciences, Moscow, Russia
[2]Palaeoecosystems Group, Department of Earth Sciences, Durham University, Durham DH1 3LE, UK
[3] Department of Physics, Lancaster University, Lancaster LA1 4YB, UK
[4]Department of Physics, Durham University, Durham DH1 3LE, UK

We report a reinvestigation of the peak at a period of 27 myr in the Fourier periodogram of Phanerozoic mass extinction data. This peak is interesting since it implies a repetitive cause of the extinctions such as might be expected from an astronomical source. In a previous paper, we showed that the peak had insufficient statistical significance to be considered a real effect. A recent paper by Melott and Bambach (2017) disagreed with our conclusion mainly on the grounds that the peak has a higher statistical significance than we considered. Here we re-examine the data and demonstrate that Melott & Bambach overstate the statistical precision of their Fourier analysis. Hence Fourier analysis alone is insufficient evidence for extinctions occurring repetitively. We show that the peak is fragile and disappears if five events with large extinction proportions are excluded from the sample of 136 events. Since the causes of these five events are unlikely to have a repetitive nature we conclude that the peak is most probably caused by a statistical fluctuation.

INTRODUCTION

The case for periodicity in mass extinctions and its relationship to astronomical drivers is not new but it remains controversial. In the light of a series of recent studies, essentially reviving the exciting discoveries of the 1980s, our paper (Erlykin, Harper, Sloan & Wolfendale 2017, referred to hereinafter as Erlykin *et al*. 2017) sought to present an alternative view. In that paper, we showed that the oscillation of the solar system about the Galactic plane is probably too variable in time to account for the possibly observed periodicity of 27 myr. Hence it is implausible as a cause of the periodicity as has been claimed (e.g. Rampino & Strothers 1984, Fox 1987). We further showed that the statistical significance of the observed 27 myr periodicity could be poorer than has been previously thought so that the possibility that it is a statistical fluctuation cannot be excluded.

In the lack of such periodicity, it is difficult then to relate the key astronomical processes that may have matched this pattern and thus driven mass extinctions. Melott & Bambach, in a series of detailed papers, have espoused the opposing view and they offer a critique of our position (Melott & Bambach 2017). They focus mainly on the statistical significance of the observed peak, claiming a higher statistical significance than we deduced in Erlykin et al. (2017). Moreover, they challenge our study on a number of grounds. Here we focus on the core of the argument: Is the 27 myr cycle real or a statistical artefact?

In the present paper, we examine the work reported in Melott & Bambach (2017). We show that they have overestimated the statistical significance of the observed 27 myr peak. Firstly we show that the statistical significance cannot be determined precisely enough to make the firm conclusion

that the peak is highly significant. This is due to statistical fluctuations from the small number of extinction events (136 in total). Secondly, we show that the noise power is much closer to white noise than the red noise assumed in Melott & Bambach 2017. This reduces considerably the statistical significance of the peak. Thirdly we show that the peak disappears if five of the 136 extinctions are excluded. Four of these events have terrestrial origins which are unlikely, in our view, to be repetitive in nature. This provides further evidence that the peak is caused by a statistical fluctuation in a small data sample.

THE PERIODOGRAMS AND THE SIGNIFICANCE OF THE 27 MYR PEAK

The peak at 27 myr arises from Fourier analysis of the extinction data (see Appendix A). As in our previous publication (Erlykin et al. 2017) and that of Melott & Bambach (2017), we use the time series of the extinctions from the work of Bambach 2006, Melott and Bambach (2011a,b, 2013 and 2014). For the analysis described here the data were selected within the 0-465 Ma range. This is identical to the range reported in Melott & Bambach (2017) and yielded a total of 136 extinction events. Figure 1 shows the extinction proportion plotted as a function of the age of the extinction with a linear fit to the data. Such linear fits are used to detrend the data.

Fourier analysis of the time series shown in Figure 1 produces the periodograms shown in Figure 2. The peak at a period of 27 myr (with extinction proportion amplitude 0.0389 or power 0.00151) is clearly obvious. Melott & Bambach (2017) produced similar plots which are in broad agreement with those shown in Figure 2. They add lines on their plots which indicate that the 27 myr peak has a significance of 98%. Here significance is defined as the probability that any peak appearing in the plot should be smaller than that observed. Such a high value of significance would allow the peak to be identified as most probably real allowing the conclusion that there is a repetitive cause to the extinction rate.

Melott & Bambach (2017) used a commercial software package (Autosignal version 1.7) for their work with a procedure which assumes that the noise is frequency dependent and biased towards low frequencies, i.e. so-called red noise. In the following we examine the validity of this assumption. One of the difficulties in assessing the noise level and the significance of a result in a periodogram is the fact that all points in it are strongly correlated since they come from one set of data. Hence determining the noise from the data is fraught with difficulty. One way round this is to attempt to mimic "independent" data sets using Monte Carlo techniques. This also is a difficult procedure since all data sets generated must resemble the single input data set and so are not truly independent.

As in our previous publication (Erlykin et al. 2017), as well as treating the real data, we pass randomly generated data sets through our software which is written by us in the computer language Fortran 77. The random data sets were generated by the Monte Carlo technique. These were used to assess the significance of any peak by counting the number of times that peaks of larger magnitude than the selected peak appeared as multiple random data sets were passed through the software. We show that the noise is almost independent of frequency (i.e. it is almost "white") so that this procedure is valid.

MONTE CARLO GENERATION OF RANDOM DATA SETS

We use the standard Monte Carlo procedure of generating a number according to a given probability distribution. The method is illustrated in Figure 3. The upper panel of Figure 3 shows a histogram of the deviations of the extinction proportions from the linear fit in Figure 1. The smooth curves in this figure represent a maximum likelihood fit (James 1972) to the data of a Gaussian plus an exponential tail (the dashed curves in Figure 3). As described in Erlykin et al. (2017) this fit was of the form for the probability $\Pi(d)$ of obtaining deviation, d, from linearity (see Figure 1)

$$\Pi(d) = N(\frac{1}{\sqrt{2\pi}p_2} \exp\left(-\frac{(d-p_1)^2}{2p_2^2}\right) + \theta(d)(p_3 exp - (p_4 d))) \qquad (1)$$

where *N=136* is the total number of events in the time series, $\theta(d)$ is the mathematical function which is zero for negative and +1 for positive d and the p_i are free parameters. The values of the latter to fit the data were $p_1 = -0.0319 \pm .0096, p_2 = 0.0469 \pm .0064, p_3 = 1.83 \pm 1.33$ and $p_4 = 7.77 \pm 3.48$. The values after the $\pm$ sign give the standard errors (68% confidence level) in each parameter from the fit. These are quite large due to the small number of 136 events in the sample fitted and are highly correlated.

The middle panel of Figure 3 shows the integral of the solid smooth curve in the upper panel from the left hand edge to the value on the x axis and normalised to unity at the largest value of x. To obtain a random data point with probability distribution $\Pi(d)$, a uniformly distributed random number between 0 and 1 is generated. The value of the deviation, *d*, is read off on the x axis from the value of this random number on the curve in the y direction. Repeating this 136 times generates a random data set which is distributed as the real data. A typical random data set generated in this way is illustrated in the lower panel of Figure 3. The data set is then completed by adding the contribution of the linear fit shown in Figure 1.

The ages of the extinctions could either be kept fixed at the observed ages or varied randomly from 0-465 Ma. The differences reported below between these two cases were found to be insignificant.

Another way to make a semi-independent data set is to randomly interchange the extinction proportions and ages. This can be repeated on a random basis many times and is referred to as shuffled below. Two kinds of shuffling were tried: one was to shuffle after detrending the data and the second was to shuffle before detrending. The latter method increases the noise which leads to a lower significance. If the linear trend of the data (see Figure 1) is an artefact of the way extinctions are detected and measured then the latter method would be correct. If the linear trend is a real effect then the former method is correct.

STATISTICAL FLUCTUATIONS IN THE DATA

Statistical fluctuations are to be expected with the small sample of 136 events i.e. if we had several samples of extinctions as a function of age they would differ due to such fluctuations. In consequence, we would compute different results and significances for each. We now address the

question of the size of such differences in significance to be expected in the Fourier analysis of the sample of 136 extinction events.

We make random data sets in different ways to see how the significance level changes between each. Note these are not truly independent data sets since they all originate from the same single available one. Hence, we use the term "semi-independent".

As shown in Appendix A the value of the mean noise power $\bar{P}$ in equation A9 plays a key role in determining the significance of an observed peak in a periodogram. Since it is exponentiated, small changes in $\bar{P}$ will lead to large changes to the significance. The values of $\bar{P}$ and the significance are compared in the following from several semi-independent data sets. As in Erlykin *et al.* (2017) we compute the significance by Monte Carlo, generating multiple data sets which are each passed through the Fourier analysis chain. The significance level is taken from the number of times a peak larger than the observed 27 myr peak appeared in the multiple random data sets. This procedure is valid when the noise is close to "white" as we show below. The procedure was criticised in Melott & Bambach (2017) and their criticism would be valid in the case that the noise is strongly "red" as they assume.

Here we use both methods described in the previous section to generate semi-independent data sets and the difference is taken as a lower limit on the statistical fluctuations in the size of the significance level from a data sample of this size. We use the term lower limit since the measurements are only semi-independent, all originating from the single given data set. Hence truly independent data sets are likely to have larger fluctuations than those we measure from these semi-independent data sets.

The fit shown in Figure 3 is used to make eight other semi-independent samples. Each parameter in the fit was fixed at values either up or down by the standard error, in turn, and the fit remade to determine a new set of parameters. Each new set of parameters was used to form a different semi-independent data set. Random data runs (5000 in number) were then generated with the new parameters and the noise power computed for each frequency and for each run.

Figure 4 shows the mean values of $\bar{P}$ (averaged over the 5000 random runs) as a function of the frequency for two of the semi-independent data sets. It can be seen that the values of $\bar{P}$ are almost independent of frequency but that the mean values from the two semi-independent methods differ by a factor of nearly two. The difference here is attributable to the statistical fluctuations between the two semi-independent samples of 136 extinction events. (The average value of $\bar{P}$ over the frequency range is denoted by $<\bar{P}>$). The fact that $\bar{P}$ is almost independent of frequency shows that the noise is approximately white.

Table 1 shows the results of this procedure. Column 1 gives the fit number. Columns 2-5 show the parameters used to generate the semi-independent data set and the integral probability (Fig. 3, middle panel). Column 6 gives the Fisher χ^2 for the fits with these parameters to the data. All are acceptable. The poorest fit was fit 4 for which the probability to find a larger value of χ^2 more than 17.8 (for 12 degrees of freedom) is 12% which is quite acceptable. In addition all 9 fits were inspected to see that the fitted curve gave a reasonable representation of the data. Column 7 gives the value of the mean power averaged over all frequencies and all 5000 runs denoted by $<\bar{P}>$. The penultimate column gives the total number of occurrences of random runs with a peak of amplitude

greater than the value 0.0389 (or power of 0.00151) i.e. that of the 27 myr peak in the data in Figure 2. The final column gives the significance $(1 - (N > 0.0389)/5000)$, here defined as the probability that all peaks in the periodogram have amplitude less than the 27 myr peak.

Table 1

Fit	P1	P2	P3	P4	χ^2	N > .0389	$1000 *$ $< \bar{P} >$	Significance(%) (S)
1	-0.0319	0.0469	1.83	7.77	12.1	490	0.226	90.2
2	-0.0254	0.0510	0.818	4.30	14.6	991	0.264	80.2
3	-0.0381	0.0432	3.31	11.3	10.5	143	0.185	97.1
4	-0.0180	0.0576	0.50	6.67	17.8	22	0.150	99.6
5	-0.0408	0.0412	3.16	8.20	15.1	1549	0.293	69.0
6	-0.0373	0.0400	2.24	7.81	14.9	766	0.250	84.7
7	-0.0275	0.0533	1.52	8.01	12.3	263	0.203	94.7
8	-0.0220	0.0510	1.43	7.51	13.5	250	0.201	95.0
9	-0.0420	0.0436	2.26	7.59	14.2	1190	0.275	76.2
Shuffled	After detrend	--	--	--	--	379	0.227	92.4
Shuffled	Before detrend	--	--	--	--	1720	0.290	65.6

Figure 5 shows the significance (final column in Table 1) plotted as the data points against the value of mean noise power, $<\bar{P}>$ (column 7 of Table 1). The strong correlation between the significance, S, and $<\bar{P}>$ is evident. The smooth curve represents the calculated variation of $1 - S$ with mean noise power $< \bar{P} >$, where S is defined in equation A10, and with $P_{pk} = 0.00151$, the observed power in the 27 myr peak. The value of M has been chosen arbitrarily as 80 to give the best visual representation of the data points. This is less than the 180 frequencies at which the Fourier analysis was made because of correlation effects. The curve gives a qualitative representation of the data.

Table 1 and Figure 5 show that the significance deduced in this way could be anything from less than 70% up to 100%. This range encompasses the values in Erlykin *et al.* (2017) and Melott & Bambach (2017). The true statistical fluctuations could be larger than this range since the data sets used to determine the range are only semi-independent.

Hence the expected statistical fluctuations resulting from the small number of extinctions are too great to say with certainty whether or not the 27 myr peak is truly statistically significant. Therefore it is invalid to state with certainty that the significance is 98% as asserted Melott & Bambach (2017) since the significance could be lower than 70%. From this we conclude that the possibility that the peak is a statistical fluctuation cannot be excluded.

THE NOISE IN THE FOURIER PERIODOGRAM

Figure 4 illustrates that the noise power $\bar{P}$ averaged over the 5000 random data sets is roughly independent of the frequency in the frequency range 0.01 to 0.1 myr^{-1} for two typical semi-independent data sets. The remaining semi-independent data sets showed similar behaviour. Regression lines fitted to the values of $\bar{P}$ with frequency for all the semi-independent data sets showed small upward slopes which indicated a variation of between a 0.0 and +1.6% over the frequency range 0.01-0.1myr^{-1} . From this we conclude that the noise power is "white" (with a tinge of blue since the slopes are slightly positive with increasing frequency).

Melott & Bambach (2017) use the software package Autosignal 1.7 in their analysis. They make the assumption that the noise in the Fourier Spectrum is red i.e. is strongest at low frequency. Their parameter AR(1) was set to be somewhat greater than 0.8 i.e. strong red noise. Autosignal then gives a 98% significance to the observed 27 myr peak. However, we have shown above that the noise is roughly independent of frequency i.e. is effectively white so their assumption of red noise is invalid.

Random noise spectra were generated as described above and fed through the Autosignal 1.7 package. When the noise is selected to be red [by clicking the automatic setting of the parameter AR(1)]many peaks appear with a significance of greater than 99%. [Note the automatic selections of AR(1) always gives a value less than that assumed in Melott & Bambach 2017]. Feeding 99 random spectra through in this manner gives a total of 161 peaks with greater than 99% significance level in the frequency range 0.01 to 0.1 myr^{-1} where the expected number is of order 1. This discrepancy between the observed and expected number of peaks with 99% significance gives a clear indication that the selection of strong red noise for this data sample is incorrect.

With the ssumption that the noise is white [i.e. setting AR(1)=0] no peaks were observed with significance level of greater than 99%.

Given our observation that the noise is effectively white the significance of the 27 myr peak should be judged assuming white noise rather than red noise. Autosignal would then give a much lower significance level of the 27 myr peak of somewhat less than 50%.

The assumption of red noise in the Fourier periodogram in Melott & Bambach (2017) without the use of Monte Carlo simulation is contrary to the warning in section 3.13 of Muller and MacDonald (2000). That their significance is overestimated can be seen from their figure 2A in which two extra peaks appear at frequencies of 0.07 and 0.09 myr^{-1} with significance judged by their procedure of greater than 90%. Given that the noise is almost white the significance of these two peaks must really be much less than 90% and so they are most probably due to random noise. The probability to find two peaks in the periodogram from random noise with significance of more than 90% is less than 1%, again illustrating that the significances Mellott and Bambach (2017) are overestmated.

SENSITIVITY OF THE 27 MYR PEAK TO LARGE EXTINCTION PROPORTION EVENTS

It can be seen from equations A2 to A5 that events with large values of the extinction proportion (the values $f(t_j)$ in these equations) play an important role in determining the Fourier power at any frequency. The sensitivity to events with large extinction proportions was therefore investigated.

There are 12 events with extinction proportions greater than 0.3 in the data. The effect of these 12 events is shown in Figure 6. Figure 6 shows the periodogram with all 136 events included with the 27 myr peak indicated (upper panel for comparison, as in Figure 2). The middle panel shows the periodogram from 124 events excluding those with extinction proportion more than 0.3. Comparing the upper and middle panels, it can be seen that the 27 myr peak has disappeared from the middle plot. This illustrates the sensitivity of the 27 myr peak to events with large extinction proportion.

Investigating in detail the 12 events showed that the 27 myr peak could be eliminated by excluding 4 of them, the remaining 8 events making little difference to its amplitude. To check that these 4 events are unique, Fourier analysis was carried out on samples of 132 extinction events drawn from the total population of 136, excluding events 4 at a time. This was repeated for every one of the 13,633,830 possible combinations of 136 events with 4 excluded. During this procedure a further one of the 12 large extinction proportion events, when combined with 3 of the four identified events, was observed to lead to the elimination of the 27 myr peak.

If none of these 5 identitied events are excluded the 27 myr peak was robust for all the combinations and the amplitude was distributed as an almost perfect Gaussian with peak at 0.0396 and standard deviation 0.0012. None of these combinations eliminated the 27 myr peak (here we define elimination of the peak as the reduction of its amplitude to less than 0.032, 6.3 standard deviations below the peak value). Such Gaussian behaviour is to be expected if the 5 identified events combined with random noise are responsible for the 27 myr peak. For combinations including any of the 5 events the distribution developed a low amplitude tail. The peak could be eliminated by 51% of the 85150 combinations which excluded 2 of the 5 events, 99.8% of the 1310 combinations excluding 3 of the 5 events and all 6 combinations with either 4 or all of the 5 events excluded. Hence, the 5 identified events are unique in that they, or a subset of them, eliminate the 27 myr peak if they are excluded from the periodogram. They are therefore the ones responsible for the observation of the peak. Fourier analysis was made of each of the subsamples of either the 5 or the 12 large events. The 27 myr peak amplitude was much less than the noise generated with such small numbers of events and therefore could not be identified.

The lower plot of Figure 6 shows the periodogram with the 5 events excluded. The 27 myr peak has disappeared.   Further investigation shows that the 5 events arrive within a time window of 8 myr around the maximum of the reconstructed 27 myr sine wave i.e. 29% of the sine wave cycle.The large amplitude events with these arrival times then reinforce each other at a period of 27 myr to generate the observed peak in the periodogram. The remaining 7 large events are more random in time and so only contribute to the noise.

These observations could mean that the 5 events are repetitive in nature or that they happened by an accident of timing.

 The probability that they happened accidentally is difficult to calculate and depends on the assumptions made. The following illustrates the difficulty. It was shown above that the 27 myr peak is eliminated if almost all combinations in which 3 of the identified events out of the 12 large extinction proportion events are excluded from the perioodogram. The binomial probability to have 3 of the events within the quoted time range out of the 12 large events is 72.2%. On the other hand if we raise the level of definition of large events to extinction proportion of more than 0.38 only 6 of the large events remain (which include the five identified events). The binomial probability to have 3

of them at the times to create the peak out of these 6 is 23.7%. If one demands combinations which exclude four of the identified events these probabilities become 47.6% and 6.3%, respectively. If one further demands to have all five events at the appropriate times the probabilities become 25.0% and 0.93%. All of these are large enough for the identified events to have arrived in time by accident except the last one. However, this is a rather extreme demand since we have shown that the peak can be eliminated by excluding combinations with either 3 or 4 of the identified events .

Hence, the probability that the identified events occurred accidentally at the appropriate times to cause the 27 myr peak is rather high and it is quite plausible that the peak is a statistical fluctuation.

To look for further evidence to decide whether or not this is a chance occurrence or a repetitive effect we look to the causes of the five events. The five identified events are the end Cretaceous event at 66 Ma, the Triassic event at 201.4 Ma, the Permian events at 252.2 Ma and 259.8 Ma and the Ordovician event at 445.2 Ma. These events have been recently reviewed (e.g. MacLeod 2014; Bond and Grasby 2017; both contain substantial bibliographies).

> 1. The end Cretaceous event at 66 Ma. An asteroid or cometary strike on the Earth (Alvarez *et al*. 1980) is thought to have been the cause of this event although there were climatic events and volcanic eruptions leading up to the final extinction. There is a significant school of thought that believes the extinctions were already well underway before the final asteroid or cometary strike, due to an enhanced greenhouse world, with unusually high sea levels, initiated by the eruption of the Deccan Traps (e.g., Li & Keller 1998; Keller *et al*. 2017). The impactor provided the ultimate assault. Significantly it is the only extinction associated with strong geological evidence for an impactor.

> 2. The Triassic event at 201.3 Ma. This event was associated with massive volcanic eruptions from the Central Atlantic Magmatic Province (CAMP) and may have lasted over an extended interval of several million years. Intense global warming was associated with elevated levels of atmospheric CO_2 driving ocean acidification (Hallam and Wignall 2004; Wignall and Bond 2008; Bond & Grasby 2017). Both marine and terrestrial biotas were affected.

> 3. The end Permian event at 252.2 Ma. This was caused by a super, run-away greenhouse event initiated by massive volcanic eruptions of the Siberian Traps (and volcanicity in South China), followed by gas hydrate exhalation; the effects were devastating on both land and in the oceans as the effects of global warming, anoxia and sulphide toxicity were felt (e.g. Wignall 2001; Benton & Twitchett 2001; Knoll *et al*. 2007; Bond & Grasby 2017).

> 4. The major extinction during latest Guadalupian (Middle Permian) at 259.8 Ma is commonly termed the Capitanian extinction event (Wignall *et al*. 2009). The event was coincident with submarine volcanicity associated with the Emeishan Large Igneous Province (Jerram *et al*. 2016). The various eruptive modes and their products had a major effect on both marine and nonmarine environments and are considered to be the key driver for the extinctions through widespread anoxia and the release of greenhouse gases (Jerram *et al*. 2017).

5. The Ordovician event at 445.2 Ma. This event was associated with a well-documented short, sharp but major ice age which generated a multi-causal event (e.g., Sheehan 2001; Harper *et al*. 2014; Finnegan *et al*. 2016, 2017).

The first of the five events could arguably be from a repetitive source if it is caused by the release of a comet from the Oort cloud by tidal effects during the oscillation of the Solar System through the Galactic plane, as has been speculated (Randall and Reece 2014). However, it is difficult to see how the other four, each with a terrestrial cause within present knowledge, could be repetitive in nature.

From this we conclude that the 27 myr peak results from the five large amplitude events which just happen to arrive at the right time by serendipity to cause the observed peak. Hence the 27 myr peak is most probably created accidentally by a statistical fluctuation. Such accidents are common when dealing with small numbers of events.

THE CLUSTER ANALYSIS IN MELOTT & BAMBACH 2017

In Melott & Bambach (2017) intervals of marked extinction were identified numbering 19 events in total. Of these events 11 were shown to be in a cluster in the quarter of a period between -1.1 to +5.9 myr from a peak in the sine wave of period 27.2 myr with the first peak at 11.3 myr. These 11 events included the four of the five events discussed in the previous section which triggered the conclusion in Melott & Bambach (2017) that the 27 myr period is real. Melott & Bambach (2017) showed that the probability to observe such clustering by chance at the frequency of the 27 myr peak is small and of order 0.2%.

We agree with this probability calculated from 11 out of 19 independent events falling within a single quarter period at a single frequency. However, the events are not independent. Four of the events discussed above were among the 11 in the first quarter period and we have shown that these four events trigger the frequency to be considered (i.e. the 27 myr period). Hence only 15 of the 19 events are truly independent and of these only seven out of the 11 in the first quarter period can be considered as independent. The binomial probability to have seven or more events in the quarter period out of a total of 15 events is 5.7%. While this is not a large probability it is too large to allow the firm conclusion that the observed clustering is a real effect and not a statistical fluctuation.

From this we conclude that the cluster argument used to justify the claimed periodicity in Melott & Bambach (2017) is weak.

CONCLUSIONS

The confirmation of the 27 myr peak seen in the Fourier periodogram of mass extinctions would be interesting because it implies that such extinctions were repetitive in time. Melott and Bambach claim that the peak is statistically significant at the 98% confidence, allowing the peak to be seen as a probable real effect. We have shown that the statistical fluctuations in the small number of extinction events (136 in total) do not allow the statistical significance to be determined with such certainty and a much lower significance is possible. Furthermore, we have shown that the noise is almost white (i.e. independent of frequency). This reduces considerably the statistical significance of

the observed peak compared to the significance level claimed by Melott and Bambach who assumed red noise.

Furthermore, the peak is fragile and can be eliminated by excluding five events of large amplitude. A contributory cause of one of these events, the end Cretaceous event 66 Ma, could be from an asteroid or comet impact on the Earth. However, the other four events are thought to have a terrestrial origin and are unlikely to be repetitive in time. This indicates that the five events create the 27 myr peak accidentally leading to the conclusion that the peak is due to a statistical fluctuation.

The absence of a significant peak when these five large extinction proportion events are removed indicates that the behaviour among the remaining 131 extinctions is at most weakly repetitive.

Given these considerations the evidence for a repetitive source of the extinctions is not compelling as claimed in Melott & Bambach (2017). The observed 27 myr peak in the Fourier analysis is most probably created by a statistical fluctuation in the small number of events from which the periodogram is derived. In conclusion it appears that 'Mass extinctions are diverse and vary in intensity, selectivity, and timing. They are not homogeneous in effect or in cause.' (Bambach 2006).

Appendix A Fourier Analysis

A function of time, $f(t)$, can be decomposed into a series of sinusoidal oscillations. The contribution of any partilcular angular frequency, ω, is given Fourier transform.

$$\widehat{f}(\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t}\, dt = a(\omega) - ib(\omega) \qquad \text{(A1)}$$

Two methods have been used here to compute the real and imaginary components of the Fourier amplitudes a and b.

The first method is that of simple averaging and has been used throught this paper (unless otherwise stated). This is done by summing the $N = 136$ readings of extinction proportion $f(t_j)$ at times t_j. A little algebra shows that the real and imaginary Fourier coefficients at angular frequency ω_i are given approximately by

$$a(\omega_i) = \frac{2}{N}\sum_{j=1}^{N} f(t_j)\cos\omega_i t \qquad \text{(A2)}$$

$$b(\omega_i) = \frac{2}{N}\sum_{j=1}^{N} f(t_j)\sin\omega_i t \qquad \text{(A3)}$$

where $\omega_i = 2\pi f_i = 2\pi/T_i$ with f_i and T_i the frequency and period, respectively. This method will be referred to as the simple averaging method.

In MB 2017 paper using Autosignal version 1.7 the Fourier coefficients are derived using the Lomb Scargle approach (Scargle 1982) where the real and imaginary coefficients (equation A1) are defined as

$$a(\omega_i) = \sqrt{\frac{2}{N}}\frac{1}{\cos\omega_i\tau}\frac{\sum_{j=1}^{N} f(t_j)\cos\omega_i(t_j-\tau)}{\sqrt{\sum_{j=1}^{N}(\cos\omega_i(t_j-\tau))^2}} \qquad \text{(A4)}$$

$$b(\omega_i) = \sqrt{\frac{2}{N}} \frac{1}{\cos \omega_i \tau} \frac{\sum_{j=1}^{N} f(t_j) \sin \omega_i(t_j - \tau)}{\sqrt{\sum_{j=1}^{N} (\sin \omega_i(t_j - \tau))^2}} \qquad \text{(A5)}$$

$$\tan 2\,\omega_i \tau = \sum_{j=1}^{N} \sin 2\omega_i t_j \,/\, \sum_{j=1}^{N} \cos 2\omega_i t_j \qquad \text{(A6)}$$

These methods of averaging each give approximate values of the integrals needed to determine the Fourier coefficients.

The power $P(\omega_i)$ is given by

$$P(\omega_i) = a(\omega_i)^2 + b(\omega_i)^2 \qquad \text{(A7)}$$

and the amplitude by

$$\text{amplitude} = \sqrt{P(\omega_i)} \qquad \text{(A8)}$$

A key factor in deciding the statistical significance of a value of the power is the statistical probability p_S of obtaining a value of the power which is larger than that observed at a particular value of angular frequency $\omega = 2\pi/T$ where T is the period.

$$\text{This is given by } p_S = \exp\left(-\frac{P(\omega)}{2\sigma^2}\right) = \exp\left(-\frac{P(\omega)}{\bar{P}}\right) \qquad \text{(A9)}$$

where σ^2 is the mean square deviation of the values of a_i and b_i (averaged over the two) (Mueller and MacDonald 2000} and $\bar{P}$ is the mean noise power in the spectrum.

Equation A9 can be used to compute the significance of a peak. For example, Scargle gives the statistical significance, S, of a peak with power P_{pk} (i.e. the probability of seeing a peak which is larger than the observed peak in a spectrum dominated by noise) as

$$S = 1 - (1 - \exp -P_{pk}/\bar{P})^M \qquad \text{(A10)}$$

where M the number of periods sampled. This formula is derived assuming that the M different periods are independent of each other. This is not the case if the significance is derived from a single spectrum since the same data sample is used to determine the power in each period. Hence the individual periods are strongly correlated. Hence care must be exercised in interpreting formulistic significance tests such as this one. (NB, the significance quoted on the plots and in table 1 is the probability that all peaks in the spectrum should be smaller than the observed peak, i.e. it is $1 - S$, defined in equation A10).

Our analysis, here as in Erlykin *et al.* (2017), proceeds by taking the input data set and detrending it by subtracting a linear least squares fit to it (as shown in Figure 1). The resulting data set is then passed through the Fourier program either by using the simple averaging or Lomb-Scargle method. No other processing of the data is made before Fourier analysis. By passing distributions of 136 random events generated with a pure sinusoidal oscillation we found that the simple averaging technique gave roughly a factor two better resolution of the peak than the Lomb-Scargle method. Hence most of the analysis reported in this paper employs the simple averaging technique.

Bibliography

ALVAREZ, L. W., ALVAREZ, W., ASARO, F. and MICHEL, H. V. 1980. Extraterrestrial cause for the Cretaceous-Tertiary extinction. *Science,* 208, 1095-1108.

BAMBACH, R.K. 2006. Phanerozoic biodiversity mass extinctions. *Annual Review Earth Planetary Science,* 34, 127-155.

BENTON, M.J. and TWITCHETT, R.J., 2003. How to kill (almost) all life: the end-Permian extinction event. *Trends in Ecology and Evolution*, 18, 358–365.

BOND, D.P.G. and GRASBY, S.E. 2017. On the causes of mass extinctions. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 478, 3–29.

BOND, D.P.G. and WIGNALL, P. 2014. Large igneous provinces and mass extinctions: an update. *Geological Society of America Special Paper*, 505, SPE505-02.

ERLYKIN A.D., HARPER D.A.T., SLOAN T. and WOLFENDALE A.W. 2017 Mass extinctions over the last 500 Myr. An astronomical cause*? Palaeontology,* 60, 159-167.

FOX W.T. 1987 Harmonic analysis of periodic extinctions*. Paleobiology*, 13, 257-271.

FINNEGAN, S., RASMUSSEN, C.M.Ø. and HARPER, D.A.T. 2016. Biogeographic and bathymetric determinants of brachiopod extinction and survival during the Late Ordovician mass extinction. *Proceedings of the Royal Society, Biological Sciences,* 283, 20160007.

FINNEGAN, S., RASMUSSEN, C.M.Ø. and HARPER, D.A.T. 2017. Identifying the most surprising victims of mass extinction events: an example using Late Ordovician Brachiopods. *Biology Letters,* 2017400.

HALLAM, A. and WIGNALL, P.B. Discussion on sea-level change and facies development across potential Triassic-Jurassic boundary horizons, SW Britain. *Journal of the Geological Society*, 161, 1053-1056.

HARPER, D.A.T., HAMMARLUND, E.U. and RASMUSSEN, C.M.Ø. 2014. End Ordovician extinctions: A coincidence of causes. *Gondwana Research*, 25, 1294–1307.

JAMES, F. 1972. MINUIT - Function Minimization and Error Analysis CERN Program Library Long Writeup D506. [This package was used for the fitting described in this paper.]

JERRAM, D.A., WIDDOWSON, M., WIGNALL, P.B., SUN YADONG, LAI XULONG, BOND, D.P.G. and TORSVIK, T. 2016. Submarine palaeoenvironments during Emeishan flood basalt volcanism, SW China: Implications for plume–lithosphere interaction during the Capitanian, Middle Permian ('end Guadalupian') extinction event. *Palaeogeography, Paleoclimatology, Palaeoecology*, 441, 65-73.

KELLER, G., MATEO, P., PUNEKAR, J., KHOZYEM, H., GERTSCH, B., SPANGENBERG, J., FONT, E., BITCHONG, A. and ADATTE, T. 2016. Volcanism in the end-Cretaceous mass extinction and Paleocene-Eocene thermal maximum events. *Palaeogeography Palaeoclimatology Palaeoecology*.

KNOLL, A.H., BAMBACH, R.K., PAYNE, J.L., PRUSS, S. and FISCHER, W.W. 2007. Paleophysiology and end-Permian mass extinction. *Earth and Planetary Science Letters*, 256, 295–313.

LI, L. and KELLER, G. 1998. Abrupt deep-sea warming at the end of the Cretaceous. *Geology*, 26, 995–998.

MACLEOD, N. 2014. The geological extinction record: History, data, biases, and testing. *In*: KELLER, G., and KERR, A.C. (eds) *Volcanism, Impacts, and Mass Extinctions: Causes and Effects*. Geological Society of America Special Paper 505, 1–28.

MELOTT, A.L. and BAMBACH, R.K. 2010. Nemesis reconsidered. *Monthly Notices of the Royal Astronomical Society,* 407, 99-102.

MELOTT, A.L. and BAMBACH, R.K. 2011. A ubiquitous ~62-myr periodic fluctuation superimposed on general trends in fossil biodiversity. I. Documentation. *Paleobiology,* 37, 92-112.

MELOTT, A.L. and BAMBACH, R.K. 2013. Do periodicities in extinction—with possible astronomical connections—survive a revision of the Geological Timescale? *The Astrophysical Journal*, 773, 6.

MELOTT, A.L. and BAMBACH, R.K. 2014. Analysis of periodicity of extinction using the 2012 geological timescale. *Paleobiology*, 40, 177-196.

RAMPINO, M.R. and STOTHERS, R.B. 1984. Terrestrial mass extinctions, cometary impacts and the Sun's motion perpendicular to the galactic plane. *Nature,* 308, 709-712.

RANDALL L. and REECE M. 2014. Dark matter as a trigger for periodic comet impacts. *Phys. Rev. Letts.* 112, 161301 (https://arxiv.org/abs/1403.0678).

MELOTT A.L. and BAMBACH R.K. 2017. Periodicity in the extinction rate and possible astronomical causes. *Palaeontology* pp 1-10, doi:10.1111/pala.12322.

MULLER R.A. and MACDONALD G.J. 2000. Ice ages and astronomical causes. Springer & Praxis Publishing, Chicheter.

SCARGLE J.D. 1982. Studies in astronomical time series II. Statistical aspects of spectral analysis of unevenly spaced data. *The Astrophysical Journal*, 40, 835 – 853.

SHEEHAN, P.M., 2001. The Late Ordovician mass extinction. *Annual Review of Earth and Planetary Science*, 29, 331–364.

WIGNALL, P.B. 2001. Large igneous provinces and mass extinctions. *Earth-Science Reviews*, 53, 1–33.

WIGNALL, P.B., SUN, Y.-D., BOND, D.P.G., IZON, G., NEWTON, R.J., VÉDRINE, S., WIDDOWSON, M., ALI, J.R., LAI, X.-L., JIANG, H.-S., COPE, H., BOTTRELL, S.H. 2009. Volcanism, mass extinction and carbon isotope fluctuations in the Middle Permian of China. *Science*, 324, 1179–1182.
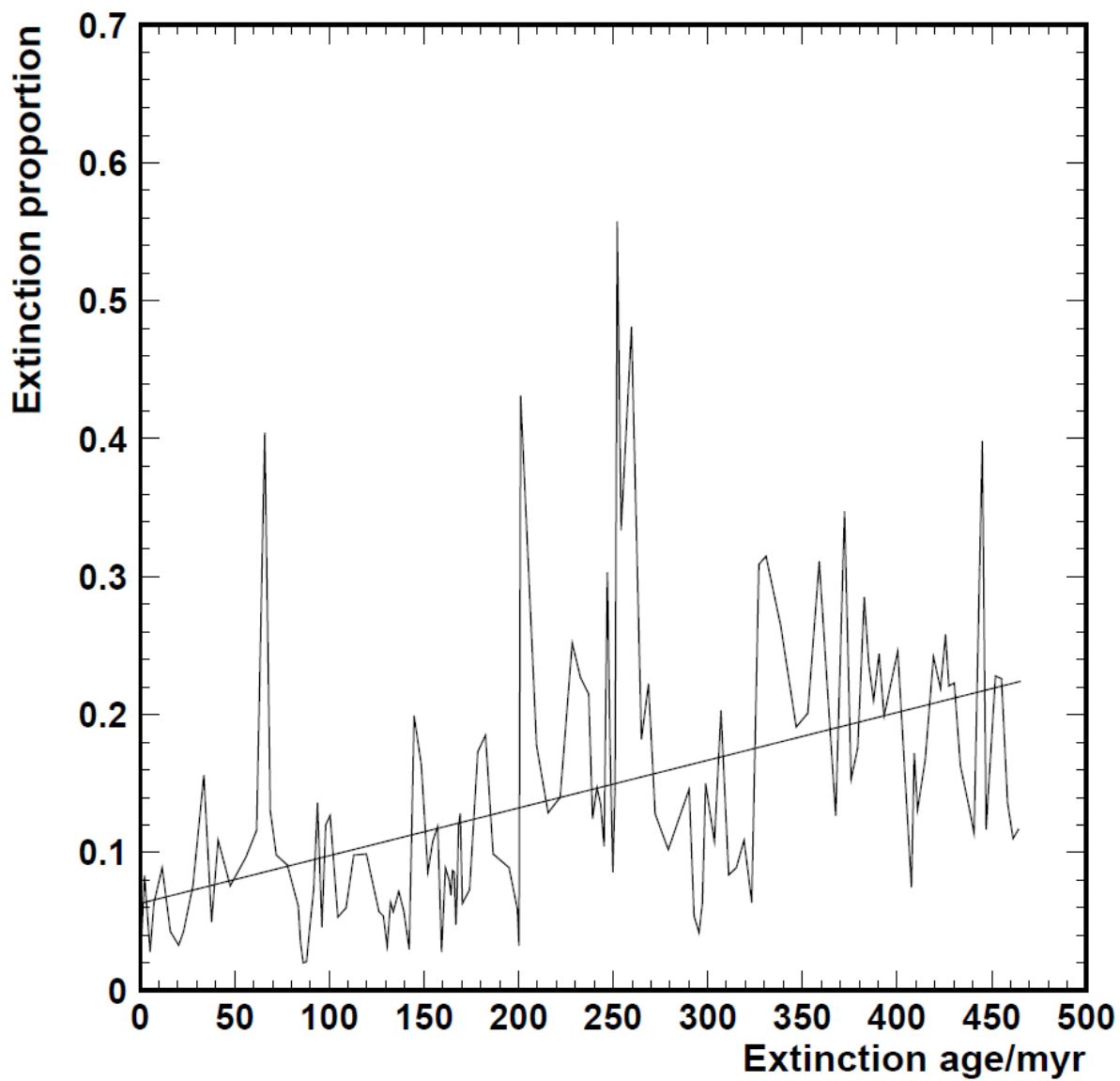
Figure 1 The extinction proportion versus extinction age for the 136 events in the data sample of ages from 0-465 Ma. The line represents a linear least squares fit to the data.
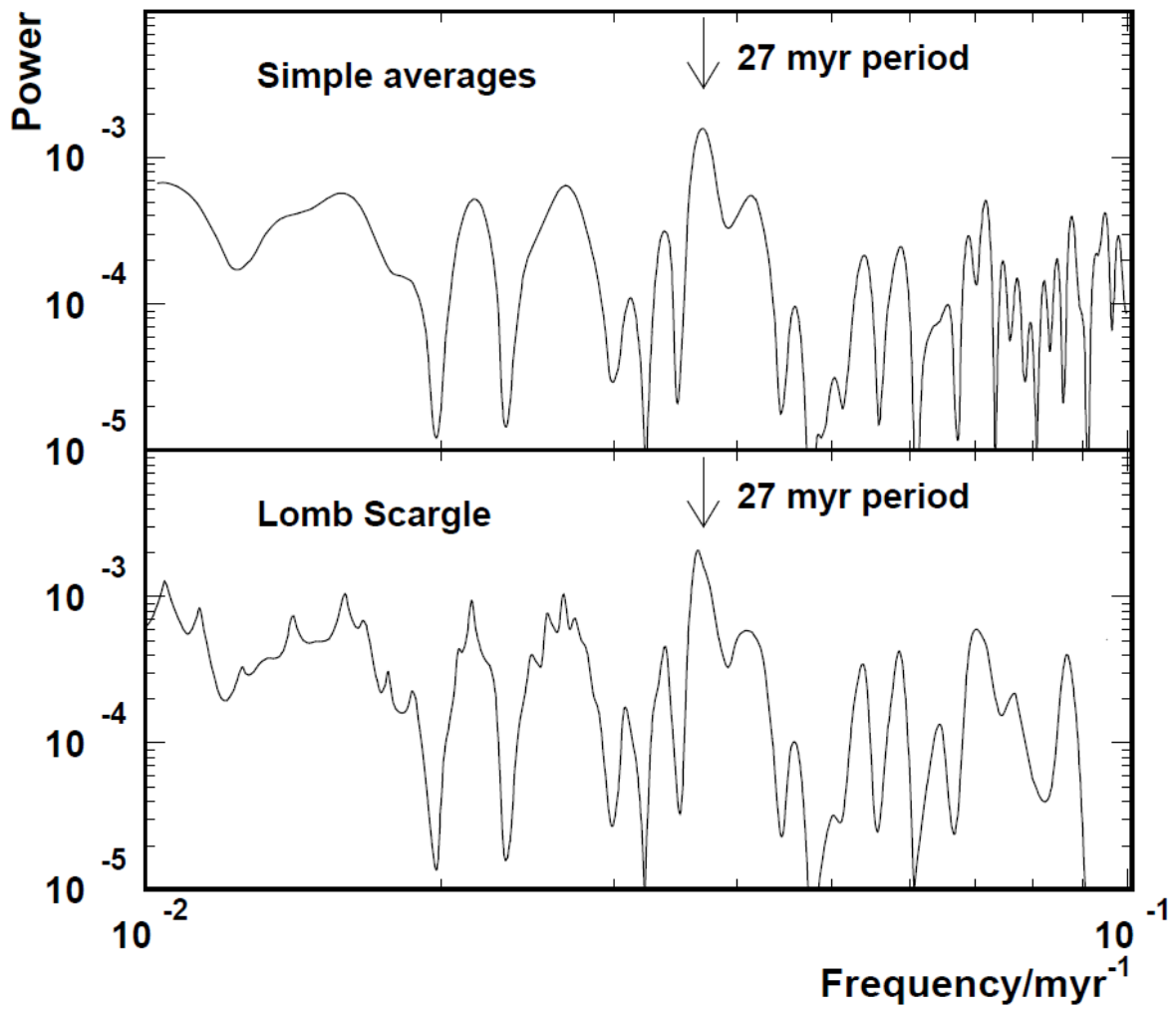
Figure 2 Periodograms of the 136 extinction events. Upper panel using the simple averaging method (equation A2 and A3). Lower panel using the Lomb-Scargle procedure (equations A4 to A6).
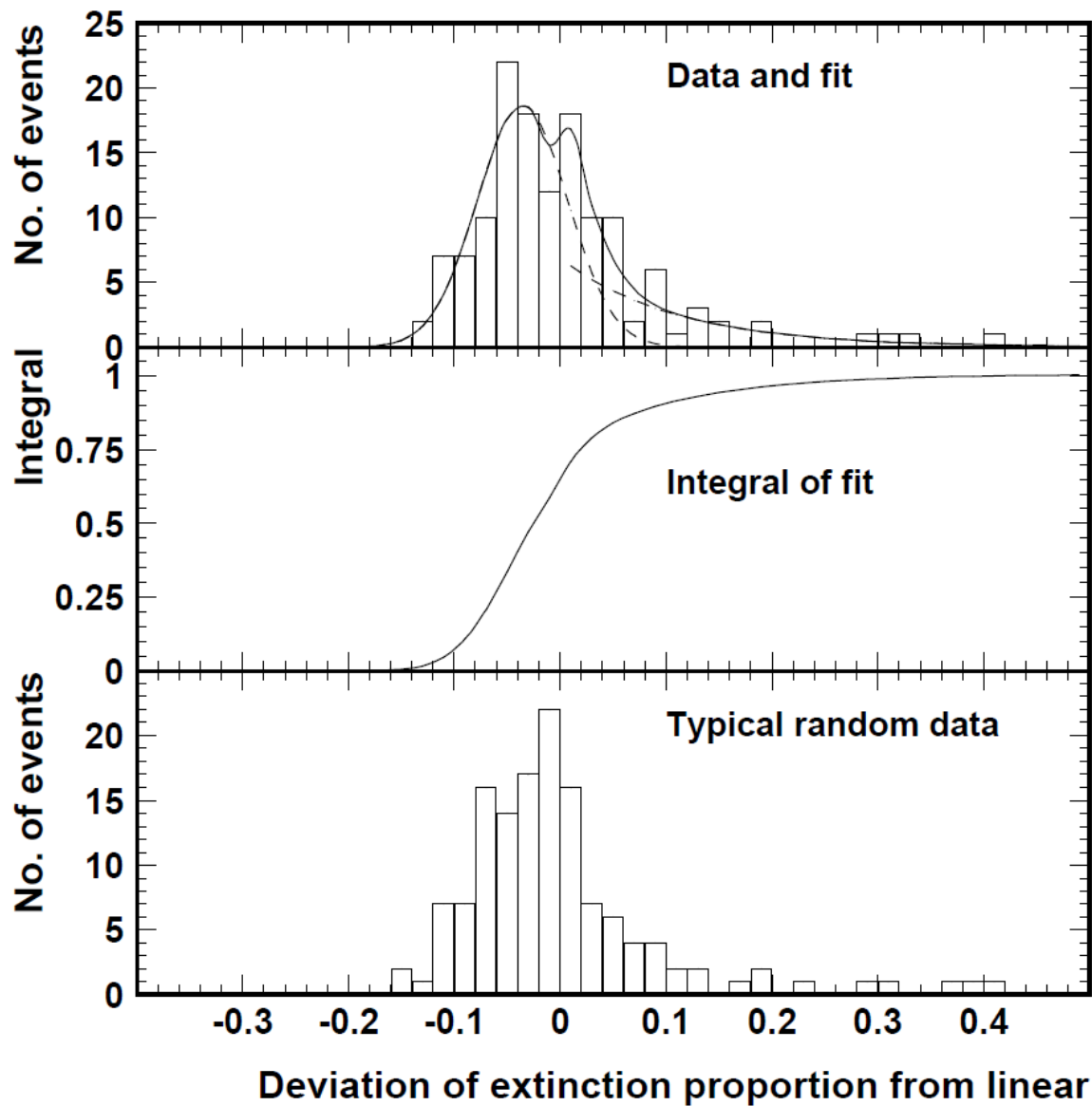
Figure 3 The upper panel shows a histogram of the deviations from the linear fit shown in Figure 1. The smooth curves shows the fit of the function given in equation (1). The dashed curves show the contributions of the Gaussian and exponential components. The middle panel shows the integral of the smooth function in the upper panel (normalised to unity). The lower panel shows a typical random spectrum, distributed as the real data, generated using the curve in the middle panel, as described in the text.
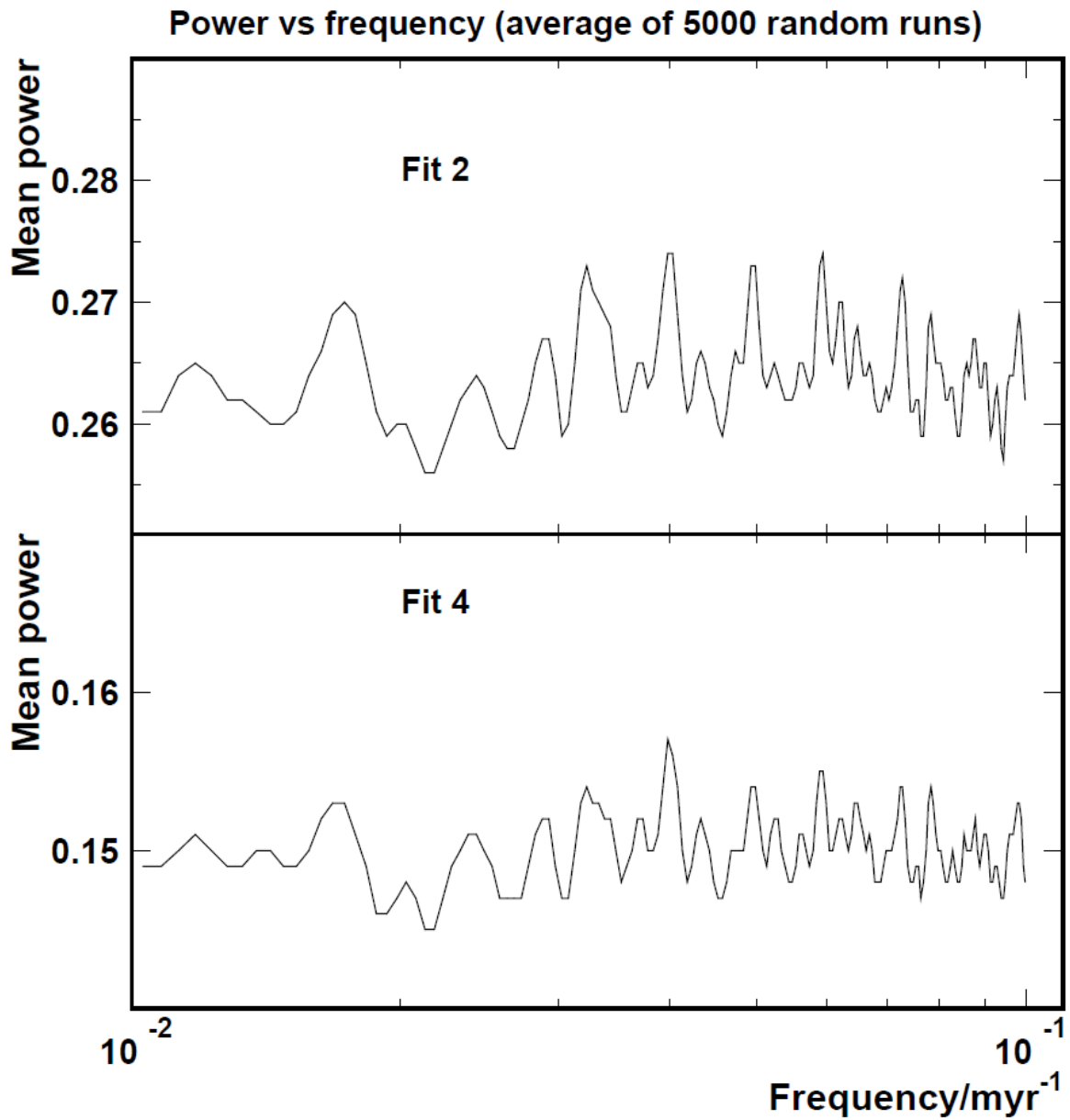
Figure 4 Values of the noise power ($\bar{P}$), averaged over 5000 random runs, against frequency for two of the semi-independent data sets described in table 1. (Note the suppressed zeros and sensitive scales).
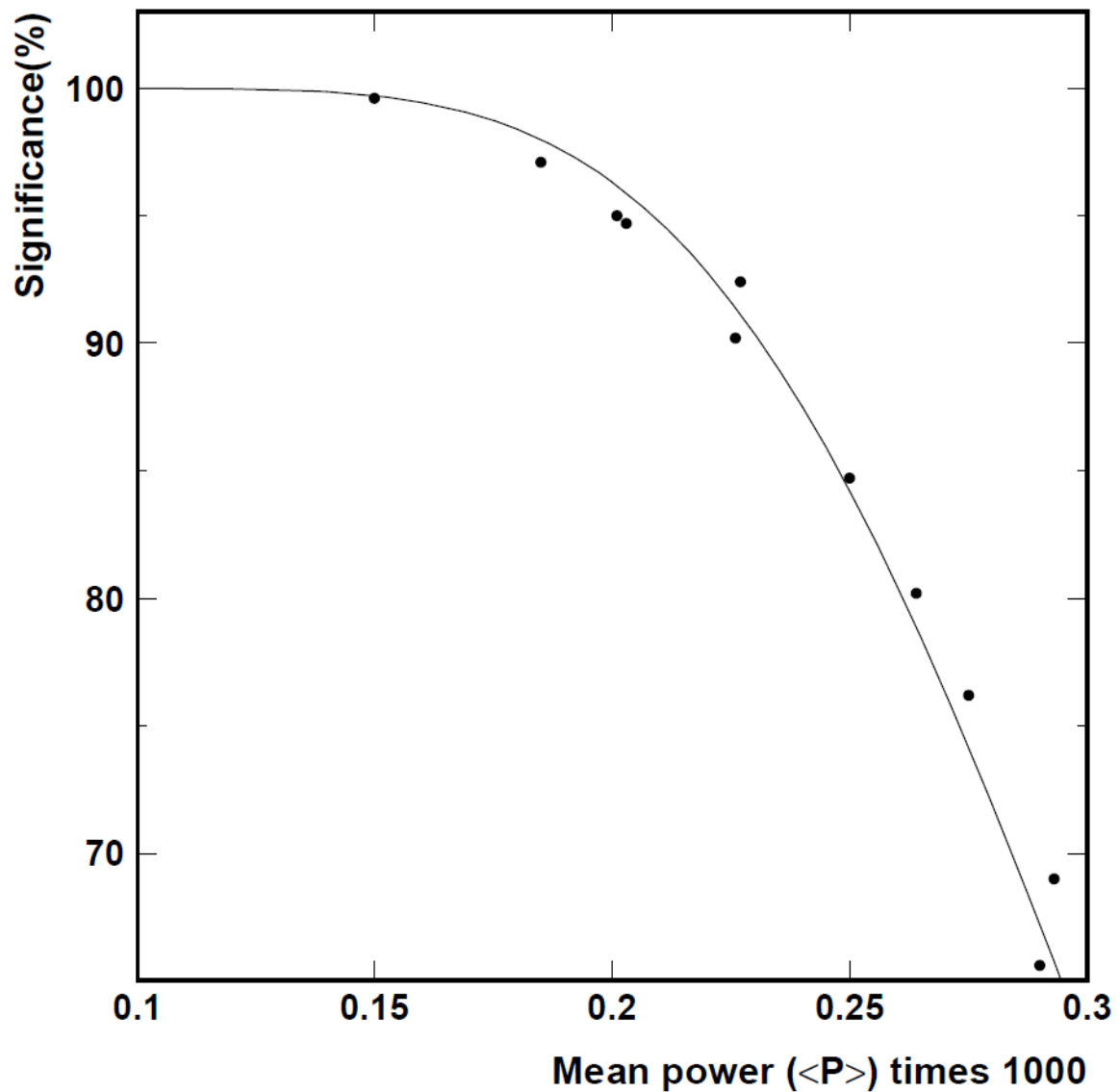
Figure 5 The data points show the significance level for the semi-independent data sets (final column of table 1) versus the mean noise power $<\bar{P}>$ for each (column 7 of table1). The smooth curve shows the calculated variation of $1 - S$ (S defined in equation A10) with $< \bar{P} >$ with $P_{pk} = 0.00151$, the peakpower observed at period 27 myr, and the value of the number of freqencies sampled M arbitrarily chosen to be 80 (see text).
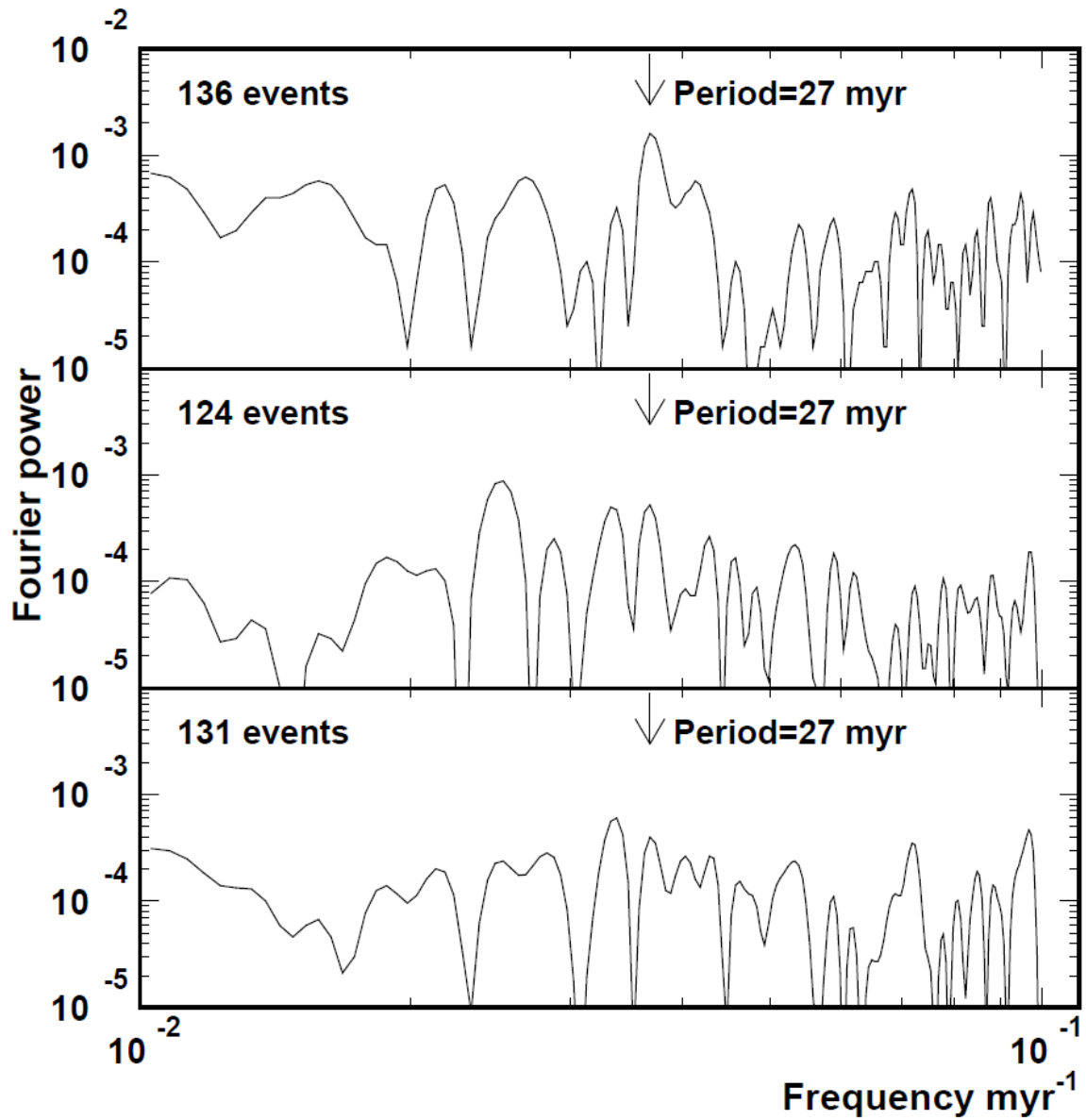
Figure 6 Periodogram of the total sample of 136 events (upper plot) to compare with that of the sample of 124 events with extinction proportion less than 0.3 (middle panel). The lower panel shows the periodogram from 131 events where the five events discussed in the text have been excluded.