Astron. Astrophys. 353, 409–418 (2000)

ASTRONOMY
AND
ASTROPHYSICS

Spurious periods in the terrestrial impact crater record

L. Jetsu[1,2] and J. Pelt[3,4]

[1] University of Helsinki, Observatory, P.O. Box 14, 00014 Helsinki, Finland (jetsu@gstar.astro.helsinki.fi)
[2] NORDITA, Blegdamsvej 17, 2100 Copenhagen, Denmark
[3] Tartu Observatory, 61 602 Toravere, Estonia (pelt@jupiter.aai.ee)
[4] Centre for Advanced Study, Drammensveien 78, 0271 Oslo, Norway

Received 8 September 1999 / Accepted 21 October 1999

Abstract. We present a simple solution to the controversy over periodicity in the ages of terrestrial impact craters and the epochs of mass extinctions of species. The first evidence for a 28.4 million year cycle in catastrophic impacts on Earth was presented in 1984. Our re-examination of this earlier Fourier power spectrum analysis reveals that the rounding of the impact crater data distorted the Monte Carlo significance estimates obtained for this cycle. This conclusion is confirmed by theoretical significance estimates with the Fourier analysis, as well as by both theoretical and Monte Carlo significance estimates with the Rayleigh method. We also apply other time series analysis methods to six subsamples of the currently available more extensive impact crater record and one sample of mass extinction epochs. This analysis reveals the spurious "human-signal" induced by rounding. We demonstrate how the data rounding interferes with periodicity analysis and enhances artificial periodicities between 10 and 100 million years. Only integer periodicities connected to irregular multimodal phase distributions reach a significance of 0.001 or 0.01. We detect no real periodicity in the ages of terrestrial impact craters, nor in the epochs of mass extinctions of species.

Key words: Earth – comets: general – minor planets, asteroids – methods: statistical – Galaxy: solar neighbourhood

1. Introduction

The interesting history of scientific or other conjectures of *how* the end of the world might arrive has been reviewed by Clube & Napier (1996). The idea by Urey (1973) that cometary collisions with the Earth could terminate geological eras was originally published in a magazine called the Saturday Review of Literature. The first physical evidence for a connection between a comet or asteroid impact and a mass extinction of species was discovered by Alvarez et al. (1980), who detected the extraterrestrial Iridium anomaly at the Cretaceous-Tertiary boundary. The Chicxulub crater in Mexico with an age of 65 Myr ($\mathrm{Myr} \equiv 10^6 \mathrm{yr}$) has been widely accepted as the site of this catastrophic impact (e.g. Pope et al. 1997).

A 26 Myr periodicity in eight epochs of major mass extinctions of species over the past 250 Myr was detected by Raup & Sepkoski (1984). A comparable cycle of 28.4 Myr was found by Alvarez & Muller (1984) in the ages of eleven larger terrestrial impact craters. Furthermore, the cycle phase of this periodicity coincided with that of the mass extinctions of species. Astronomical "clocks" capable of triggering comet showers from the Oort Cloud at regular intervals were soon invented: an unseen solar companion (Davis et al. 1984; Whitmire & Jackson 1984) or the oscillation of the Solar System in the galactic plane (Rampino & Stothers 1984; Schwartz & James 1984). The periodicity in these catastrophies has remained a controversial subject for over a decade (e.g. Clube & Napier 1996; Grieve & Pesonen 1996; Harris 1996; Matsumoto & Kubotani 1996; Rampino & Haggerty 1996; Yabushita 1996; Leitch & Vasisht 1997; Napier 1997; Stothers 1998). A highly significant "human-signal" was recently discovered in the terrestrial impact crater record by Jetsu (1997, hereafter Paper I). Were this detection interpreted as those of *any* earlier studies on the subject, comets and asteroids would bombard Earth at exact intervals like 3, 5, 10, 15 or 20 Myr, and with a high statistical certainty. The rounding of impact crater ages induces these spurious periodicities.

As already noted in Paper I, here we concentrate on the influences of the "human-signal" on the statistics of time series analysis. We present a more detailed analysis of the data from Paper I. Those eight different data samples are described in Sect. 2. The applied methods are summarized in Sect. 3, while the formulation of every method is thoroughly explained in a separate appendix. In the next two sections we demonstrate how rounding distorts the statistics of time series analysis, and enhances spurious periodicities. We confirm that Alvarez & Muller (1984) neglected these bias effects in the first periodicity detection in the ages of terrestrial impact craters. In Sect. 4 we repeat their Fourier power spectrum analysis and then apply the Rayleigh method analysis to their sample. In Sect. 5 we describe the sensitivity of different methods in detecting periodicities connected to different types of phase distributions. Our conclusions are given in Sect. 6.

Send offprint requests to: L. Jetsu

2. Data

Astronomical data frequently contain periodic gaps, because the objects can be observed a few minutes earlier each night and some only during certain seasons. Thus the windows of a sidereal day ($0.^{\rm d}9972$) and year ($365.^{\rm d}2564$) interfere with periodicity detection in ground-based astronomical observations. An analogy arises when a series of time points is rounded, like the ages of terrestrial impact craters. Depending on the achieved accuracy, the age of some arbitrary crater may be rounded, e.g. from 66.7 Myr to 67 ± 1, 65 ± 5 or $70 \pm 10\mathrm{Myr}$.

Our data from Table 1 of Paper I are $n = 74$ craters, or crater pairs, with an age $t_{\rm i} \pm \sigma_{t_{\rm i}}$ [Myr] and a diameter $D_{\rm i}$ [km]. These 82 impact structures were chosen from the database constantly updated by the Geological Survey of Canada[1]. The selected values represent those available in October 1996, except that the $t_{\rm i}$ values with no $\sigma_{t_{\rm i}}$ estimate have been discarded. Hence we do not analyse ambiguous data with no error estimates, nor craters with only an upper or lower limit for their age. The remaining data of 82 craters contained eight pairs of simultaneous events. The phases for the two events of each pair would be equal for any periodicity. Because such pairs would deteriorate the statistics of any period finding method, they were combined in Paper I. The geographical coordinates of these pairs did not rule out the possibility of a double impact, except for the pair of Carswell (Canada) and Zapadnaya (Ukraine) craters with an age of $t_{\rm i} = 115 \pm 7.1$ Myr. Evidence for multiple impacts on Earth was recently presented by Spray et al. (1998), who claimed that a chain of five craters with an age of about 214 Myr was formed by the collision of a fragmented object during the Norian stage of the Triassic period. While several doublet craters have been reliably identified, the existence of crater chains on Earth is still debated (Melosh 1998).

The subsamples

C_1: $5 \le t_{\rm i}$, n=61
C_2: $t_{\rm i} \le 250$, $\sigma_{t_{\rm i}} \le 20$, $D \ge 5$, n=34
C_3: $5 \le t_{\rm i} \le 300$, $\sigma_{t_{\rm i}} \le 20$, n=35
C_4: $t_{\rm i}$ of C_1 that are not multiples of 5, n=27
C_5: $t_{\rm i}$ of C_2 that are not multiples of 5, n=25
C_6: $t_{\rm i}$ of C_3 that are not multiples of 5, n=23.

were selected from Table 1 of Paper I. We shall now shortly explain the reasons for the above selection criteria. About 49% or 72% of all $t_{\rm i}$ values are below 100 Myr or 250 Myr, respectively. Thus these data are strongly biased towards younger craters, erosion and other geological processes having removed or buried the older ones. The thirteen youngest craters have been removed from the first subsample C_1. The selection criteria of C_2 are the same as in the period analysis by Grieve & Pesonen (1996), while those of C_3 were applied by Matsumoto & Kubotani (1996). The reasons for selecting only the most accurate crater ages are obvious. As for the crater diameter criterion, e.g. Yabushita (1991) has detected a periodicity of 16.5 Myr in small impacts on Earth. The diameter criterion also recognizes the expected higher impact velocities for comets than asteroids. A study of the orbital and size distributions of Near Earth Objects (NEO) by Shoemaker et al. (1990) indicated that most of the craters with $D_{\rm i} > 50$ km have been caused by cometary impacts, while those with $D_{\rm i} < 30$ km mainly by asteroids. The subsamples C_4, C_5 and C_6 were selected, because many ages in C_1, C_2 and C_3 have been rounded into multiples of 5 Myr.

Like in Paper I, two additional samples

C_7: Alvarez & Muller (1984), n=11
C_8: Matsumoto & Kubotani (1996), n=8

are also studied here. The 28.4 Myr cycle was detected by Alvarez & Muller (1984, their Table 1) in the seventh sample (C_7). This first periodicity detection in the terrestrial impact crater record will be thoroughly re-examined in our Sect. 4. The last sample (C_8) consists of the eight mass extinction epochs analysed by Matsumoto & Kubotani (1996, their Table 1). These extinction events of marine animal families during the past 250 Myr were considered significant by Raup & Sepkoski (1984, 1986).

There are certainly several bias effects that cannot be simply eliminated from the C_1, C_2,..., C_7 samples of crater ages. For example, the number of craters in each age group should be corrected for the surface area available at the time of impacts, the corresponding surface area currently available for sampling, and the past geological processes in these areas that modify the detection probability of craters. In other words, some geological eras are better presented than others. Even the selection criteria for the samples C_2 or C_3, which were already applied in earlier studies (Grieve & Pesonen 1996; Matsumoto & Kubotani 1996), cannot account for this type of complicated bias effects. Furthermore, the crater ages and their error estimates are constantly revised, e.g. like in the case of C_7 discussed in Sect. 4.

3. Methods

For any period P, the phases of $t_{\rm i}$ in C_1,..., C_8 are $\phi_{\rm i} = \mathrm{FRAC}[(t_{\rm i} - t_0)P^{-1}]$, where FRAC removes the integer part of $(t_{\rm i} - t_0)P^{-1}$ and t_0 is the zero phase epoch. These data represent one particular type of *circular data*, because the phases can be projected onto the circumference of a circle, e.g. by using the direction angles $\theta_{\rm i} = 2\pi\phi_{\rm i}[\mathrm{rad}]$.

For example, Batschelet (1981) describes eleven methods to analyse circular data. Surprisingly, the impact crater or mass extinction data have not been studied with these standard methods. Only techniques sensitive to unimodal phase distributions have been applied, like the Fourier methods (e.g. Alvarez & Muller 1984; Rampino & Haggerty 1996) or modifications of the Broadbent (1955, 1956) method (e.g. Raup & Sepkoski 1984; Yabushita 1991; Grieve & Pesonen 1996). Another striking similarity between these earlier studies is that they have relied on Monte Carlo significance estimates, instead of theoretical ones.

Here we shall search for periodicity with the Rayleigh (uni- and bimodal versions), Kuiper (1960), Scargle (1982) and Swanepoel & De Beer (1990) methods. Note that the Scargle

[1] http://gdcinfo.agg.emr.ca/toc.html?/crater/

Table 1. The two best periods (P) between 10 and 100 Myr detected with the SD-, K-, R1- and R2-methods in C_1, ..., C_8, and their critical levels for m independent frequencies (Eqs. A.2, A.3, A.4 and A.5). The phase distributions in Figs. 4a–u are the cases, where H_0 is rejected with $\gamma = 0.001$ ($\star$), 0.01 (†) or 0.1 (‡).

Sample	SD-method		K-method		R1-method		R2-method	
	$P(Q_{\mathrm{SD}})$	Fig.	$P(Q_{\mathrm{K}})$	Fig.	$P(Q_{\mathrm{R1}})$	Fig.	$P(Q_{\mathrm{R2}})$	Fig.
C_1	20.00 (0)	4a*	10.00 ($4.6\cdot10^{-5}$)	4b*	10.49 (0.82)		10.00 (0.0037)	4b†
($m = 198$)	10.00 (0)	4b*	15.00 (0.0017)	4c†	17.93 (0.99)		10.97 (0.54)	
C_2	10.00 ($2.0\cdot10^{-9}$)	4d*	11.67 (0.17)		17.60 (0.24)		35.21 (0.24)	
($m = 24$)	14.02 (0.012)	4e‡	17.50 (0.19)		11.65 (0.29)		23.30 (0.29)	
C_3	10.00 ($\ll 10^{-20}$)	4f*	11.62 (0.30)		10.59 (0.46)		21.16 (0.46)	
($m = 28$)	14.97 (0.038)	4g‡	10.64 (0.37)		11.60 (0.47)		23.18 (0.47)	
C_4	10.00 ($6.8\cdot10^{-7}$)	4h*	13.20 (0.32)		13.30 (0.23)		26.60 (0.23)	
($m = 198$)	11.00 ($1.2\cdot10^{-6}$)	4i*	14.06 (0.56)		13.49 (0.47)		10.96 (0.44)	
C_5	14.02 (0.015)	4j‡	17.43 (0.077)	4k‡	17.54 (0.17)		14.12 (0.12)	
($m = 24$)	70.11 (0.018)	4l‡	14.12 (0.15)		10.61 (0.53)		35.06 (0.17)	
C_6	14.01 (0.017)	4m‡	14.00 (0.14)		13.42 (0.14)		26.84 (0.14)	
($m = 23$)	10.88 (0.056)	4n‡	10.45 (0.55)		10.56 (0.54)		10.92 (0.28)	
C_7	10.00 (0)	4o*	28.75 (0.015)	4p‡	28.54 (0.034)	4q‡	57.08 (0.034)	4r‡
($m = 19$)			12.09 (0.092)	4s‡	21.13 (0.36)		12.15 (0.24)	
C_8			11.03 (0.079)	4t‡	19.55 (0.47)		39.10 (0.45)	
($m = 23$)			13.80 (0.10)	4u‡	29.40 (0.48)		58.80 (0.49)	

(1982) method is applied only to C_7 in Sect. 4. The formulation of all these methods is thoroughly described in our separate appendix. The abbreviations for each method and their sensitivity in detecting different ϕ_{i} distributions are summarized in Table A.1. The notations and the method abbreviations from this table will be hereafter used throughout this paper. Instead of Monte Carlo simulations, we determine the theoretical significance estimates for the detected periodicities (Eqs. A.2, A.3, A.4 and A.5). The K- and SD-methods are also sensitive to the multimodal phase distributions connected to the "human-signal", which is enhanced by the rounding of the data. The presence of this bias could not be detected with methods sensitive only to unimodal phase distributions, like the R1- or LS-methods. These sensitivity effects will be illustrated in Sect. 5.

The analysis in Paper I was performed between $P_{\min} = 2.2\mathrm{Myr}$ and $P_{\min} = 250\mathrm{Myr}$. In the current paper, the SD-, K-, R1- and R2-methods are applied to search for periodicity between $P_{\min} = 10\mathrm{Myr}$ and $P_{\max} = 100\mathrm{Myr}$. The detection of the smallest "human-signal" periodicities is avoided by increasing $P_{\min}$ from 2.2 to 10 Myr. The $P_{\max}$ reduction from 200 Myr to 100 Myr is made, because the former value is comparable to $t_{\max} - t_{\min}$ of C_2, C_3, C_5, C_6, C_7 and C_8. Detection of real periodicity below 10 or above 100 Myr is difficult, because these data are rounded and heavily biased towards younger craters.

Some models predict comets showers within the selected range of $10\mathrm{Myr} \leq P \leq 100\mathrm{Myr}$, like the vertical oscillation of the Solar System in the galactic plane (e.g. Bahcall & Bahcall 1985; Clube & Napier 1996; Stothers 1998).

Our statistical point of view is to test the "null hypothesis" (H_0): "The phases ϕ_{i} with an arbitrary period P are randomly distributed between 0 and 1." This H_0 is rejected at the preassigned significance level of $\gamma = 0.001$, as explained in greater detail in our appendix. The two best periods detected in subsamples C_1, C_2,..., and C_8 are given in Table 1. Although it is a convention to fix γ *before* the test, the less significant periodicities that *would* be detected with $\gamma = 0.01$ or 0.1 are also specified in Table 1. In other words, one should not reject H_0 with these less significant periodicities. Note that the "human-signal" cannot be eliminated by excluding the t_{i} that have been rounded into multiples of 5 Myr, because integer periodicities like 10, 11 or 14 Myr are also present in subsamples C_4 or C_6.

4. Re-examination of the first periodicity detection

We decided to repeat the earlier Fourier power spectrum analysis by Alvarez & Muller (1984), because the 28.4 Myr cycle detection in C_7 undoubtedly motivated the subsequent studies on the subject. That earlier analysis began by replacing the time series $t_{\mathrm{i}} \pm \sigma_{t_{\mathrm{i}}}$ with gaussian distributions. These unity area gaussians with a standard deviation of $\sigma_{t_{\mathrm{i}}}$ were centered at t_{i}. An interpolated event-rate function was then constructed by superimposing these gaussian distributions, as depicted in their Fig. 1b. This event-rate function is denoted here by $G_\Sigma(t_{\mathrm{i}})$.

The Fourier power spectrum is also called the "classical periodogram". Several refinements of this method were introduced by Scargle (1982), who defined the "classical" and "revised" periodograms in his Eqs. 3 and 10, respectively. Because the event-rate function $G_\Sigma(t_{\mathrm{i}})$ is evenly spaced in time, all our results for the Monte Carlo simulations presented in this section were the same for the "classical" and "revised" periodograms. Hence we present the results for the "revised" periodogram z_{LS} (Eq. A.6), which allowed us to apply the theoretical significance

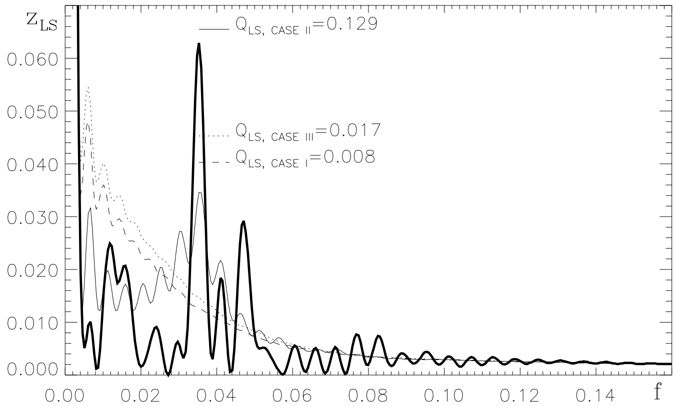

Fig. 1. The thick continuous line displays the LS-method periodogram $z_{\rm LS}$ of $\rm C_7$. The thinner dashed, continuous, and dotted lines show the mean periodograms $\langle z_{\rm LS} \rangle$ of CASE I, CASE II and CASE III. The respective short horizontal lines indicate the level that the $z_{\rm LS}$ peak at $f = 0.035\,{\rm Myr}^{-1}$ ($P \approx 28.4\,{\rm Myr}$) should exceed in order to reach the $\gamma = 0.1$ significance in CASE I, CASE II and CASE III. The simulated critical levels of this peak are $Q_{\rm LS,CASE\ I}$, $Q_{\rm LS,CASE\ II}$ and $Q_{\rm LS,CASE\ III}$. Note that the $z_{\rm LS}$, the CASE I simulation for $\langle z_{\rm LS} \rangle$ and the $Q_{\rm LS,CASE\ I}$ estimate are as in Alvarez & Muller (1984).

estimates of Eq. A.7. This particular Fourier method based on the "revised" periodogram, i.e. the LS-method, is also known as the Lomb–Scargle periodogram. Our Fig. 1 shows the LS-method periodogram $z_{\rm LS}$ for $G_\Sigma(t_{\rm i})$. Like Alvarez & Muller (1984), we generated 5000 random samples:

CASE I: $n = 11$ random $t^*_{\rm i}$ between 5 Myr and 250 Myr.

The $\sigma_{t_{\rm i}}$ of $\rm C_7$ were assigned to these $t^*_{\rm i}$ in random order. The $z_{\rm LS}$ were determined for each $G_\Sigma(t^*_{\rm i})$, and the mean $\langle z_{\rm LS} \rangle$ of these 5000 periodograms was obtained. Instead of 1000 random samples (Alvarez & Muller 1984), we simulated 5000 to ascertain the stability of our critical level estimates $Q_{\rm LS,CASE\ I}$. The significance of the peak at 28.4 Myr was determined by checking how frequently "such peaks occurred for frequencies equal to or above $0.035\,{\rm Myr}^{-1}$". The result was $Q_{\rm LS,CASE\ I} = 40/5000 = 0.008$. Thus our CASE I analysis with the LS-method confirmed this result by Alvarez & Muller (1984).

One must emphasize that the above analysis was restricted *only* to $f \geq 0.035\,{\rm Myr}^{-1}$. Perhaps Alvarez & Muller (1984) chose this limit because $\langle z_{\rm LS} \rangle$ is monotonously decreasing in f, which complicates significance estimates for $f < 0.035\,{\rm Myr}^{-1}$. But a theoretical estimate can be obtained by transforming the $z_{\rm LS}$ periodogram into "noise units" (Scargle 1982). There are $m = 19$ independent frequencies when searching for periodicity between $f = 0.01{\rm Myr}^{-1}$ and $0.10\,{\rm Myr}^{-1}$ (Table 1: Eq. A.1 for $\rm C_7$). Hence the theoretical critical level estimate for the 28.4 Myr peak with $z_{\rm LS}/\langle z_{\rm LS} \rangle$=4.91 is $Q_{\rm LS} = 0.13$ (Eq. A.7). The LS-method analysis being "exactly equivalent to least squares fitting of sinusoids to the data" (Scargle 1982), this $Q_{\rm LS}$ estimate implies that a sinusoid is an inappropriate model for the event-rate function $G_\Sigma(t_{\rm i})$. Furthermore, this function represents data interpolation *before* analysis, and, e.g., Raup & Sepkoski (1986) and Bai (1992) have warned against such interpolations.

All $N = (n-1)n/2$ time difference pairs $t_{\rm i,j} = t_{\rm j} - t_{\rm i}$ in $\rm C_7$ were also studied by Alvarez & Muller (1984). Their $t_{\rm i,j}$ errors were $\sigma^2_{t_{\rm i,j}} = \sigma^2_{t_{\rm i}} + \sigma^2_{t_{\rm j}}$. To display the 28.4 Myr cycle, these $t_{\rm i,j} \pm \sigma_{t_{\rm i,j}}$ were replaced by gaussian distributions that were then superimposed in their Fig. 3. Because the $n-1$ time differences $t_{\rm i,i+1} = t_{\rm i+1} - t_{\rm i}$ determine all N time differences $t_{\rm i,j}$ uniquely, we decided to generate 5000 random samples:

CASE II: $t_{\rm i,i+1}$ in random order determine a new $t^*_{\rm i}$ sample.

The LS-method Monte Carlo analysis for $f \geq 0.035{\rm Myr}^{-1}$ was then repeated as in CASE I. This randomization scheme gave $Q_{\rm LS,CASE\ II} = 0.129$ for the 28.4 Myr cycle. But a $\langle z_{\rm LS} \rangle$ peak also appeared at $f \approx 0.035\,{\rm Myr}^{-1}$ (Fig. 1). This result is trivial. Just like $G_\Sigma(t_{\rm i})$, the $G_\Sigma(t^*_{\rm i})$ of CASE II always have eight peaks or less, while $t^*_{\rm max} - t^*_{\rm min}$ remains constant. Eight peaks are equal to seven full 27.9 Myr cycles between $t^*_{\rm min} = 14.8$ Myr and $t^*_{\rm max} = 210$ Myr. If the $G_\Sigma(t^*_{\rm i})$ of some random sample has *less* than eight peaks, the best modelling frequency will be below $0.035\,{\rm Myr}^{-1}$. But such f were not tested, nor detected! The theoretical critical level estimate for 28.4 Myr with $z_{\rm LS}/\langle z_{\rm LS} \rangle$=1.82 was $Q_{\rm LS} = 0.96$ (Eq. A.7).

Rounding was not tested in CASE I and CASE II simulations. An effect resembling the "human-signal" was created by rounding CASE I to imitate $\rm C_7$. The following 5000 random samples were generated:

CASE III: $t^*_{\rm i}$ as in CASE I, but four values rounded with 1, and six with 5 (and two of these made equal).

The LS-method Monte Carlo analysis for $f \geq 0.035{\rm Myr}^{-1}$ gave $Q_{\rm LS,CASE\ III} = 0.017$ for the $P = 28.4$ Myr cycle. Rounding decreased the significance of this cycle, because the $\langle z_{\rm LS} \rangle$ level increased at low f (Fig. 1). The above cycle had $z_{\rm LS}/\langle z_{\rm LS} \rangle$=4.31, which gave a theoretical critical level of $Q_{\rm LS} = 0.23$ (Eq. A.7).

The Monte Carlo simulations of CASE I, CASE II and CASE III for the R1-method (Fig. 2) differed slightly from those for the LS-method (Fig. 1). The R1-method does not utilize $\sigma_{t_{\rm i}}$, nor the questionable data interpolation. The $f \geq 0.035{\rm Myr}^{-1}$ restriction was unnecessary, because the CASE I mean periodogram fulfills $\langle z_{\rm R1} \rangle \approx 1$. Thus the R1-method simulations could be performed over $0.01{\rm Myr}^{-1} \leq f \leq 0.10{\rm Myr}^{-1}$. In other words, the simulations revealed how frequently a $z_{\rm R1}$ peak of a given height occurred within the above frequency interval. The respective theoretical and CASE I Monte Carlo simulation estimates of the critical level for the 28.5 Myr peak were $Q_{\rm R1} = 0.034$ (Table 1: Eq. A.4) and $Q_{\rm R1,CASE\ I} = 0.045$ (Fig. 2). This result confirmed that our $m = 19$ estimate for the number of independent frequencies is correct. In these R1-method simulations the "human-signal"($Q_{\rm R1,CASE\ III} = 0.111$) had a stronger influence on the critical level than the randomization of $t_{\rm i,i+1}$ ($Q_{\rm R1,CASE\ II} = 0.082$).

The main idea in the study by Alvarez & Muller (1984) was to search for the best sinusoid to fit the eight $G_\Sigma(t_{\rm i})$ peaks in their Fig. 1b. But a sinusoid is a poor model for $G_\Sigma(t_{\rm i})$ at high

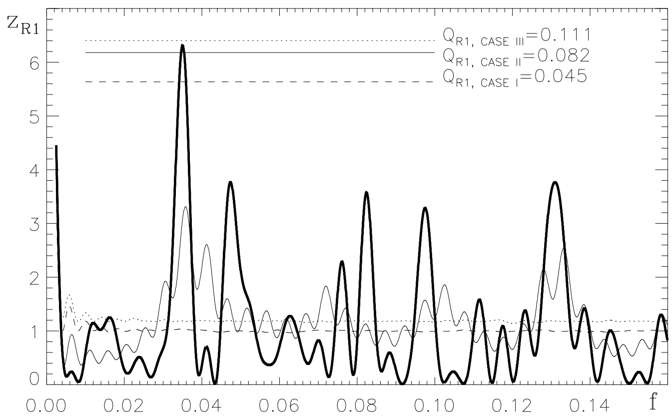

Fig. 2. The thick continuous line is the R1-method periodogram $z_{\rm R1}$ of $\textsc{c}_7$. The notations for the mean periodograms $\langle z_{\rm R1}\rangle$ of CASE I, CASE II and CASE III are as in Fig. 1. The respective horizontal lines outline the $\gamma = 0.1$ significance level between 0.01 and $0.10 \mathrm{Myr}^{-1}$. The simulated critical levels for the peak at $f = 0.035\,\mathrm{Myr}^{-1}$ ($P = 28.5\,\mathrm{Myr}$) are $Q_{\rm R1,CASE\ I}$, $Q_{\rm R1,CASE\ II}$ and $Q_{\rm R1,CASE\ III}$.

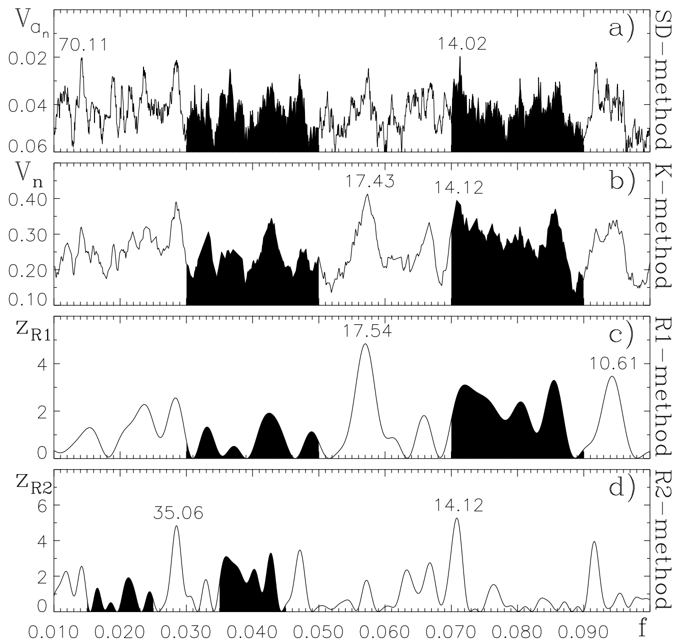

Fig. 3a–d. The $V_{\rm a_n}(f)$, $V_{\rm n}(f)$, $z_{\rm R1}(f)$ and $z_{\rm R2}(f)$ periodograms of $\textsc{c}_5$ between $f = 0.01\,\mathrm{Myr}^{-1}$ and $0.1\,\mathrm{Myr}^{-1}$. The two best P are indicated above each periodogram. To allow for easier comparison between these periodograms, frequency intervals of 0.02 are shaded in **a**–**c** and their counterparts at $f/2$ in **d**. Note that the $V_{\rm a_n}(f)$ scale has been reversed in **a**, because the best periodicity candidates minimize this test statistic, while the best periodicities for the other methods maximize $V_{\rm n}(f)$, $z_{\rm R1}(f)$ and $z_{\rm R2}(f)$.

f (i.e. small P), where the similarities of the periodograms of Figs. 1 and 2 disappear. The R1-method periodogram indicates several periodicity candidates at higher f, while that of the LS-method does not. The latter method relies on data interpolation and assumes a sinusoidal model. The obvious reason for the featureless $z_{\rm LS}$ periodogram at higher f is that when the distances between the peaks of some $G_\Sigma(t_{\rm i}^*)$ are small, then $G_\Sigma(t_{\rm i}^*) \approx 0$ elsewhere. Hence a model sinusoid has large residuals and $z_{\rm LS}$ approaches zero.

As already noted in Paper I, the scatter of $\sigma_{t_{\rm i}}$ in $\textsc{c}_7$ is so large that the statistics of the weighted K- and SD-method versions become unreliable. This problem could be anticipated when there are only eleven $t_{\rm i}$ values, and the sum of two largest weights is 9.5, the sum for the other nine being 1.5. It is therefore logical to argue that the LS-method statistics suffer from the very same problem. Furthermore, although the small size of $\textsc{c}_7$ prevents an SD-method analysis (see Paper I), these data contain too many $t_{\rm i,j} = 10\,\mathrm{Myr}$ to "qualify" as a random sample. The critical level for this 10 Myr periodicity is $Q_{\rm SD} = 0$, i.e. never under H_0!

We decided to reanalyse $\textsc{c}_7$, although comparison with Table 1 in Paper I revealed that nine $t_{\rm i}$ values out of eleven have been revised since the 28.4 Myr cycle detection. Four revisions exceed $2\sigma_{t_{\rm i}}$ given in Alvarez & Muller (1984). Had their $\sigma_{t_{\rm i}}$ estimates been reliable, one $2\sigma_{t_{\rm i}}$ change might have happened, but not the largest ones of $5.2\sigma_{t_{\rm i}}$ and $3\sigma_{t_{\rm i}}$.

The main results of our re-examination of $\textsc{c}_7$ are: (1) None of the theoretical LS-method critical level estimates for the 28.4 Myr cycle was even close to that of 0.008 determined by Alvarez & Muller (1984). (2) For any period finding method, the significance estimates depend on the chosen Monte Carlo scheme. (3) Rounding deteriorates these Monte Carlo schemes, e.g. CASE I Monte Carlo significance estimates are simply invalid for rounded data. (4) CASE I, CASE II and CASE III are all randomization schemes. But if the data are rounded into multiples of fixed units (e.g. 1, 5 or 10 Myr), then the interference of these "unit" frequencies introduces spurious periodicities. Furthermore, if real periodicity is present, the interference with these "unit" periodicities enhances additional spurious periodicities. Both processes increase the periodogram fluctuations. The former process explains the higher mean level of the periodograms determined for the random samples $t_{\rm i}^*$ of CASE II and CASE III (Figs. 1 and 2). The end of our appendix describes this interference.

5. Detectability of different types of phase distributions

This section should clarify why some methods succeed in detecting the "human-signal", while others fail. For example, the methods sensitive only to unimodal phase distributions certainly fail. The R1-method applied in this paper is one such method, as also are the Fourier methods or the modifications of the Broadbent (1955, 1956) method, which were applied in numerous studies before Paper I. The best periodicities detected with different methods usually differ for the same data, like in our Table 1. These results merely confirm that a "universal" method equally sensitive to all types of $\phi_{\rm i}$ distributions does not exist. Table A.1 of our appendix summarizes these sensitivity effects for all methods applied in this paper. We chose the periodograms of $\textsc{c}_5$ to illustrate the sensitivity of each method in detecting different types of phase distributions (Fig. 3). Regardless of these sensitivity effects, the $V_{\rm a_n}(f)$, $V_{\rm n}(f)$ and $z_{\rm R1}(f)$ periodograms of $\textsc{c}_5$ in Figs. 3a, b and c share the same overall shape. Furthermore, comparison between Figs. 3c and d shows that $z_{\rm R1}(f) \approx z_{\rm R2}(f/2)$, because unimodal phase distributions become bimodal at $f/2$. For example, the $\sim 14 \rm Myr$ periodicity is connected to an approximately bimodal $\phi_{\rm i}$ distribution (Fig. 4j). This periodicity is among the two best ones detected in $\rm C_5$ with the SD-, K- and R2-methods (Table 1). All differences between these periodograms can be explained by sensitivity effects.

The $V_{\rm a_n}(f)$ periodogram in Fig. 3a appears "noisy". This test statistic depends on a small fraction of $N = n(n-1)/2$ phase differences $\phi_{\rm i,j} = \mid \phi_{\rm i} - \phi_{\rm j} \mid$, i.e. small f changes cause large $V_{\rm a_n}(f)$ changes. But the $z_{\rm R1}(f)$, $z_{\rm R2}(f)$ and $V_{\rm n}(f)$ periodograms appear "smoother", because these depend on all $\phi_{\rm i}$ (see our appendix). The $V_{\rm a_n}$ periodogram fluctuations in Fig. 3a are not "noise", but an indication of that the SD-method can detect regular and irregular uni- or multimodal $\phi_{\rm i}$ distributions. This method was designed to search for periodicity in sources exhibiting short duration bursts at unknown intervals, like the γ-ray events of the Vela pulsar studied in Swanepoel & De Beer (1990). It turned out that the multimodal phase distributions connected to the rounded $t_{\rm i}$ in the impact crater record represent an analogy of such bursts. That the "human-signal" remained undetected for over a decade was due to favouring methods sensitive only to unimodal $\phi_{\rm i}$ distributions. These rounding induced periodicities are connected to multimodal phase distributions, as displayed by the $\phi_{\rm i}$ distributions reaching $\gamma = 0.001$: all detected by the SD-method (Figs. 4a, b, d, f, h, i and o). All periodicities to reject H_0 with $\gamma = 0.001$ are actually integers, as well as for $\gamma = 0.01$ (Table 1).

As for the K-method, it is a generalization of the Kolmogorov-Smirnov test in the phase domain, capable of detecting irregular and regular uni- or multimodal $\phi_{\rm i}$ distributions (Figs. 4b, c, k, p, s, t and u). The R1-method was already compared to the LS-method in Sect. 4. It detects only one unimodal phase distribution reaching $\gamma = 0.1$ (Fig. 4q). The R2-method detects the bimodal counterpart of this distribution (Fig. 4r).

6. Conclusions

We detect no real periodicity in the data of terrestrial impact crater ages or the epochs of mass extinctions of species. The highly significant "human-signal" represents spurious periodicity enhanced by rounding. The phase distributions connected to the "human-signal" are *not* uni- or bimodal, but simply irregular. No significant unimodal $\phi_{\rm i}$ distributions were detected. This contradicts the models of periodic comet showers triggered by an unseen solar companion or the oscillation of the Solar System in the galactic plane. But the impact crater record may contain a random and periodic component, there being no reasons to suspect that asteroid impacts on Earth are periodic (e.g. Trefil & Raup 1987). For example, resonant phenomena may constantly transport main belt asteroids into Earth crossing orbits (Gladman et al. 1997). The study by Shoemaker et al. (1990) indicated that the $D \geq 50 \rm km$ craters are *mostly* caused by cometary impacts, and that the probability for a 10 km diameter comet colliding with Earth is about three times larger than that for an asteroid of the same size. Such impacts may cause mass extinctions of species. But Table 1 of Paper I contains only 12 craters equal to or greater than $50 \rm km$ in diameter, the age difference between the oldest and youngest one being about 2000 Myr. Three of these may be nonperiodic asteroid impacts. Searching for periodicity in these $n = 12$ rounded times points would be a wasted effort.

Unlike Alvarez & Muller (1984) or Matsumoto & Kubotani (1996), we did not examine the temporal connection between impacts and mass extinctions. The mass extinction epochs of $\rm C_8$ betray no signatures of significant periodicity (Table 1), like that of 26 Myr suggested by Raup & Sepkoski (1984). Indications of the "human-signal" are also present. For example, Table 1 does not list the third best periodicity detected with the K-method, which is $P = 20 \rm Myr$ with $Q_{\rm K} = 0.39$. The reality of the 26 Myr periodicity in the mass extinctions of species was already criticised a decade ago (see e.g. Patterson & Smith 1987; Stigler & Wagner 1987; Raup & Sepkoski 1988). Moreover, the connections between the numerous causes that may lead to these catastrophies resemble "a tangled web" (Erwin 1994). Benton (1995) concluded that five major mass extinctions have occurred during the last 600 Myr, and that those during the past 250 Myr reveal no periodicity. The recent analysis by Bowring et al. (1998) revealed that the most profound one of these "big five" at the Permian-Triassic boundary 251.4 Myr ago may have lasted only $10^4 \rm yr$. They discuss evidence for both direct and indirect volcanism induced effects, but also speculate that this sudden annihilation of from 70 to 90% of all species on Earth may have been caused by a huge comet impact. But even the most violent impacts do not necessarily cause mass extinctions. Two large impacts occurred in quick succession during the late Eocene about 36 Myr ago (Bottomley et al. 1997), and signatures of a comet shower before and after these impacts were recently detected in the He^3 isotope data (Farley et al. 1998). But as emphasized by Stöffler & Claeys (1997) or Kerr (1998), this "combination punch" did *not* cause a mass extinction, the closest one having been dated about two million years later! Thus the 65 Myr old Chicxulub crater at the Cretaceous-Tertiary boundary still remains the only case, where the connection between an impact and a mass extinction can be more or less firmly established.

Our conclusions are: (1) Rounding enhances the spurious "human-signal". (2) The statistics based on CASE I simulations or H_0 are not valid for rounded data. (3) We did not determine the rounding limit that would still allow detection of real periodicity, if present. This problem is connected to $\sigma_{\rm t_i}$ (e.g. Heisler & Tremaine 1989). For example, the last digits of "less" rounded $t_{\rm i}$ would be closer to random numbers, say 10.26 ± 5.36 Myr instead of 10 ± 5 Myr. Statistics based on CASE I simulations or H_0 would be reliable when searching for periodicity in such $t_{\rm i}$. An alternative solution would be to redetermine the terrestrial impact craters ages more precisely. (4) Rounding has undoubtedly interfered with some earlier periodicity detections, especially through erroneous Monte Carlo significance estimates. Some earlier claims may have been spurious, like the 30 and

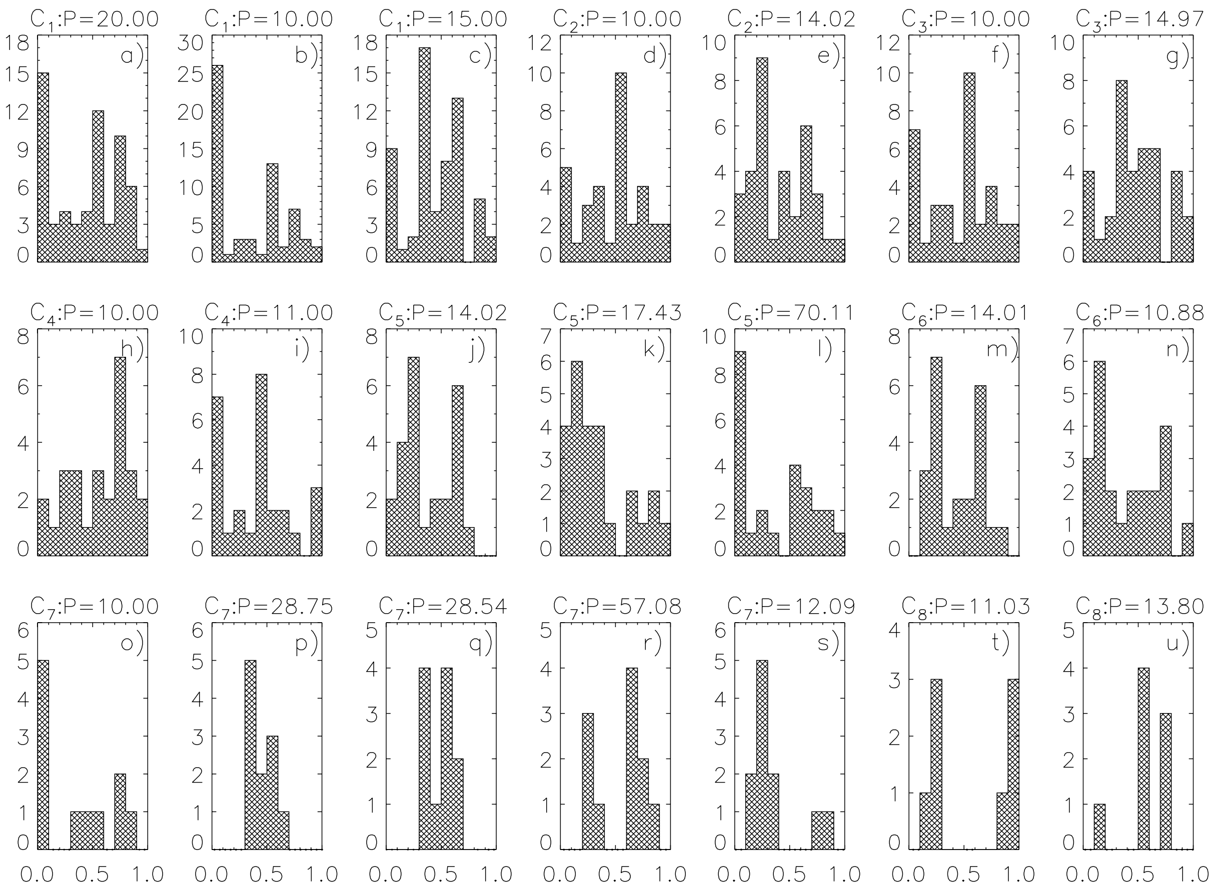

Fig. 4a–u. The ϕ_i distributions of C_1, ..., C_8 for the periods reaching $\gamma \leq 0.1$ in Table 1. The histogram bin width is 0.1 in ϕ.

50 Myr integer cycles in Yabushita (1991). This enhancement of spurious periodicities into rounded data is discussed in greater detail at the end of our appendix (Eq. A.8). (5) Alvarez & Muller (1984) overestimated the significance of the 28.4 Myr cycle in the terrestrial impact crater ages. Nevertheless, their analysis motivated numerous subsequent studies on the subject of periodic catastrophic impacts on Earth, including ours.

Appendix A: formulation of the methods and related statistical topics

This appendix presents a detailed description all methods applied in this paper, i.e. the SD-, K-, R1-, R2- and LS-methods. These abbreviations are explained in Table A.1. The same table also summarizes the sensitivity of these methods in detecting different types of ϕ_i distributions, as well as the test statistic and critical level notations. The weighted versions of the K- and SD-methods from Jetsu & Pelt (1996) were applied in Paper I. The crater diameters were used as weights, because the scatter of the $w_\mathrm{i} = \sigma_{\mathrm{t_i}}^{-2}$ weights was so large that the statistics of these weighted methods became unreliable. Since the "human-signal" was detected with or without weights, we applied here only the nonweighted versions with stable statistics. Our appendix terminates with a description of how rounding enhances spurious periodicities.

The structure of our statistical tests was:

1. Make the "null hypothesis" (H_0): "*The ϕ_i with an arbitrary period P are randomly distributed between 0 and 1.*"
2. Fix the preassigned significance level γ to reject H_0. We selected $\gamma = 0.001$ for a test between $P_\mathrm{min} = 10$ Myr and $P_\mathrm{max} = 100$ Myr.
3. Determine the *test statistic* for each tested P. Under H_0, compute the *critical level* Q, i.e. the probability for the occurrence of this, or an even more extreme value, for the test statistic. The notations for the test statistic and critical level of each method are given in Table A.1.
4. Reject H_0, if and only if $Q \leq \gamma$.

The *periodogram* is the test statistic as a function of tested frequencies $f = P^{-1}$ between $f_\mathrm{min} = P_\mathrm{max}^{-1} = 0.01$ [Myr^{-1}] and $f_\mathrm{max} = P_\mathrm{min}^{-1} = 0.10$ [Myr^{-1}]. The step in tested frequencies is $f_\mathrm{step} = f_0/G$, where $f_0 = (t_\mathrm{max} - t_\mathrm{min})^{-1}$ and the integer $G > 1$ is called an overfilling factor. All integer multiples of f_step between f_min and f_max were tested. The estimate for the number of independent tested frequencies between f_min and f_max was

$$m = \mathrm{INT}[(f_\mathrm{max} - f_\mathrm{min}) f_0^{-1}], \quad \text{(A.1)}$$

where INT removed the fractional part of $(f_\mathrm{max} - f_\mathrm{min}) f_0^{-1}$. If some tested frequency changes from f to $f \pm f_0$, the phase

Table A.1. The methods: abbreviation, sensitivity to different $\phi_{\rm i}$ distributions, test statistic and critical level.

Method	Abbreviation	Sensitivity	Test statistic	Critical level
Swanepoel & Be Beer (1990)	SD-method	uni- and multimodal	$V_{\rm a_n}$	$Q_{\rm SD}$
Kuiper (1960)	K-method	uni- and multimodal	$V_{\rm n}$	$Q_{\rm K}$
Rayleigh	R1-method	unimodal	$z_{\rm R1}$	$Q_{\rm R1}$
Rayleigh	R2-method	bimodal	$z_{\rm R2}$	$Q_{\rm R2}$
Scargle (1982)	LS-method	unimodal (sinusoidal)	$z_{\rm LS}$	$Q_{\rm LS}$

difference during the total time span of a time series changes by $f_0(t_{\max}-t_{\min})=\pm1$. In other words, a change of f_0 in f completely rearranges the phases $\phi_{\rm i}$. Hence the values of any test statistic that depends on $\phi_{\rm i}$ should become uncorrelated within an interval of $\pm f_0/2$. To obtain accurate estimates for the best periods, the periodograms were determined with an overfilling factor of $G=100$. This does not deteriorate the statistics, because performing 100 tests within $\pm f_0/2$ amounts to testing only one independent frequency. These m estimates could be verified with the empirical approach outlined in Jetsu & Pelt (1996, $r(k)$ in their Eq. 14). Furthermore, one m estimate was confirmed with simulations (Fig. 2: CASE I for the R1-method).

A.1. The SD-method

The three stages to determine the SD-method test statistic $V_{\rm a_n}$ are

1. Derive the $N=n(n-1)/2$ phase differences $\phi_{\rm i,j} = |{\rm FRAC}\left[f(t_{\rm i}-t_{\rm j})\right]|$,S where $i=1,...,n-1$ and $j=i+1,...,n$. The absolute value means here that all $\phi_{\rm i,j}$ greater than 0.5 are converted to $1-\phi_{\rm i,j}$.
2. Arrange $\phi_{\rm i,j}$ into ascending (i.e. rank) order and denote them by $V_1,...,V_{\rm N}$, e.g. the i:th smallest one is $V_{\rm i}$.
3. Derive $a_{\rm n} = {\rm INT}\left[N\beta_{\rm n}^{-1}\right]$, where $\beta_{\rm n}=2^{\frac{1}{3}}n^{\frac{2}{3}}$. The SD-method test statistic is $V_{\rm a_n}$, i.e. the $a_{\rm n}$:th smallest value of all $V_{\rm i}$.

It is relatively easy to understand why the SD-method is sensitive to both uni- and multimodal, as well as to regular and irregular, phase distributions. The test statistic $V_{\rm a_n}$ is minimized when the $a_{\rm n}$ smallest phase differences are obtained from phases located on one or many concentrations, while the remaining other $N-a_{\rm n}$ phase differences do not contribute to this test statistic. Even small changes in the tested frequency may cause large changes of $V_{\rm a_n}$, because $a_{\rm n}\ll N$ for larger samples (Jetsu & Pelt 1996, their Eq. 7).

The critical level of the SD-method when testing m independent frequencies is

$$Q_{\rm SD}=1-[U(s)]^{\rm m}, \tag{A.2}$$

where

$$U(s)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{s}{\rm e}^{-t^2/2}dt$$

and $s=0.125-\beta_{\rm n}(\beta_{\rm n}V_{\rm a_n}-0.5)$.

A.2. The K-method

The three stages that determine the K-method test statistic $V_{\rm n}$ are

1. Derive the n phases $\phi_{\rm i}={\rm FRAC}[f(t_{\rm i}-t_0)]$, and arrange them into rank order.
2. Determine the sample distribution function
$$F_{\rm n}(\phi)=\begin{cases}0, & \phi<\phi_1\\ in^{-1}, & \phi_{\rm i}\le\phi<\phi_{\rm i+1},\ 1\le i\le n-1\\ 1, & \phi\ge\phi_{\rm n}.\end{cases}$$
3. Under H_0, the cumulative distribution function is $F(\phi)=\phi$. The K-method test statistic is
$$V_{\rm n}=D^{+}+D^{-},$$
where D^+ and D^- denote the maximum values of $F_{\rm n}(\phi)-F(\phi)$ and $F(\phi)-F_{\rm n}(\phi)$, respectively.

The best periodicity candidates maximize the K-method test statistic $V_{\rm n}$, which is the sum of D^+ and D^-. Every $\phi_{\rm i}$ concentration can contribute to D^+, while phase interval(s) devoid of $\phi_{\rm i}$ contribute to D^-. Hence the K-method is sensitive in detecting both uni- and multimodal phase distributions.

For m independent frequencies, the critical level of the K-method is

$$Q_{\rm K}=1-[1-P(z)]^{\rm m}, \tag{A.3}$$

where $z=n^{1/2}V_{\rm n}$,

$$P(z)=\sum_{\rm k=1}^{\infty}2\left(4k^2z^2-1\right){\rm e}^{-2k^2z^2}\ -$$
$$(8z/3)n^{-1/2}\sum_{\rm k=1}^{\infty}k^2\left(4k^2z^2-3\right){\rm e}^{-2k^2z^2}\ +\ {\rm O}(n^{-1}),$$

and the residual ${\rm O}(n^{-1})$ is negligible for larger samples ($n\ge\sim 20$). An exact solution for smaller samples can be found in Stephens (1965).

A.3. The R1- and R2-methods

The R1-method test statistic is

$$z_{\rm R1}=n^{-1}\left[\left(\sum_{\rm i=1}^{\rm n}\cos\theta_{\rm i}\right)^2+\left(\sum_{\rm i=1}^{\rm n}\sin\theta_{\rm i}\right)^2\right],$$

where $\theta_{\rm i}=2\pi f t_{\rm i}$. The R2-method test statistic, $z_{\rm R2}$, is obtained by doubling the angles $\theta_{\rm i}$ (e.g. Batschelet 1981, his Fig. 4.2.2).

The R1-method test statistic z_{R1} depends on the length of the sum for the vectors $[\cos\theta_{\mathrm{i}}, \sin\theta_{\mathrm{i}}]$, i.e. these vectors are more or less parallel for a good periodicity candidate. Thus the R1-method is sensitive *only* to unimodal phase distributions. A maximum of the R2-method test statistic is obtained when the above vectors concentrate into two groups with an angular separation of 180 degrees, which explains the sensitivity in detecting regular bimodal phase distributions.

The critical levels of the R1- and R2-methods are

$$Q_{\mathrm{R1}} = 1-(1-\mathrm{e}^{z_{\mathrm{R1}}})^{\mathrm{m}} \qquad \text{(A.4)}$$

$$Q_{\mathrm{R2}} = 1-(1-\mathrm{e}^{z_{\mathrm{R2}}})^{\mathrm{m}} \qquad \text{(A.5)}$$

when testing m independent frequencies.

We wish to point out that the method in Clube & Napier (1996) was nearly equivalent to the R1-method, because their vector R satisfied $z_{\mathrm{R1}} = n^{-1}R^2$. Nevertheless, they applied a test statistic $I = 2n^{-1}R^2$ and Monte Carlo significance estimates in their period analysis of mass extinctions of species.

A.4. The LS-method

When Alvarez & Muller (1984) applied the Fourier power spectrum analysis, they replaced each $t_{\mathrm{i}} \pm \sigma_{\mathrm{i}}$ with a gaussian distribution of unity area. These distributions were then superimposed to construct an event-rate function $G_\Sigma(t_{\mathrm{i}})$. This function is denoted by $X_{\mathrm{j}} = X(t_{\mathrm{j}})$ below. Note that this interpolated function X_{j} consists of $N_0 > n$ values. The LS-method test statistic is

$$z_{\mathrm{LS}} = \frac{\left[\sum_{\mathrm{j=1}}^{\mathrm{N_0}} X_{\mathrm{j}} \cos 2\pi f(t_{\mathrm{j}}-\tau)\right]^2}{2\sum_{\mathrm{j=1}}^{\mathrm{N_0}}[\cos 2\pi f(t_{\mathrm{j}}-\tau)]^2} + \frac{\left[\sum_{\mathrm{j=1}}^{\mathrm{N_0}} X_{\mathrm{j}} \sin 2\pi f(t_{\mathrm{j}}-\tau)\right]^2}{2\sum_{\mathrm{j=1}}^{\mathrm{N_0}}[\sin 2\pi f(t_{\mathrm{j}}-\tau)]^2}, \qquad \text{(A.6)}$$

where τ is obtained from

$$\tan(4\pi f\tau) = \left(\sum_{\mathrm{j=1}}^{\mathrm{N_0}} \sin 4\pi f t_{\mathrm{j}}\right)\left(\sum_{\mathrm{j=1}}^{\mathrm{N_0}} \cos 4\pi f t_{\mathrm{j}}\right)^{-1}.$$

We converted this test statistic z_{LS} into "noise units" by dividing it with $\langle z_{\mathrm{LS}}\rangle$, which was the average of all random sample periodograms of CASE I, CASE II or CASE III. Thus for m independent frequencies, the critical level of the LS-method is

$$Q_{\mathrm{LS}} = 1-(1-\mathrm{e}^{-z_{\mathrm{LS}}/\langle z_{\mathrm{LS}}\rangle})^{\mathrm{m}}. \qquad \text{(A.7)}$$

The sensitivity of the LS-method to unimodal phase distributions follows from the assumption that the interpolated function X_{j} can be modelled with a sinusoid, which maximizes z_{LS}. That this interpolated function resembles a sinusoid requires a unimodal $t_{\mathrm{i}} \pm \sigma_{t_{\mathrm{i}}}$ distribution, e.g. for the sum of superimposed gaussian distributions $G_\Sigma(t_{\mathrm{i}})$.

In conclusion, regular unimodal phase distributions can be detected with the SD-, K-, R1- and LS-methods, and regular bimodal ones with the R2-, K- and SD-methods. The K- and SD-methods are also sensitive to irregular bi- or multimodal distributions.

A.5. The spurious periods

We do not regard the "human-signal" as real periodicity, although these regularities were detected just like those that were claimed to be real in several earlier studies. But one might (erroneously) argue that rounding cannot enhance periodicities such as the noninteger 28.4 Myr cycle in C7. A related argument would be that rounding to an accuracy of 1, 2, 5 or 10 Myr does not interfere with periodicity analysis at larger P, say between 30 and 100 Myr. Let us select C7 to illustrate how spurious periodicities arise from rounding. The periodicities of $P_1 = 10\mathrm{Myr}$ and $P_2 = 4\mathrm{Myr}$ in C7 reach $Q_{\mathrm{SD}} = 0$, because the probability for $V_{\mathrm{a_n}} = 0$ is *never* under H_0. These two periodicities enhance a well known set of spurious periodicities

$$P_3 = [P_1^{-1} + k_1(k_2P_2)^{-1}]^{-1}, \qquad \text{(A.8)}$$

where $k_1 = \pm1, \pm2, ...$ and $k_2 = 1, 2, ...$ This relation states that the phase distributions for t_{i} separated by integer multiples of P_2 are identical with P_1 and P_3. Some interesting combinations in C7 are

P_1	P_2	k_1	k_2	P_3
10	4	−1	4	$+26\frac{2}{3}$
10	4	−1	5	+20
20	4	−1	3	−30
20	10	−1	6	+30
20	10	−1	7	+28
20	10	−1	8	$+26\frac{2}{3}$.

The $P_3 = 20\mathrm{Myr}$ periodicity from the second line can replace P_1 in the third line above, because the 4 or 10 Myr periodicities were also enhanced through the same relation by periodicities like 1, 2 or 5 Myr. The number of possible combinations in Eq. A.8 is unlimited. It is therefore next to impossible to identify the particular combinations that may have induced the 28.4 Myr cycle, but the interplay of the above periodicities represents one potential alternative.

But the relation of Eq. A.8 does not certainly cover all spurious periodicities. For example, because C7 satisfies $V_{\mathrm{a_n}} = 0$ for $P_1 = 10\mathrm{Myr}$, all periodicities P_1/k_4 ($k_4 = 1, 2, ...$) also satisfy $V_{\mathrm{a_n}} = 0$. Periodicities like $P_1/2 = 5\mathrm{Myr}$, $P_1/3 = 3\frac{1}{3}\mathrm{Myr}$, $P_1/4 = 2\frac{1}{2}\mathrm{Myr}$, $P_1/5 = 2\mathrm{Myr}$,... are therefore connected to $P_1 = 10\mathrm{Myr}$. In other words, rounding to an accuracy of 1, 2 or 5 Myr increases the probability that the data fit some multiple of 1x2x5 Myr, like 10, 20 or 30 Myr. This effect, as well as the interplay of numerous spurious periodicities expressed in Eqs. A.8, increases the probability for obtaining extreme test statistic values for rounded data. This was nicely illustrated in Figs. 1 and 2, where the mean periodograms of CASE II or CASE III exceeded those of CASE I.

We conclude this appendix with a contradictory remark that is hopefully not misunderstood. Rounded data do not only cause unreliability in the CASE I Monte Carlo significance estimates, but the theoretical significance estimates based on H_0 suffer the same fate. This "null hypothesis" assumes a *continuous* phase distribution between 0 and 1. But rounded t_{i} yield a *discrete* ϕ_{i} distribution! Hence none of our critical level estimates in Table 1 is actually correct. Those values overestimate significance, because the probability for obtaining extreme test statistic values is higher for rounded data, as explained above or illustrated earlier in connection with Figs. 1 and 2. The solutions for theoretical critical level estimates with a revised "null hypothesis", which would incorporate the effects of rounding, are beyond the scope of this study.

Acknowledgements. We wish to thank our referee, Prof. R.A. Muller, for valuable comments on the manuscript. The research by J.P. at Copenhagen was financed by NORDITA.

References

Alvarez L.W., Alvarez W., Asaro F., Michel H.V., 1980, Sci 208, 1095
Alvarez W., Muller R.A., 1984, Nat 308, 718
Bahcall J.N., Bahcall S., 1985, Nat 316, 706
Bai T., 1992, ApJ 397, 584
Batschelet E., 1981, Circular statistics in biology. Academic Press, London
Benton M.J., 1995, Sci 268, 52
Bottomley R., Grieve R., York D., Masaitis V., 1997, Nat 388, 365
Bowring S.A., Erwin D.H., Jin Y.G., et al., 1998, Sci 280, 1039
Broadbent S.R., 1955, Biometrika 42, 45
Broadbent S.R., 1956, Biometrika 43, 32
Clube S.V.M., Napier W.M., 1996, QJRAS 37, 617
Davis M., Hut P., Muller R.A., 1984, Nat 308, 715
Erwin D.H., 1994, Nat 367, 231
Farley K.A., Montanari A., Shoemaker E.M., Shoemaker C.S., 1998, Sci 280, 1250
Gladman B.J., Migliorini F., Morbidelli A., et al., 1997, Sci 277, 197
Grieve R.A.F., Pesonen L.J., 1996, Earth, Moon and Planets 72, 357
Harris A.W., 1996, Earth, Moon and Planets 72, 489
Heisler J., Tremaine S., 1989 Icarus 77, 213
Jetsu L., Pelt J., 1996, A&AS 118, 587
Jetsu L., 1997, A&A 321, L33 (Paper I)
Kerr R.A., 1998, Sci 279, 652
Kuiper N.H., 1960, Proc. Koningkl. Nederl. Akad. Van Wettenschappen, Series A 63, 38
Leitch E.M., Vasisht G., 1997, New Astronomy 3, 51
Matsumoto M., Kubotani H., 1996, MNRAS 282, 1407
Melosh H.J., 1998, Nat 394, 221
Napier W.M., 1997, Celestial Mechanics and Dynamical Astronomy 69, 59
Patterson C., Smith A.B., 1987, Nat 330, 248
Pope K.O., Baines K.H., Ocampo A.C., Ivanov B.A., 1997, Journal of Geophysical Research 102, 21,645
Rampino M.R., Stothers R.B., 1984, Nat 308, 709
Rampino M.R., Haggerty B.M., 1996, Earth, Moon and Planets 72, 441
Raup D.M., Sepkoski J.J., Jr, 1984, Proc. Natl. Acad. Sci. USA 81, 801
Raup D.M., Sepkoski J.J. Jr, 1986, Sci 231, 833
Raup D.M., Sepkoski J.J. Jr, 1988, Sci 241, 94
Scargle J.D., 1982, ApJ 263, 835
Schwartz R.D., James P.B., 1984, Nat 308, 712
Shoemaker E.M., Wolfe R.F., Shoemaker C.S., 1990, Geol. Soc. Am. Spec. Pap. 247, 155
Spray J.G., Kelley S.P., Rowley D.B., 1998, Nat 392, 171
Stephens M.A., 1965, Biometrika 52, 309
Stigler S.M., Wagner M.J., 1987, Sci 238, 940
Stothers R.B., 1998, MNRAS 300, 1098
Stöffler D., Claeys P., 1997, Nat 388, 331
Swanepoel J.W.H., De Beer C.F., 1990, ApJ 350, 754
Trefil J.S., Raup D.M., 1987, Earth and Planetary Science Letters 82, 159
Urey H.C., 1973, Nat 242, 32
Whitmire D.P., Jackson A.A., 1984, Nat 308, 713
Yabushita S., 1991, MNRAS 250, 481
Yabushita S., 1996, MNRAS 279, 727