Periodicity in extinction and the problem of catastrophism in the history of life

J. JOHN SEPKOSKI, JR.

Department of the Geophysical Sciences, University of Chicago, 5734 South Ellis Avenue, Chicago, Illinois 60637, USA

Abstract: The hypothesis that extinction events have recurred periodically over the last quarter billion years is greatly strengthened by new data on the stratigraphic ranges of marine animal genera. In the interval from the Permian to Recent, these data encompass some 13,000 generic extinctions, providing a more sensitive indicator of species-level extinctions than previously used familial data. Extinction time series computed from the generic data display nine strong peaks that are nearly uniformly spaced at 26 Ma intervals over the last 270 Ma. Most of these peaks correspond to extinction events recognized in more detailed, if limited, biostratigraphic studies. These new data weaken or negate most arguments against periodicity, which have involved criticisms of the taxonomic data base, sampling intervals, chronometric time scales, and statistical methods used in previous analyses. The criticisms are reviewed in some detail and various new calculations and simulations, including one assessing the effects of paraphyletic taxa, are presented. Although the new data strengthen the case for periodicity, they offer little new insight into the driving mechanism behind the pattern. However, they do suggest that many of the periodic events may not have been catastrophic, occurring instead over several stratigraphic stages or substages.

The claim that mass extinctions recur periodically through geological time (Fischer & Arthur 1977; Fischer 1981; Raup & Sepkoski 1984, 1986; Rampino & Stothers 1984a,b) is an empirical hypothesis concerning pattern in the fossil record. It is an inductive claim based on analysis of several types of palaeontologic data, and not one derived from any *a priori* model for how the history of life is organized. However, the hypothesis of periodicity has become intimately associated with catastrophic models of mass extinction. The reasons for this are twofold: first, few, if any, terrestrial processes have cycle times around 26 Ma, the observed periodicity, leaving astronomical mechanisms as likely forcing agents (Raup & Sepkoski 1984; Raup 1986; Sepkoski & Raup 1986a); and, second, one event in the periodic array, the Cretaceous–Tertiary mass extinction, may have been caused by an extraterrestrial impact, suggesting that the other, serially dependent events may have been caused similarly (Flessa *et al.* 1986; Sepkoski 1986a).

Because of the inferred link between periodicity and catastrophism, the hypothesis has become quite controversial, and a large number of criticisms and counterarguments have been published. In this paper, I will review both the empirical evidence for periodicity and the arguments against it. As will be demonstrated, efforts to obtain better data have strengthened the case for periodicity and have made many of the counterarguments irrelevant. However, the new data also show some patterns that weaken the link between periodicity and catastrophism.

Evidence for periodicity in extinction

The primary evidence for periodicity has come from compilations of the stratigraphic ranges of marine taxa. Fischer & Arthur (1977) claimed that there was a 32 Ma cycle in diversity minima in pelagic ecosystems on the basis of global data for globigerinid species, ammonite genera, and large, open-ocean predators. Raup & Sepkoski (1984),

on the other hand, argued for a 26 Ma periodicity in mass extinctions on the basis of data on fossil marine animal families.

The attraction of using taxonomic data is that these data provide long time series on global evolutionary patterns summed over all marine environments. Detailed biostratigraphic data, consisting of species ranges sampled at zonal or subzonal levels, clearly would be superior to data for higher taxa, which are usually resolved only to the stage level. However, detailed data rarely are available over long time intervals (i.e. tens to hundreds of million years, although see House (1985)) and usually involve only selected taxa; furthermore, the most detailed sets of biostratigraphic data are for only limited environments and geographic regions. The global taxonomic data used in the analyses of periodicity avoid some of these problems, but in so doing lose resolution by using higher taxa, which may be arbitrarily delimited and poorly sampled, and generalized stratigraphic intervals, which may not be precisely defined and accurately correlated worldwide (Sepkoski 1986b). The test of whether such data are reflecting real evolutionary patterns rather than random noise comes from how well these data correspond to patterns or events recognized in available detailed biostratigraphic data (cf. Sepkoski & Raup 1986a).

In their initial analysis of periodicity, Raup & Sepkoski (1984) used a highly culled subset of Sepkoski's (1982a) family compendium. This culling was designed to enhance the signal of extinction by eliminating all data for extant families, families known only from Lagerstätten, and families with extinctions not resolved to the stage level. The culled data set consisted of 567 extinct families in 39 stages of Late Permian to Miocene age. (Too few families were in the subset to extend the analysis through the Pleistocene.) Extinction intensity was measured as the percent of families becoming extinct in each stage. The resulting time series exhibited 12 local maxima, or 'peaks,' of extinction, many of

which corresponded to well-documented mass extinctions. Other peaks, however, were small and possibly spurious, but in order to avoid potential subjective bias, Raup & Sepkoski chose to analyse the complete set of peaks. Unfortunately, they referred to all peaks as 'mass extinctions.'

Several statistical analyses were performed on the time series in order to test whether the peaks were uniformly distributed. The most powerful of these was a randomization procedure (Connor 1986), related to Stothers' (1979) residuals index method (see formal definition in Stigler & Wagner 1987). This procedure fit periodic functions to the actual data and to randomized versions of the data for comparison. The randomization shuffled the 39 data points to produce what was essentially a random walk with the same number of peaks as in the actual time series. Results of the test showed that a periodic function fit the observed data better than 99.99% of random walks at a 26 Ma cycle length. This led Raup & Sepkoski to conclude that there was a significant periodicity to mass extinctions.

Subsequent analyses (Raup & Sepkoski 1986; Sepkoski & Raup 1986a) using unculled familial data with corrected extinction times, more stratigraphic stages, and statistical selection of 'significant' extinction events corroborated this result. These analyses also permitted a more precise measurement of the period length: 26.2 Ma with a standard error of estimate of ±1 Ma. Connor (1986) found essentially the same periodicity in the complete familiar data set with bootstrapped statistical tests, although he judged the standard error on period length to be closer to ±3 Ma. Rampino & Stothers (1984a,b), using only nine of Raup & Sepkoski's (1984) original 12 extinction peaks, also found the 26 Ma periodicity with Stothers' residuals index method of analysis. However, with a regression-based technique (criticized by Stigler 1985), they calculated a 30 Ma periodicity, which they favoured because it was closer to the 33 Ma half oscillation of the solar system about the galactic plane.

The use of familial data in such analyses presents several major problems. The number of familial extinctions counted in any stratigraphic interval tends to be small and subject to

considerable random variation. In frequency data, the standard error about any count, n, approximates $\sqrt{n}$ (see Sepkoski & Raup 1986a); therefore, the coefficient of variation, $1/\sqrt{n}$, increases as the count decreases, making it more difficult to ascertain signal above noise. A more serious problem is that familial data tend to reduce fluctuations in species-level extinction. Most families are polytypic, and all constituent species must disappear if a polytypic family is to register an extinction event (cf. McLaren 1983, 1984). If species extinctions are approximately independent of familial membership (but not necessarily nonselective), then the damping induced by families should follow a rarefaction relationship, as argued by Raup (1979). Figure 1 illustrates rarefaction curves for families as well as genera relative to species. The curve for families has a very shallow slope over low levels of species loss, indicating little sensitivity to extinction; even at a 50% loss of species, families record only a 10% extinction. Genera, which on average contain nearly an order of magnitude fewer species than do families in the modern oceans, have a less convex rarefaction curve and are therefore more sensitive to species extinction; a 50% loss of species should be reflected in an approximately 25% extinction among genera.

In order to perform more robust tests of periodicity, a new, genus-level data set comparable to the previous family compendium has been compiled. Several preliminary analyses of these data have already been published (Raup & Sepkoski 1986; Sepkoski 1986a, b). The compilation uses the *Treatise on Invertebrate Paleontology* (Moore *et al.* 1953–83) and *Vertebrate Paleontology* (Romer 1966) as preliminary sources of information on genera and their stratigraphic ranges. However, more than 750 additional published sources, including journal articles, monographs, and symposium volumes, have been used to update classification, add new genera, delete synonymies, correct errors, and enhance stratigraphic resolution. At present, the compilation contains more than 30,000 fossil marine animal genera, with times of origination and extinction of more than three-quarters resolved to internationally recognized stages and substages of the Phanerozoic record.

Figure 2 illustrates extinction time series computed from the most recent (April 1988) version of the generic data. The upper time series displays the per-genus rate of extinction for the entire data set of approximately 17,500 genera of mid-Permian to Recent age. These extinction rates were computed for 49 sampling intervals, exceeding the 39 to 43 used for families over the same time range. Most of the intervals are stages and include the previously excluded Aalenian. Several of the longer stages have been divided into two (Leonardian, Norian, Bajocian, Tithonian, Campanian, Lower Miocene) or three (Albian) substages, and three pairs of shorter stages (Induan and Olenekian, Turonian and Coniacian, Pliocene and Pleistocene) have been amalgamated. This manipulation was designed to provide a more even and detailed sampling through time, with a mean interval duration of 5.5 Ma and standard deviation of 1.2 to 2.3 Ma, depending on time scale. Approximately 68% of the 13,000 extinct genera in the data set have their times of extinction resolved to one of these intervals, and another 7% are resolved to stages that have been subdivided; among the remaining, 21% are resolved only to series and 4% to systems.

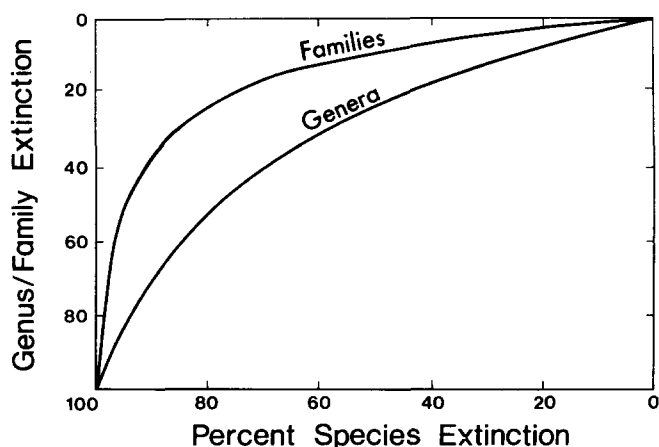All low-resolution data were distributed among the

Fig. 1. Rarefaction curves displaying the expected relationships of percent generic and familial extinction to percent species extinction. The curves show that families are not very sensitive to species extinctions below about 75%, whereas genera are much more sensitive. After Raup (1979).
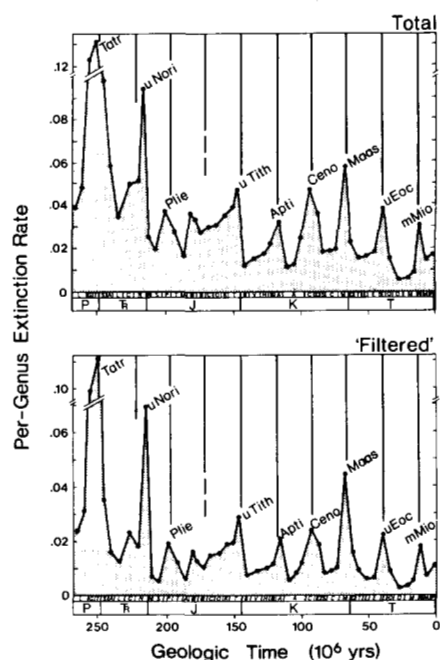
Fig. 2. Per-genus extinction rate (in units of extinctions/genus/Ma) for 49 sampling intervals from the mid-Permian (Leonardian) to Recent. A 26 Ma periodicity, represented by the vertical lines, is superimposed in best-fit position to illustrate the conformity of the data to the hypothesis of periodic extinction. The upper time series, labelled Total, is for the entire data set of 17,500 genera, whereas the lower, 'filtered' time series is for a subset of 11,000 from which genera confined to single stratigraphic intervals have been excluded. Systems and stages (but not sampling intervals) are indicated along the abscissa of each graph. The time scale is from Harland *et al.* (1982) with the durations of Jurassic stages rescaled using Westermann (1984) in order to eliminate the unrealistic uniformity of estimated durations.

sampling intervals in proportion to high-resolution data. Although it could be argued that the low-resolution data are of no aid in identifying extinction events (cf. Patterson & Smith 1987), they are important in (1) estimating relative magnitudes of events and (2) defining extinction peaks when series differ in the quantity of contained low-resolution data (which is a problem for the series of the Tertiary). My experience with the generic data is that the identity and magnitudes of extinction peaks with distributed data have varied little as the quantity of low-resolution data has been diminished over the last several years of work.

The lower, 'filtered' time series in Fig. 2 is for a subset of data from which genera confined to single stratigraphic intervals have been excluded. These excluded genera often (although not always) are rare and known only from single localities, often Lagerstätten; their inclusion in the time series adds monographic noise to extinction patterns (cf. Boucot 1986; Patterson & Smith 1987) and induces a spurious correlation between extinctions and originations (Sepkoski 1986a). Their exclusion reduces the data set to just over 11,000 genera of mid-Permian to Recent age, of which 6500 are extinct. This improves the resolution of extinctions, with 76% known to one of the 49 sampling intervals, 7% to subdivided stages, 16% to series, and only 1% to systems. This 'filtering' of the data actually has

minimal effect on the patterns displayed by the data, and the two time series in Fig. 2 appear quite similar. The same is true for the other measures of extinction intensity illustrated in Sepkoski (1986a), indicating that the generic data contain more robust patterns of extinction than seen in the earlier familial data.

The fit of the 26 Ma periodicity to extinction peaks is much more striking for genera than for families, as shown by the vertical lines in Fig. 2. This has also been demonstrated by several statistical tests, including the randomization procedure (Raup & Sepkoski 1986), autocorrelation analysis (Sepkoski 1986a), and Fourier analysis (Fox 1987). It should be noted that this most recent generic data set seems to resolve questions about the age of several extinction events. With the familial data, it was uncertain whether the Late Permian peak of extinction fell in the Guadalupian or Tatarian (= Djulfian + Dorashamian) and whether the Late Triassic peak fell in the Norian or Rhaetian (= upper Norian, in part, in this analysis). Because of this uncertainty, Raup & Sepkoski (1986) tested periodicity with four sets of ages for these two events. The new generic data indicate that the two events fall in the last stages of their respective systems. This possibility was tested by Raup & Sepkoski and it was found that the null hypothesis of randomness could be rejected in favor of periodicity with $p < 0.044$ after correction for multiple testing.

The vertical lines in Fig. 2 show that nine of the ten periodic cycles fall on or close to distinct peaks of extinction, with the fit being best over the more accurately dated late Mesozoic and Cenozoic. Only the mid-Jurassic cycle lacks a corresponding peak of extinction. (The lower Bajocian peak, to the left of the predicted cycle, is small and may be spurious.) Very importantly, most of the extinction peaks falling along the periodic function correspond to events recognized in detailed biostratigraphic studies conducted at the species level (Sepkoski 1982b; Sepkoski & Raup 1986a, b). These include the Upper Permian (Guadalupian–Tatarian), upper Norian, upper Tithonian, Cenomanian, Maastrichtian, and Upper Eocene. The Pliensbachian peak probably reflects the event documented by Hallam (1976, 1977) in the lowest zone of the Toarcian, with global extinction records smeared backward into the Pliensbachian as a result of incomplete sampling (Hallam 1986; Sepkoski 1986b). The only well-documented marine (as opposed to exclusively terrestrial) extinction event not represented among the nine periodic peaks in Fig. 2 is the Pliocene event, documented by Stanley (1984, 1986a, b). This event is present in the data, as evident in the slight increase in extinction intensity at the right-hand margins of the graphs; but it is considerably smaller than the periodic events. There is also a small peak in the Carnian, along the left flank of the upper Norian maximum in Fig. 2b; an extinction event has been argued to occur in the Carnian by Benton (1986) using familial data, but no event has yet been demonstrated in detailed biostratigraphic studies of marine sections.

In all, the generic data in Fig. 2 display 11 to 12 peaks of extinction, of which the nine most distinct are periodic. Of the 12, eight correspond to well-documented extinction events. If the generic data reflected only noise, we would expect only about two of the eight extinction events to be sampled by the 12 peaks. The probability of finding eight out of eight events among 12 random peaks scattered over
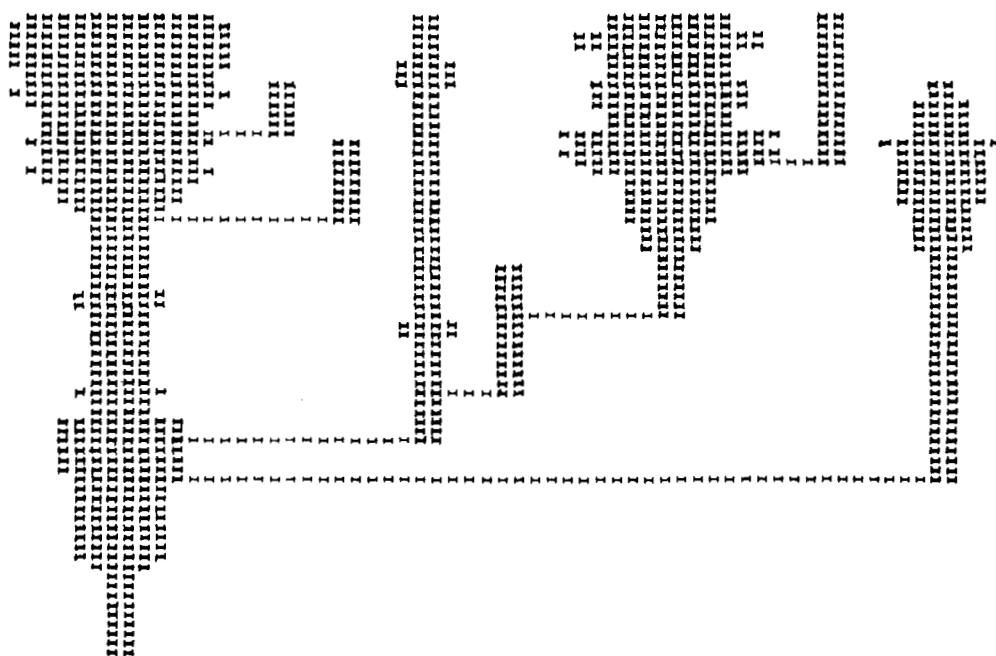
Fig. 3. Examples of paraclades produced by a stochastic Monte Carlo simulation of phylogeny. Time runs from bottom to top in the figure, and widths of the spindle diagrams indicate the number of lineages in a paracalde during each time interval; connecting lines indicate phylogeny. Note that the paraclades vary considerably in size from one to many lineages and can be either monophyletic or paraphyletic.

48 intervals (the Guadalupian and Tatarian must be combined) is only 1.3×10^{-6}. And, the probability of finding seven or more of the eight documented extinction events in nine periodic peaks among the 48 intervals is less than 0.5×10^{-7}. These minute probabilities imply that the generic data are reflecting real phenomena in the history of extinction.

Arguments against periodicity

The hypothesis of periodicity in extinction has provoked a variety of criticisms and counter-arguments. Most fall into four broad categories: criticisms of the taxonomic data base, of the sampling intervals, of the chronometric time scale, and of the statistical methods. Each of these is considered below.

Taxonomic data base

Raup & Sepkoski's (1984) initial use of a highly culled data set of fossil families has been criticised frequently, most extensively by Hoffman (1985). He argued that different treatments of the familial data yield time series with different peaks of extinction, which he claimed falsified the hypothesis of periodicity. However, much of the apparent noisiness of the familial data reflects damping of the extinction signal. The generic data in Fig. 2 damp the signal to a much lesser degree, as argued above, and indeed display a much more obvious and robust pattern of extinction peaks. Furthermore, as demonstrated by Sepkoski & Raup (1986b) and Sepkoski (1986a), even Hoffman's time series exhibit a strong 26 Ma periodicity when the noise is filtered out and only consistent peaks are considered.

A more intriguing criticism of the familial data has been presented by Patterson & Smith (1987). They performed a much-needed thorough review of the extinct families of Mesozoic–Cenozoic echinoderms and fishes in Sepkoski's (1982a) compendium. In addition to a number of errors in listed times of extinction, they discovered that approximately 55% of the families are not clades but rather monotypic, paraphyletic, or even polyphyletic groups. The last appearances of such families, they argued, constitute 'pseudo-extinctions' and thus represent noise in the data. However, while the first part of that statement is true for phylogenetic units, the second part does not necessarily follow. Families reflect extinction events not because they are real evolutionary units in any sense but because they represent a sample of species (cf. Sepkoski 1984). Each familial (or generic) extinction, if correctly determined, encompasses at least one species extinction, whether or not the taxon is strictly monophyletic, so long as the last member(s) of the family has not been arbitrarily cleaved at some point in a phyletic transition. Thus, if the number of families in a data set is substantial, it should constitute a sufficiently large sample of species to provide an estimate of underlying patterns of species extinction.

This argument can be supported with Monte Carlo simulations of phylogeny, using methods developed by Raup *et al.* (1973). Model phylogenies are generated by random branching and extinction of 'lineages' and are then divided into 'clades' (or, more properly, 'paraclades' (Raup 1985a)) using any of several algorithms. The algorithm I have used divides the phylogenies at random, as described in Sepkoski (1978). Figure 3 illustrates spindle diagrams for several of the resulting paraclades. As evident, the paraclades can be either paraphylectic or truly monophyletic and either monotypic or polytypic. Despite this total arbitrariness in how the paraclades are delimited, their patterns of

extinction, origination, and diversity tend to parallel patterns among lineages, so long as sample sizes are not too small. Figure 4 illustrates time series of extinction rates for lineages and paraclades in a simulation that produced 1462 lineages and 147 random paraclades. (Only part of the simulation is shown; and equilibrium constraint was imposed on the lineage diversity, and only data following attainment of 85% of the equilibrium are illustrated.) No extinction events were modeled in the simulation, so that all variation in rates represents stochastic fluctuations. Despite the relatively small numbers of extinctions (an average of 7.5 per sampling interval for the paraclades), there is still a fair correspondence between extinction patterns in the lineages and paraclades, with an overall linear correlation of 0.791 between the two rates. Thus, paraphyletic groups, even at fairly low numbers, can reflect the extinction histories of their constituent lineages.

Patterson & Smith went on to argue that periodicity appeared only in the 'noise component' of the echinoderm and vertebrate familial data. They suggested that this might result from taphonomic variation rather than any real pattern of extinction: apparent extinction in the data is caused when a stage with a good fossil record is followed by one or more with very poor records. The generic data presented here provide three arguments against this contention. First, if the data reflected mere fluctuations in fossiliferousness, we would expect an association of *apparent extinctions* (last documented occurrences) and apparent originations (first documented occurrences). But no such association exists among the genera, especially when those confined to single stratigraphic intervals are filtered from the data. The correlation between raw numbers of extinctions and originations for the 49 intervals from the mid-Permian to Plio–Pleistocene is only 0.194, with the positive sign reflecting general increase in diversity
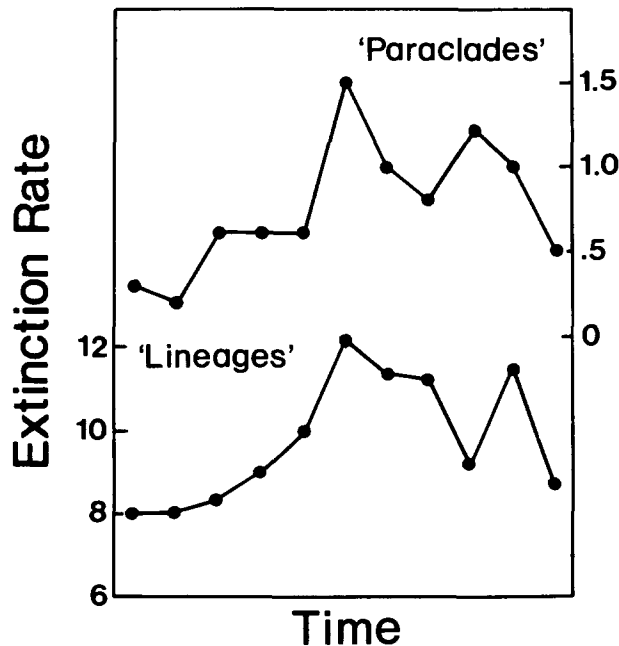
over the time (providing more genera for both origination and extinction). Second, if apparent extinction patterns were caused by fluctuations in fossiliferousness, we would expect peaks to occur in stages with spectacular Lagerstätten. But this is not the case (Sepkoski 1987): peaks occur in the Pliensbachian (and not the Toarcian containing the Posidonienschiefer), the upper Tithonian (and not the lower Tithonian containing the Solnhofen), and the Upper Eocene (and not the Middle Eocene containing the Monte Bolca fish beds). Finally, if the generic extinction pattern were mere artefact, we would not expect any correspondence with extinction events recognized in restricted but detailed biostratigraphic studies of species lineages. But, as argued above, there is a strong correspondence and a vanishingly small probability that this could occur by chance.

This same line of argumentation applies to Stigler & Wagner's (1987) observation that periodicity might be induced by imperfect sampling of the fossil record. They generated a large number of random time series of extinction and smoothed them with moving averages to mimic the backward-smearing of extinction records resulting from failure to sample taxa in their last stage of existence (Signor & Lipps 1982). This filtering tended to eliminate smaller extinction peaks and cause the remaining peaks to appear more evenly spaced than expected for a Poisson (random) distribution. Stigler & Wagner found that smoothed time series, in which 50% of extinction records were displaced by one stage, favored a 26 Ma periodicity in 8% of their simulations and any periodicity in the range 25–30 Ma in 30%.

Whether Stigler & Wagner's model is applicable to the Mesozoic–Cenozoic fossil record can be questioned, however. Most biostratigraphically documented extinction events are well spaced in time, and few smaller events missed by the generic data intervene. I know of only three documented cases from the marine record in which this might be true: the Upper Eocene which seems to encompass several steps, with the earliest occurring late in the Middle Eocene (Hut *et al.* 1987); the Pliocene which represents an event intervening between the Middle Miocene and Recent; and the Upper Triassic which arguably includes a small event in the Carnian (Benton 1986). Only the first case does not show through in the generic data. With the other two events included, Raup & Sepkoski (1986) still found that randomness could be rejected in favour of the 26 Ma periodicity.

Harper (1987), like Stigler & Wagner, argued that the appearance of periodicity could result from a moving-averages process. He based his argument on the work of Slutsky (1937), who showed that summations of random time series can display quasiperiodicities in the absence of any periodic forcing. Harper noted that the number of extinctions, E_t, in any time interval t can be considered the sum of extinctions from cohorts with initial diversities O_{t-i} that originate $t - i$ ($i = 0, 1, 2, \ldots$) intervals in the past:

$$E_t = kO_t + k^2O_{t-1} + k^3O_{t-2} + \cdots + k^{i+1}O_{t-i} \qquad (1)$$

where k is the proportion of a cohort becoming extinct from one time interval to the next. If each O_{t-i} is essentially random, then E_t becomes a sum of random time series and might display the type of quasiperiodicity discovered by Slutsky. This tendency will be enhanced in fossil data,

Fig. 4. Total extinction rates (in units of extinctions/unit time) for lineages (below) and paraclades (above) produced in the simulation illustrated in Fig. 3. Data points are for intervals of ten simulated time units, paralleling the measurement of extinction rates over intervals of several million years duration in the fossil record.

according to Harper, because extinction metrics are measured for stratigraphic stages of several million years duration and are thus averages of several E_i's. (Harper incorrectly identified this last effect as a moving average; but the metric measured for any one stage in fact does not include values of E_i from adjacent stages.)

This argument, while ingenious, seems to have two flaws: (1) the quasiperiodicity from moving averages should change under change of scale, but no change in the 26 Ma periodicity was found when stratigraphic stages were subdivided and the density of sampling intervals increased from 39 to 49; and (2) most of the quasiperiodic peaks should reflect extinctions spread throughout their respective stages and substages and should not be manifested in biostratigraphically recognizable short-term events, which, again, is not the case. An additional argument against Harper's model is that the deviations from periodicity in the new generic data are much smaller than those measured in the previous data sets. Figure 5 presents a circular plot of the ages of the extinction peaks in Fig. 2 similar to that used by Harper. This kind of plot, which has a circumference scaled to period length, provides a good visual display of dispersion around best-fit periodicities. Harper, using the data of Raup & Sepkoski (1986), found that 10 of 13 extinction peaks fell within a 140° sector, which he noted was comparable to the arc formed by 8 of 11 peaks in Slutsky's (1937) random data. In the case of the new generic data, however, 8 of 11 extinction peaks fall within 66° of arc,
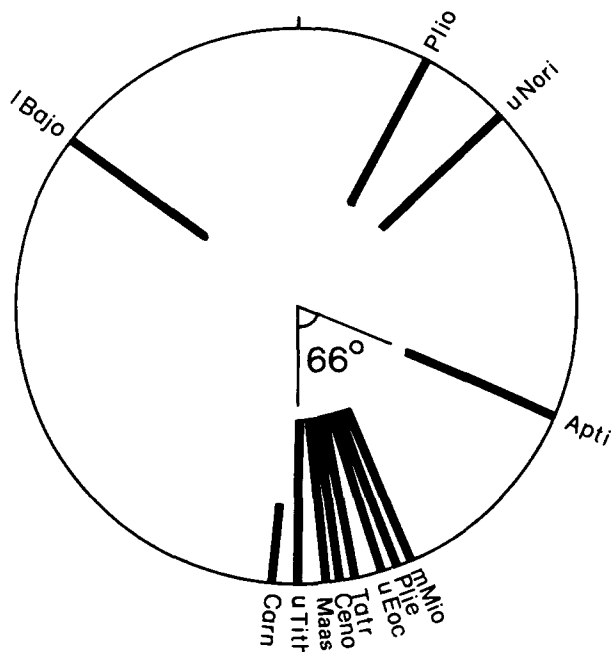
less than half the dispersion of the moving averages data; with the questionable Carnian peak of Fig. 2b included, 9 of 12 peaks fall within 71°. Thus, the generic data seem to be more periodic than expected for the type of quasiperiodicities generated from moving averages of random data.

Sampling intervals

A major difficulty working with global taxonomic data is that the sampling intervals (i.e. stratigraphic units) are not equal. This precludes rigorous application of many statistical tests developed for classical time series analysis which assume uniformly spaced data (Connor 1986). The sampling intervals in the familial data of the early analyses are the stratigraphic stages used in Harland et al. (1967) and copied by Sepkoski (1982a). The durations of these stages within the Mesozoic and Cenozoic range from 1 to 15.5 Ma with a standard deviation of 2.7 to 2.9 Ma, depending on time scale. Because of this, Raup & Sepkoski (1984, 1986) based their analyses on estimates of absolute time rather than numbers of intervals between events and developed the randomization test to accommodate the irregular template imposed by the stratigraphic column.

Several subsequent authors (Kitchell & Pena 1984; Hoffman & Ghiold 1985; Gilinsky 1986; Noma & Glass 1987) have treated the sampling intervals as if they were equal. However, their results are not necessarily general and may become irrelevant if the density of sampling intervals is changed, as in the generic data. For example, Noma & Glass (1987) counted the frequency of intervals of length d ($d = 1$, 2, $\geqslant 3$) between 'turning points' (minima and maxima) in Raup & Sepkoski's (1984) familial time series; on the basis of a chi-square test, they argued that these frequencies did not differ from the expectations of randomness (although there are counting errors in Noma & Glass's table 1). They also counted the simple number of turning points and argued these, too, were consistent with expectations of randomness. In a random time series, the expected number of turning points is $2(n-2)/3$ with a standard deviation of $[(16n - 29)/90]^{\frac{1}{2}}$, where n is the total number of observations. Raup & Sepkoski's original familial data had 39 observations with 22 turning points, compared to an expectation of 24.67 and standard deviation of 2.57; this yields a z-statistic (observed minus expected value divided by standard deviation) of -1.04, which is not significant.

It is not clear, however, whether Noma & Glass's tests demonstrate anything other than variance in stage durations. Harper (1987) applied these tests to Raup & Sepkoski's (1986) later familial time series (which Noma & Glass commented upon but did not test) and obtained results that differed significantly from random. The generic data in Fig. 2, with more even sampling intervals than in the familial data, also give nonrandom results: the upper time series, with 49 observations, has 20 turning points, yielding a z-statistic of -3.90, and the lower time series has 22 turning points, for a z-statistic of -3.21. Both differ from the expectation of randomness with $p < 0.001$. Similarly, the frequencies of intervals between turning points in the generic data, tabulated in Table 1, yield chi-square statistics that differ from expectations of randomness at high levels of confidence. In the conclusion to their critique, Noma & Glass (1987) stated that 'The existence of a cyclical pattern can only be established through discovery off [sic] more data

Fig. 5. A circular graph of the ages of all generic extinction peaks in Fig. 2, plotted in the manner of Lutz (1985) and Harper (1987). The circumference of the circle is 26.2 Ma which is the best-fit periodicity found by Sepkoski & Raup (1986a). The Recent (0 Ma) is placed at the 12 o'clock position, and extinction peaks (arbitrarily considered to fall at the ends of stages) are plotted clockwise at a position equal to 360° × age/26.2. Note that 8 of 11 peaks (or 8 of 12 if the variable Carnian peak is included) fall within less than 70° of arc. Peaks falling outside are the periodic upper Norian mass extinction, the aperiodic small Pliocene event, and the questionable lower Bajocian peak. The time scale is the same as used in Fig. 2.

Table 1. *Frequencies of intervals between 'turning points' (minima and maxima) in the genus-level extinction time series in Fig.* 2.

Number of intervals	Observed frequencies		Expected frequency
	'filtered'	'total'	
1	9	5	19.17
2	5	6	8.25
≥3	7	8	2.92

Expected frequencies were calculated with the formula used by Noma & Glass (1987): $2(n - d - 2)(d^2 + 3d + 1)/(d + 3)!$, where n is the number of observations (≈ 49) and d is the number of intervals. Chi-square statistics, with 2.5 degrees of freedom, are $X^2 = 12.38(p < 0.001)$ for the 'filtered' data set and $X^2 = 19.93(p < 0.001)$ for the total set.

or refinement of the temporal resolution of the existing data.' The generic data presented here provide both and indeed negate many of their criticisms.

In an earlier critical paper, Hoffman & Ghiold (1985) also claimed that the apparent periodicity of extinction might simply reflect a random walk through the stages (despite the fact that Raup and Sepkoski's randomization test actually compared the observed data to random walks, a point also ignored by Noma & Glass (1987)). Hoffman & Ghiold argued that in a random time series, there is a 0.5 probability of extinction going either up or down from one stage to the next; thus, the probability of going up and then down to produce an extinction peak is 0.25, or 1 in 4. This, they claimed, was sufficient to account for the frequency of peaks in the familial data, since the average stage duration is 6.4 Ma and $4 \times 6.4 = 25.6 \approx 26$ (see also Hoffman 1985, 1986a).

This argument has been criticized on *a priori* grounds by several authors. Harper (1987) argued that peaks in a *random model should in fact have an average occurrence of one in three intervals*, generating a mean peak spacing of about 19 Ma ($\approx 3 \times 6.4$ Ma). On the other hand, Gould (1985), Gilinsky (1986), and Sepkoski & Raup (1986b) pointed out that similarity in frequency does not necessitate similarity in pattern. Kitchell & Estabrook (1986) used Monte Carlo simulations to show that the 12 peaks in Raup & Sepkoski's (1984) initial data could appear periodic about 7% of the time when randomly distributed among the 39 stages. However, Raup & Sepkoski (1986) demonstrated that when unculled familial data and 43 stages are used, random peaks appear periodic no more than 4.8% of the time and virtually never fit a periodic function as well as the significant peaks in the data. Using a slightly earlier version of the generic data in Fig. 2, Sepkoski (1986a) showed that autocorrelation analysis revealed a statistically significant cycle at lags of five intervals (which is consistent with a 26 Ma periodicity) and not at four intervals, as predicted by Hoffman & Ghiold's (1985) argument.

Autocorrelation analysis involves computing the product-moment correlation coefficient for a time series with itself at a lag of $t - i$ time intervals:

$$r_{t,t-i} = \sum_{t=i}^{n} (x_t - \bar{x})(x_{t-i} - \bar{x})/[(n - i)s_x^2] \qquad (2)$$

where $\bar{x}$ is the mean of the time series with n observations and s_x^2 is its variance. Autocorrelation is related to autoregression techniques which attempt to predict values at

time t in terms of values in previous time intervals; the simplest autoregressive models are linear of the form:

$$\hat{x}_t = a_0 + a_1 x_{t-1} + a_2 x_{t-2} + \ldots + a_i x_{t-i} \qquad (3)$$

where each a is a constant fit by least-squares methods. Kitchell & Pena (1984) fit a series of more complicated autoregressive models to the familial data, treating the stages as if equal in length. The first model they tested was a simple impulse function with a five-stage period, which they found did not fit Raup & Sepkoski's (1984) data. This is hardly surprising, however, since the separation of the periodic events, especially as identified in Raup & Sepkoski (1986), varies from two to six stages, and the simple impulse function attempts to fit the amplitude as well as timing of the events. (Raup & Sepkoski were concerned only with timing.)

Kitchell & Pena obtained their best fit to both amplitude and timing with a fifth-order autoregressive model that exhibited pseudoperiodic behavior. But this model, and its pseudo-periodicity, are difficult to interpret, since the incorporated five-stage interval is not of fixed length, varying from 20 to 40 Ma through the Mesozoic and Cenozoic. Furthermore, it is difficult to understand exactly what a five-stage memory of extinction intensity means in any evolutionary sense. Kitchell & Pena did note that equilibrial systems can exhibit endogenous oscillations free of any external forcing if these systems have internal lags or operate in discrete steps. However, this behavior occurs only in the vicinity of the system's equilibrium, and there is no evidence that diversity was anywhere near equilibrium during the Mesozoic–Cenozoic interval (Kitchell & Carr 1985).

Stanley (1986b) has presented another evolutionary argument for how internal lags might produce a pseudo-periodicity in extinction data. He suggested that there might be a recovery time after each extinction events during which ecosystems would be dominated by generalist taxa that are relatively immune to massive extinction. This recovery time would impose a minimum spacing between extinction events, so that time series would appear more regular than expected for random events and might even approach the appearance of periodicity (see also McKinney 1987). However, the familial and generic data indicate that such recovery times, as measured by rebounds in diversity, last for only a stage or two, except after the larger mass extinctions; and even after the Maastrichtian event, diversity recovered to previous levels within four stages (see Sepkoski 1984). The randomization test used by Raup & Sepkoski (1984, 1986) actually incorporated a two-stage recovery time: because peaks in extinction can be identified only when separated by one or more intervals of low extinction, the randomization procedure was constrained to prevent peaks from falling into contiguous stages during shuffling. Thus, the test distribution was somewhat less variable than expected for a truly random (i.e. Poisson) distribution, with the minimum spacing between peaks being equal to the durations of two stages. This distribution still did not fit the data as well as a periodic distribution, suggesting that Stanley's explanation is insufficient.

Chronometric time scale

Perhaps the strongest argument against the hypothesis of periodicity is that available estimates of the absolute ages of

extinction events are still inadequate and might contain too much error to permit rigorous statistical testing. As pointed out by Hallam (1984) and repeated by Hoffman (1985), the various chronometric time scales published over the last several years differ considerably in estimated ages of various stages, especially for those older than the mid-Cretaceous. This, however, does not preclude the discovery of periodicity. As Raup & Sepkoski (1986) have argued, if differences among time scales reflect random error rather than systematic bias, they can only degrade the signal of periodicity; they cannot produce it.

Shoemaker & Wolfe (1986) assessed the ages of Raup & Sepkoski's (1984) extinction peaks and concluded most were too inaccurately dated to support the hypothesis of periodicity. They accepted only three events (the Cenomanian, Maastrichtian, and Upper Eocene) as being reliably dated and argued this was too small a sample to distinguish periodicity from randomness. One of the events they rejected, however, is the Middle Miocene peak, which was represented by few extinctions in the familial data. But this event is represented by many more extinctions in the generic data and should not be excluded. Raup & Sepkoski (1986) included the Middle Miocene event with the three others in an analysis of periodicity over the accurately-dated last 100 Ma of geologic time. Their randomization test for this interval overwhelmingly rejected randomness in favor of the 26 Ma periodicity. Fox (1987), working with the last 130 Ma, showed that Fourier analysis also indicates a strong 26 Ma periodicity in the generic data; this is important because Fourier analysis does not require explicit identification of peaks. Fox also demonstrated that the older, less accurately dated portion of the generic data, from 130 to 260 Ma, has the same periodicity, with a nearly identical amplitude, wavelength, *and phase* as found in the better-dated, younger portion.

Table 2 documents the effects that differences in time scale have on the fit of a 26 Ma periodicity. The time scales listed are the four most recently published, with a fifth version added by modifying the Harland *et al.* (1982) scale so that Jurassic stages have varying lengths. Fits to periodicity were measured as the standard deviation of differences between observed and expected positions of the nine periodic extinction peaks illustrated in Fig. 2. For each time scale, fits were computed placing the peaks at the end, middle, and bottom of the stages. This manipulation produces only minor changes in fits, even for time scales

with variable stage durations (contra Hoffman 1986*b*). The best fits to the 26 Ma periodicity generally occur when the peaks are placed at the middles of stages, indicating that Raup & Sepkoski (1984, 1986) were somewhat conservative in their testing when they measured goodness of fit to the ends of stages.

Differences in fits among time scales are fairly large, with the scale of Harland *et al.* (1982) producing the best fits and those of Odin (1982) and Snelling (1985) the worst. Yet, Raup & Sepkoski (1984) found that even the Odin scale provided fits to periodicity that were significantly better than random distributions. Much of the deviation from periodicity in all time scales results from a few poorly fit events in the older part of the Mesozoic where estimates of ages have fairly substantial uncertainties. For example, the upper Norian event alone contributes 22–57% of the lack of fit, depending on time scale (see Fig. 5). Table 2 also lists fits of periodicity to the last four extinction events over the well-dated last 100 Ma, using the same conventions as before. Again, the Harland time scale provides the best fit and the Odin scale the worst. But the differences in fits are rather small, and even the Odin scale yields standard deviations that are no more than 10% of the period length. Thus, although uncertainties in estimated ages of extinction events are problematic, they are not fatal to the hypothesis of periodicity.

But there is one other potential problem with the time scales, discovered by Stigler & Wagner (1987). They placed a large number of random time series over the geologic time scale of Harland *et al.* (1982) and used the nonparametric method of Raup & Sepkoski (1984) to find best-fit period lengths (regardless of whether or not these fits were significant). They found that approximately one-quarter of best-fit periodicities longer than 20 Ma occurred in the 26–27 Ma band, leading them to conclude that the 26 Ma periodicity is embedded in the time scale. This result is hardly surprising, however, since a number of stratigraphic boundaries are placed where there is substantial faunal turnover resulting from mass extinction (e.g. the Palaeozoic–Mesozoic and Mesozoic–Cenozoic boundaries). Thus, if there is a periodicity in extinction, it should be reflected, albeit weakly, in the time scale. This fact was recognized by Raup & Sepkoski (1984, 1986), who shuffled the stratigraphic stages in their randomization procedure in order to eliminate any potential coupling between the stratigraphic and biological records. (This shuffling was not

Table 2. *Fits of a 26 Ma periodicity to all principal peaks of generic extinction over the last 270 Ma* ($n = 9$) *and to peaks in the well-dated last 100 Ma* ($n = 4$).

Time scale	Last 270 Ma			Last 100 Ma		
	u	*m*	*l*	*u*	*m*	*l*
Harland *et al.* (1982)	2.7	2.3	2.5	0.7	1.7 (0.6)	2.6 (0.6)
Harland (modified)	2.8	2.9	3.0	0.7	1.7 (0.6)	2.6 (0.6)
Odin (1982)	5.8	5.0	4.8	2.0	2.6 (1.9)	2.6 (1.8)
Palmer (1983)	4.0	3.0	3.2	1.5	2.4 (1.2)	3.4 (1.0)
Snelling (1985)	5.8	4.9	5.0	1.6	1.9 (1.3)	2.3 (1.0)

Fits are measured as the standard deviation of differences (in Ma) between observed and predicted positions of the peaks, with the peaks placed at the tops (*u*), middles (*m*), and bottoms (*l*) of the stages in which they occur. Numbers in parentheses indicate fits that result if events are more realistically assumed to be confined to the last 2.5 Ma of the Cenomanian and Maastrichtian (see Kauffman 1986). The modified version of the Harland *et al.* time scale is explained in the caption to Fig. 2.

performed by Sepkoski & Raup (1986a), however, so that any periodic components in the time scale could be directly assessed in the randomization test.) It can be shown that the 49-interval time scale used here for the generic data, with several longer stages subdivided, displays no tendency toward embedded periodicity (Raup & Sepkoski 1988; Stigler & Wagner 1988), validating the tests of Raup & Sepkoski (1986), who used a slight variant of these data.

Statistical methods

Various technical aspects of the statistical tests conducted by Raup & Sepkoski have been criticized. Noma & Glass (1987) and Quinn (1987) have presented extended, and in places vituperative, critiques, and Stigler & Wagner (1987), as summarized above, have questioned the power of the randomization test.

Noma & Glass argued that the procedure by which Raup & Sepkoski (1986) (and also Sepkoski & Raup 1986a) selected 'significant' extinction peaks for analysis was flawed on two counts: (1) the procedure should have compared adjacent points in the time series and not closest maxima and minima, and (2) the significance level in the procedure should have adjusted for multiple comparisons. On the first count, it is unfortunate that Noma & Glass did not refer to the explicit rationale presented by Sepkoski & Raup (1986a, p. 10). They argued that peaks should be compared to local minima rather than adjacent stages because backward smearing of extinction records (Signor & Lipps 1982) can influence stages preceding extinction events and high turnover in the wake of mass extinctions can influence succeeding stages.

The second criticism, that Raup & Sepkoski should have adjusted their 95% significance level to account for 42 possible tests, is valid. Thus, Sepkoski & Raup (1986a) and Raup & Sepkoski (1986) merely eliminated demonstrably insignificant peaks of extinction from their analyses. But it cannot be claimed that the eight analyzed peaks were necessarily insignificant (or even that four of the eight might be insignificant); these peaks appear even more prominently in the generic data and most correspond to independently identified extinction events in the biostratigraphic record. I resent Noma & Glass's published insinuation that our evaluations of the data were done to bias the statistical tests and make the data appear more regular.

Quinn (1987) reviewed Raup & Sepkoski's (1984) randomization test and delivered three criticisms. First, he argued that randomization of data underestimates variance and that bootstrapping procedures, in which data are randomly sampled with replacement, should have been used. This, however, is not a fatal flaw since Connor (1986) corroborated the periodicity in the familial data using true bootstrapping procedures. Second, Quinn argued that randomization assumes extinction intensities from stage to stage are independent while in fact the familial time series displayed substantial autocorrelation. However, he ignored Raup & Sepkoski's (1984) explicit discussion of the Markovian properties of their time series and the fact that they based their test only on randomized time series (Quinn's 'pseudorecords') with the same number of peaks as in the observed series in order to accommodate autocorrelation. Furthermore, Stigler & Wagner (1987) found that autocorrelation was not a problem in the test and could not induce an artificial appearance of periodicity.

Quinn's third argument was that Raup & Sepkoski (1984) did not adjust their significance levels for multiple tests. (He ignored the explicit treatment in Raup & Sepkoski (1986)). Quinn claimed that because all integer period lengths from 12 to 60 Ma were tested, a total of 49 independent tests were conducted; thus, the significance level was not 0.99 but $(0.99)^{49} = 0.61$ (or $p \approx 0.39$). But there are two mistakes in this conclusion. First, Quinn used a lower confidence level than the level of 0.9999 computed by Raup & Sepkoski for the fit of the 26 Ma period. Second, and more important, he ignored Raup & Sepkoski's (1984) demonstration that tests over longer period lengths were not independent but strongly correlated (i.e. a significant test result at a period length P_i will tend to be accompanied by significant results at P_{i-1} and P_{i+1}). Thus, there were far fewer than 49 independent tests, although Raup & Sepkoski (1984) did not know precisely how many.

Tremaine (1986) solved this problem by demonstrating that the tests become independent in the frequency domain (frequency is the reciprocal of period length). He argued that the 12–60 Ma band examined by Raup & Sepkoski was equivalent to 18 independent frequencies so that the significance level adjusted for independent tests (given Tremaine's recalculation of the significance level of the 26 Ma period) was actually $(0.9974)^{18} = 0.954$, or $p < 0.05$. This result is at variance with Quinn's conclusion. (Tremaine went on to argue, however, that Monte Carlo tests of the eight extinction events identified by Sepkoski & Raup (1986a) did not display a significant periodicity when multiple testing was taken into account. But this result seems to have been sensitive to the ages Tremaine assigned to the events and especially to his assumption that the measure of goodness of fit in Raup & Sepkoski's randomization procedure was exactly proportional to period length. Raup & Sepkoski (1986) repeated Tremaine's test without this assumption and with several sets of ages for extinction events and obtained significant [$p < 0.05$] results for all but one set of ages, as discussed above.)

Quinn (1987) proposed an alternative method of testing for periodicity which he used to show no difference from randomness in data sets for both families and genera. This test involved measuring waiting times (or 'time lags') between extinction events rather than fitting periodic functions; the waiting times were then compared to expectations for random events (i.e. a broken-stick distribution) using the standard Kolmogorov–Smirnov test. Unfortunately, Quinn used several arbitrary definitions of mass extinction in his analyses: either stages with extinction intensities exceeding 75% of measured values in the data set or peaks of extinction exceeding the mean intensity after adjustment for declining extinction toward the Recent. He used the former definition because he saw no reason why mass extinctions could not occur in adjacent stages, although this ignores both the possible backward-smearing of extinction intensities due to sampling (Signor & Lipps 1982) and the biostratigraphic literature on the nature of mass extinctions.

There is a more serious problem with Quinn's testing procedure, however. His random test distribution was constructed with the same number of waiting times as measured from the data. This assumes that any periodicity will have precisely the same number of cycles in the data as the number of observed events (which, again, is an arbitrary number in Quinn's selection procedure). If a sine curve

were to pass Quinn's test, the number of points chosen for testing would have to equal the number of cycles in the time series; but if the sine curve were noisy with some small spurious peaks, Quinn's test could well fail to distinguish it from random data. (In general, tests based on waiting times seem to be sensitive to missing events, admixed aperiodic events, and spurious peaks. I once rejected Fischer & Arthur's (1977) claim of periodicity with a test based on waiting times because my data were missing several events in the periodic sequence (see Sepkoski 1982b).)

All of the statistical test procedures reviewed thus far, including the randomization procedure, have used a Poisson model of independent events as the test distribution for comparison to periodicity. An alternative distribution was used by Lutz (1987) in one of the more constructive critical assessments of periodicity. He argued that periodic functions should also be compared to distributions reflecting constrained episodicity in which events tend to have some minimum spacing resulting from an embedded recovery time. Distributions of this nature can be generated with a gamma function, which is given by the parameters the mean and standard deviation of waiting times between events. If the standard deviation and mean are equal, the gamma function is equivalent to a Poisson distribution, whereas if the standard deviation is less than 20% of the mean, the function is equivalent to a noisy periodicity with some wobble in cycle length. Between these extremes, the gamma function describes constrained episodicity without clocklike properties. Lutz used various gamma distributions to test Raup & Sepkoski's (1984) data. He found that the data rejected both Poisson distributions and gamma distributions with short recovery times (recovery times in the range implied by Stanley's (1986b) hypothesis); but Lutz also found that the data rejected regular periodicities with standard deviations less than 15% of mean cycle length. However, he could reject neither a noisier periodicity with a 26 Ma cycle length nor a constrained eposidicity with standard deviation less than 30% of mean waiting time. He concluded that the record of extinction events certainly was not random but still could reflect either a periodic forcing with a severe wobble or an episodic forcing with a long recovery time. These conclusions are important and Lutz's results need to be further investigated to determine their sensitivity to selection of extinction peaks (Lutz used 11) and to error of estimate in the time scale.

The question of catastrophism

Two principal conclusions can be drawn from the review above:
(1) As palaeontological data and stratigraphic resolution have improved from the familial to the generic data sets, so too has the appearance of periodicity in extinction.
(2) Although many arguments against periodicity and for randomness of extinction have been presented, all either are flawed or are negated by the new data.
Thus, periodicity remains a strong hypothesis for the ordering of extinction events over the last quarter billion years of Earth history, although there is still a problem of why no periodicity is evident among genera in the Palaeozoic (Sepkoski 1986a).

This conclusion constitutes a fundamentally descriptive statement. However, it does have one strong implication for the nature of extinction events: the periodic events cannot be independent of one another (Flessa et al. 1986). Independent events, produced by a multitude of ultimate causal processes, tend to be irregularly spaced in time, with considerable variation in the waiting times between them (Raup & Sepkoski 1986; Sepkoski 1986a). This is clearly not the pattern observed for Mesozoic–Cenozoic extinction events. Therefore, there must be some single ordering principal. While this conceivably could be some sort of endogenous lag or oscillation (Kitchell & Pena 1984; Stanley 1986b), the such models do not generate patterns consistent with the data or with our understanding of diversification, as argued above. The alternative, therefore, is a single, ultimate forcing agent that has either clocklike behavior or some form of constained episodicity (cf. Lutz 1987).

Note that the existence of a single ultimate cause does not necessitate that each extinction event has precisely the same chain of proximate causes or the same pattern of taxonomic selectivity. Two factors can make the biotic response to perturbation quite complex. Nonlinearities or threshold responses in the chain of causation can produce marked differences in proximate causes of extinction (sensu McLaren 1983) under different intensities of forcing. In impact scenarios, for example, differences in bolide size or velocity (i.e. energy) can produce very different effects on the Earth's surface and atmosphere; medium-sized bolides may have local kill areas and small, short-term climatic effects, whereas large bolides may blanket the Earth in dust and produce severe atmospheric disturbances. A second factor that may effect response is the situation of the biota and of Earth conditions at the time of perturbation. For example, perturbations at times of high provinciality or diverse tropical faunas with numerous rare, localized taxa may have very different biotic effects than at times of low provinciality and common, widespread taxa.

What the nature of the ultimate forcing agent of periodic extinction might be is still unknown. Raup & Sepkoski (1984) suggested that it might be extraterrestrial for the reasons outlined in the introduction to this paper: there are no known terrestrial mechanisms with cycle times around 26 Ma, and at least one event in the periodic series, the Maastrichtian mass extinction, is associated with strong evidence of an extraterrestrial impact. This suggestion was later strengthened by observation of a similar periodicity in ages of terrestrial impact craters (Alvarez & Muller 1984; Shoemaker & Wolfe 1986; Trefil & Raup 1987). Response to Raup & Sepkoski's hypothesis of periodicity was immediate, and a variety of possible solar system and galactic mechanisms were hypothesized (Davis et al. 1984; Rampino & Stothers 1984a, b; Schwartz & James 1984; Whitmire & Jackson 1984; Muller 1985; Whitmore & Matese 1985). However, all these potential mechanisms have been shown to be flawed (Shoemaker & Wolfe 1986; Tremaine 1986), and to date only two periodic events, the Maastrichtian and the Upper Eocene, have been demonstrated to be associated with positive (but still debated) evidence of extraterrestrial events (although the Cenomanian and Middle Miocene events have also been found to encompass iridium anomalies; see Kerr 1987).

Alternative suggestions of terrestrial driving mechanisms are also still rather speculative. These are based, in part, on two observations: (1) an arguable periodicity of around 30 Ma in the reversal frequency of the Earth's magnetic field (Raup 1985b; Pal & Creer 1986; Stothers 1986), suggesting

some kind of regularity in deep-Earth dynamics (although see Muller & Morris 1986); and (2) the association of several periodic extinction events with immense flow basalts, such as the Siberian traps, Deccan traps, and Columbia River basalts. Loper & McCartney (1986) and Loper et al. (1988) suggested that both phenomena could reflect quasiperiodic processes in the lower mantle. They hypothesized that the thermal layer at the base of the mantle increases in thickness over time as a result of heating from the core. When the thickness reaches a critical threshold, plumes are spawned that rise through the mantle. These are manifested at the surface in widespread volcanism and basaltic outpourings that may inject large quantities of particulates, sulphates, and carbon dioxide into the atmosphere, thereby perturbing climate and inducing extinction. Such events will not be strictly clocklike but rather episodic with long intervening lag times, since release of mantle plumes will thin the thermal layer and re-establish stability until the heated material again builds to critical thickness. Thus, this hypothesis is consistent with the kind of constrained eposidicity discussed by Lutz (1987) or even the nonlinear dynamics with internal quasiperiodicities discussed by Shaw (1987). However, alternative hypotheses still must be evaluated: release of mantle plumes could be stimulated by seismic shocks caused by large bolide impacts.

Most of the extraterrestrial scenarios assume that periodic extinction events would result from cometary impacts occurring during comet showers of 1 to 3 Ma duration (Alvarez 1986; Hut et al. 1987). If these impacts were large, extinctions should be fairly catastrophic and disastrous. Similarly, if periodicity were driven by widespread catastrophic volcanism, extinction events should also be fairly abrupt and disastrous. However, the generic data suggest that only a few of the periodic events could be considered true disasters for the global marine fauna. Table 3 lists observed magnitudes of generic extinction, measured as percent extinction over each sampling interval, for both the total and the filtered data sets illustrated in Fig. 2. These observed magnitudes are two to eight times larger than the corresponding percentages for families presented by

Sepkoski & Raup (1986a) and are even larger than some of the species-level estimates made by Hoffman (1986c).

The estimated magnitudes of species extinction listed in Table 3 were derived from the rarefaction curve in Fig. 1 using the method developed by Raup (1979). The smaller estimates are from the filtered generic data and are probably somewhat conservative, whereas the higher estimates are for the total data and are liberal. All estimates need to be compared to average background levels of species extinction, which have a parallel range of 10 to 17% per sampling interval for the Tertiary and Cretaceous and 15 to 33% for the Jurassic and Triassic. Thus, most of the periodic events enhance extinction by factors of only two to three above background, resulting in disappearance of one-quarter to one-third of marine species over 3 to 5 Ma. However, the larger events do seem more significant. Both the upper Norian and the Maastrichtian mass extinctions eliminated two-thirds to three-quarters of marine species, the latter probably occurring over less than 1 Ma at the end of the Cretaceous. And, consistent with Raup's (1979) estimate, the Upper Permian event seems to have eliminated more than 90% of species, which must have been devastating to the marine fauna. Thus, these three large events may still need special partial explanations.

There are two other patterns in the generic data that might be interpreted as evidence for a noncatastrophic agent of periodicity:

(1) Most of the peaks of extinction, especially for the filtered data in Fig. 2, have nearly equivalent amplitudes, well below the heights of the three major events.

(2) Most of the peaks have widths that span several stages, perhaps indicating that these reflect long-term oscillations rather than short-term events (although much of this pattern could reflect sampling errors of the kind discussed by Signor & Lipps (1982)).

These observations might suggest that there are at least two causative components to the observed pattern of extinction: a 26 Ma oscillation of the Earth's climates and/or oceans, such as posited by Fischer & Arthur (1977) and Fischer (1981), that causes increases in extinction rates (either

Table 3. *Intensities of generic extinction, measured from the data in Fig. 2, and of species extinction, estimated from the generic rarefaction curve in Fig. 1*

Stratigraphic interval	Generic extinction		Estimated species extinction
	filtered	total	
Middle Miocene	6 ± 1	10 ± 2	16 ± 2–24 ± 4
Upper Eocene	9 ± 1	15 ± 2	22 ± 2–35 ± 4
Maastrichtian	36 ± 2	47 ± 3	63 ± 2–77 ± 3
Cenomanian	14 ± 1	26 ± 4	33 ± 2–53 ± 6
Aptian	13 ± 1	19 ± 3	31 ± 2–41 ± 5
Upper Tithonian	13 ± 2	21 ± 4	31 ± 4–45 ± 7
Pliensbachian	14 ± 1	26 ± 3	33 ± 2–53 ± 4
upper Norian	36 ± 3	47 ± 5	63 ± 3–75 ± 4
Upper Permian	78 ± 1	84 ± 1	93 ± 1–95 ± 1

Intensities were measured as percent extinction (= numbers of extinctions times 100 divided by diversity). Standard errors for generic extinction were estimated using the method outlined by Raup and Sepkoski (1986); standard errors for species were estimated from the generic data using the rarefaction curve.

continuously or in high-frequency, 'stepwise' episodes) over long intervals of time; and independent agents or constraints on extinction, such as extraterrestrial impacts, volcanic episodes, and/or sealevel changes, that occasionally amplify the periodic oscillations when they operate at times of increasing extinction.

This proposal is clearly speculative and is meant only to demonstrate the possibility of compromise between catastrophic and noncatastrophic models permitted within the constraint of periodicity. The proposal is supported by precious little evidence and still begs the question of the ultimate cause of the long-term oscillations—they could still have either a terrestrial or extraterrestrial forcing. I suspect we are still a long way off from solving this problem, given our present ignorance of the detailed nature of each of the periodic extinction events.

Conclusions

The hypothesis that extinction events have recurred with a 26 Ma periodicity over the last quarter billion years has withstood a variety of tests and has been greatly strengthened by refinements in palaeontologic data. However, the hypothesis still tells us little about the cause of mass extinction or whether they are catastrophic or gradual. Answers to these questions must await increase in our knowledge of biological processes of species extinction and more comparative information, derived from detailed field studies, on the duration, internal configuration, taxonomic and environmental selectivity, biogeography, and geological correlates of each of the periodic extinction events.

References

ALVAREZ, W. 1986. Toward a theory of impact crises. *Eos,* **67,** 649–58.
—— & MULLER, R. A. 1984. Evidence from crater ages for periodic impacts on the Earth. *Nature,* **308,** 718–20.
BENTON, M. J. 1986. More than one event in the Late Triassic mass extinction. *Nature,* **321,** 857–61.
BOUCOT, A. J. 1986. Ecostratigraphic criteria for evaluating the magnitude, character and duration of bioevents. *In:* WALLISER, O. (ed.) *Global Bio-Events.* Springer-Verlag, Berlin, 25–45.
CONNOR, E. F. 1986. Time series analysis of the fossil record. *In:* RAUP, D. M. & JABLONSKI, D. (eds) *Patterns and Processes in the History of Life.* Springer-Verlag, Berlin, 119–48.
DAVIS, M., HUT, P. & MULLER, R. A. 1984. Extinction of species by periodic comet showers. *Nature,* **308,** 715–7.
FISCHER, A. G. 1981. Climatic oscillations in the biosphere. *In:* NITECKI, M. H. (ed.) *Biotic Crises in Ecological and Evolutionary Time.* Academic Press, New York, 103–31.
—— & ARTHUR, M. A. 1977. Secular variations in the pelagic realm. *In:* COOK, H. E. & ENOS, P. (eds) *Deep-Water Carbonate Environments.* Society of Economic Palaeontologists and Mineralogists, Special Publication, **25,** 19–50.
FLESSA, K. W., ERBEN, H. K., HALLAM, A., HSÜ, K. J., HÜSSNER, H. M., JABLONSKI, D., RAUP, D. M., SEPKOSKI, J. J., JR., SOULÉ, M. E., SOUSA, W., STINNESBUCK, W. & VERMEIJ, G. J. 1986. Causes and consequences of extinction. *In* RAUP, D. M. & JABLONSKI, D. (eds) *Patterns and Processes in the History of Life.* Springer-Verlag, Berlin, 235–57.
FOX, W. T. 1987. Harmonic analysis of periodic extinctions. *Palaeobiology,* **13,** 257–71.
GILINSKY, N. L. 1986. Was there 26-Myr periodicity of extinctions? *Nature,* **321,** 533–4.
GOULD, S. J. 1985. All the news that's fit to print and some opinions that aren't. *Discover,* November 1985, 86–91.
HALLAM, A. 1976. Stratigraphic distribution and ecology of European Jurassic bivalves. *Lethaia,* **9,** 245–59.

—— 1977. Jurassic bivalve biogeography. *Paleobiology,* **3,** 58–73.
—— 1984. The causes of mass extinction. *Nature,* **308,** 686–7.
—— 1986. The Pliensbachian and Tithonian extinction events. *Nature,* **319,** 765–8.
HARLAND, W. B., COX, A. V., LLWELLYN, P. G., PICKTON, C. A. G., SMITH, A. G. & WALTERS, R. 1982. *A Geologic Time Scale.* Cambridge University Press, Cambridge.
HARLAND, W. B., HOLLAND, C. H., HOUSE, M. R., HUGHES, N. F., REYNOLDS, A. B., RUDWICK, M. J. S., SATTERTHWAITE, G. E., TARLO, L. B. H. & WILLEY, E. C. 1967. *The Fossil Record.* Geological Society of London.
HARPER, C. W., JR. 1987. Might Occam's canon explode the Death Star?: A moving-average model of biotic extinctions. *Palaios,* **2,** 600–4.
HOFFMAN, A. 1985. Patterns of family extinction: dependence on definition and geologic time scale. *Nature,* **315,** 659–62.
—— 1986a. Mass extinctions, diversification, and the nature of paleontology. *Revista Española de Paleontologia,* **1,** 101–7.
—— 1986b. Was there 26-Myr periodicity of extinctions? [Hoffman replies]. *Nature,* **321,** 535–6.
—— 1986c. Neutral model of Phanerozoic diversification: implications for macroevolution. *Neues Jahrbuch für Geologie und Paläontologie Abhandlungen,* **172,** 219–44.
—— & GHIOLD, J. 1985. Randomness in the pattern of 'mass extinctions' and 'waves of originations'. *Geological Magazine,* **122,** 1–4.
HOUSE, M. R. 1985. Correlation of mid-Paleozoic ammonoid evolutionary events with global sedimentary perturbations. *Nature,* **313,** 17–22.
HUT, P., ALVAREZ, A., ELDER, W. P., HANSEN, T., KAUFFMAN, E. G., KELLER, G., SHOEMAKER, E. M. & WEISSMAN, P. R. 1987. Comet showers as a cause of mass extinctions. *Nature,* **329,** 118–26.
KAUFFMAN, E. G. 1986. High resolution event stratigraphy: regional and global Cretaceous bioevents. *In:* WALLISER, O. (ed.) *Global Bio-Events.* Springer-Verlag, Berlin, 279–336.
KERR, R. A. 1987. Beyond the K–T boundary. *Science,* **236,** 667.
KELLER, G. 1986. Stepwise mass extinctions and impact events: Late Eocene to Early Oligocene. *Marine Micropaleontology,* **10,** 267–93.
KITCHELL, J. A. & CARR, T. R. 1985. Nonequilibrium model of diversification: faunal turnover dynamics. *In:* VALENTINE, J. W. (ed.) *Phanerozoic Diversity Patterns: Profiles in Macroevolution.* Princeton University Press, Princeton, 277–309.
—— & ESTABROOK, G. 1986. Was there 26-Myr periodicity of extinctions? *Nature,* **321,** 534–5.
—— & PENA, D. 1984. Periodicity of extinctions in the geologic past: deterministic versus stochastic explanations. *Science,* **226,** 689–92.
LOPER, D. E. & McCARTNEY, K. 1986. Mantle plumes and the periodicity of magnetic field reversals. *Geophysical Research Letters,* **13,** 1525–8.
——, —— & BUZYNA, G. 1988. A model of correlated episodicity in magnetic-field reversals, climate, and mass extinctions. *Journal of Geology,* **96,** 1–16.
LUTZ, T. M. 1985. The magnetic reversal record is not periodic. *Nature,* **317,** 404–7.
—— 1987. Limitations to the statistical analysis of episodic and periodic models of geologic time series. *Geology,* **15,** 1115–7.
McKINNEY, M. L. 1987. Extinction periodicity through random processes: the 'rebound effect'. *Geological Society of America Abstracts with Program,* **19**(7), 767.
McLAREN, D. J. 1983. Bolides and biostratigraphy. *Geological Society of America Bulletin,* **94,** 313–24.
—— 1984. Abrupt extinctions. *Terra Cognita,* **4,** 27–32.
MOORE, R. C., TEICHERT, C. & ROBINSON, R. A. (eds) 1953–1983. *Treatise on Invertebrate Paleontology.* Geological Society of America and University of Kansas Press, Lawrence, Kansas.
MULLER, R. A. 1985. Evidence for a solar companion star. *In:* PAPAGIANNIS, M. D. (ed.) *The Search for Extraterrestrial Life: Recent Developments.* Reidel, Dordrecht, 233–43.
—— & MORRIS, D. E. 1986. Geometric reversals from impacts on the Earth. *Geophysical Research Letters,* **13,** 1177–80.
NOMA, E. & GLASS, A. L. 1987. Mass extinction pattern: result of chance. *Geological Magazine,* **124,** 319–22.
ODIN, G. S. (ed.) 1982. *Numerical Dating in Stratigraphy.* Wiley, Somerset, New Jersey.
PAL, P. C. & CREER, K. M. 1986. Geomagnetic reversal spurts and episodes of extraterrestrial catastrophism. *Nature,* **320,** 148–50.
PALMER, A. R. 1983. The Decade of North American Geology 1983 geologic time scale. *Geology,* **11,** 503–4.
PATTERSON, C. & SMITH, A. B. 1987. Is the periodicity of extinctions a taxonomic artefact? *Nature,* **330,** 248–51.
QUINN, J. F. 1987. On the statistical detection of cycles in extinctions in the marine fossil record. *Paleobiology,* **13,** 465–78.
RAMPINO, M. R. & STOTHERS, R. B. 1984a. Terrestrial mass extinctions,

cometary impacts and the Sun's motion perpendicular to the galactic plane. *Nature,* **308,** 709–12.

—— & —— 1984*b.* Geological rhythms and cometary impacts. *Science,* **226,** 1427–31.

RAUP, D. M. 1979. Size of the Permo-Triassic bottleneck and its evoutionary implications. *Science,* **206,** 217–8.

—— 1985*a.* Mathematical models of cladogenesis. *Paleobiology,* **11,** 42–52.

—— 1985*b.* Magnetic reversals and mass extinctions. *Nature,* **314,** 341–3.

—— 1986. *The Nemesis Affair.* W. W. Norton & Co., New York.

—— & SEPKOSKI, J. J., JR. 1984. Periodicity of extinctions in the geologic past. *Proceedings of the National Academy of Sciences USA,* **81,** 801–5.

—— & —— 1986. Periodic extinction of families and genera. *Science,* **231,** 833–6.

—— & —— 1988. Testing for periodicity of extinction. *Science,* **241,** 94–96.

RAUP, D. M., GOULD, S. J., SCHOPF, T. J. M. & SIMBERLOFF, D. S. 1973. Stochastic models of phylogeny and the evolution of diversity. *Journal of Geology,* **81,** 525–42.

ROMER, A. S. 1966. *Vertebrate Paleontology,* 3rd ed. University of Chicago Press, Chicago.

SCHWARTZ, R. D. & JAMES, P. B. 1984. Periodic mass extinctions and the Sun's oscillation about the galactic plane. *Nature,* **308,** 712–3.

SEPKOSKI, J. J., JR, 1978. A kinetic model of Phanerozoic taxonomic diversity I. Analysis of marine orders. *Paleobiology,* **4,** 223–51.

—— 1982*a.* A compendium of fossil marine families. *Milwaukee Public Museum Contributions to Biology and Geology,* **51.**

—— 1982*b.* Mass extinctions in the Phanerozoic oceans: a review. *In:* SILVER, L. T. & SCHULTZ, P. H. (eds) *Geological Implications of Impacts of Large Asteroids and Comets on the Earth.* Geological Society of America Special Paper, **190,** 283–9.

—— 1984. A kinetic model of Phanerozoic taxonomic diversity III. Post-Palcozoic families and mass extinctions. *Paleobiology,* **10,** 246–67.

—— 1986*a.* Global bioevents and the question of periodicity. *In:* WALLISER, O. (ed.) *Global Bio-Events.* Springer-Verlag, Berlin, 47–61.

—— 1986*b.* Phanerozoic overview of mass extinction. *In:* RAUP, D. M. & JABLONSKI, D. (eds) *Patterns and Processes in the History of Life.* Springer-Verlag, Berlin, 277–96.

—— 1987. Is the periodicity of extinction a taxonomic artefact? Response. *Nature,* **330,** 251–2.

—— & RAUP, D. M. 1986*a.* Periodicity in marine extinction events. *In:* ELLIOTT, D. K. (ed.) *Dynamics of Extinction.* Wiley, New York, 3–36.

—— & —— 1986*b.* Was there 26-Myr periodicity of extinctions? *Nature,* **321,** 533.

SHAW, H. R. 1987. The periodic structure of the natural record, and nonlinear dynamics. *Eos,* **68,** 1651–65.

SHOEMAKER, E. M. & WOLFE, R. F. 1986. Mass extinctions, crater ages and comet showers. *In:* SMOLUCHOWSKI, R. S., BAHCALL, J. N. & MATTHEWS, M. S. (eds) *The Galaxy and the Solar System.* University of Arizona Press, Tucson, 338–86.

SIGNOR, P. W., III & LIPPS, J. H. 1982. Sampling bias, gradual extinction patterns, and catastrophes in the fossil record. *In:* SILVER, L. T. & SCHULTZ, P. H. (eds) *Geological Implications of Impacts of Large Asteroids and Comets on the Earth.* Geological Society of America Special Paper, **190,** 291–6.

SLUTSKY, E. 1937. The summation of random causes as the source of cyclic processes. *Econometrica,* **5,** 105–47.

SNELLING, N. J. (ed.) 1985. *The Chronology of the Geological Record.* Geological Society, London, Memoir, **10.**

STANLEY, S. M. 1984. Marine mass extinctions: a dominant role for temperature. *In:* NITECKI, N. H. (ed.) *Extinctions.* University of Chicago Press, Chicago, 69–117.

—— 1986*a.* Anatomy of a regional mass extinction; Plio-Pleistocene decimation of the Western Atlantic bivalve fauna. *Palaios,* **1,** 17–36.

—— 1986*b. Extinction.* Scientific American Books, New York.

STIGLER, S. M. 1985. Terrestrial mass extinctions and galactic plane crossings. *Nature,* **313,** 159.

—— & M. J. WAGNER. 1987. A substantial bias in nonparametric tests for periodicity in geophysical data. *Science,* **238,** 940–5.

—— & —— 1988. Testing for periodicity of extinction (response). *Science,* **241,** 96–99.

STOTHERS, R. 1979. Solar activity cycle during classical antiquity. *Astronomy and Astrophysics,* **77,** 121–7.

—— 1986. Periodicity of the Earth's magnetic reversals. *Nature,* **322,** 444–6.

TREFIL, J. S. & RAUP, D. M. 1987. Numerical simulations and the problem of periodicity in the cratering record. *Earth and Planetary Science Letters,* **82,** 159–64.

TREMAINE, S. D. 1986. Is there evidence for a solar companion star? *In:* SMOLUCHOWSKI, R. S., BAHCALL, J. N. & MATTHEWS, M. S. (eds) *The Galaxy and the Solar System.* University of Arizona Press, Tucson, 409–16.

WESTERMANN, G. 1984. Gauging the duration of stages: a new appraisal for the Jurassic. *Episodes,* **7,** 26–8.

WHITMIRE, D. P. & JACKSON, A. A., IV. 1984. Are periodic mass extinctions driven by a distant solar companion? *Nature,* **308,** 713–5.

—— & MATESE, J. J. 1985. Periodic comet showers and Planet X. *Nature,* **313,** 36–8.