Earth and Planetary Science Letters 325–326 (2012) 100–107

Contents lists available at SciVerse ScienceDirect

Earth and Planetary Science Letters

journal homepage: www.elsevier.com/locate/epsl

To tune or not to tune: Detecting orbital variability in Oligo-Miocene climate records

Cristian Proistosescu [a,*], Peter Huybers [a], Adam C. Maloof [b]

[a] Department of Earth and Planetary Sciences, Harvard University, 20 Oxford Street, Cambridge, MA 02138, USA
[b] Department of Geosciences, Guyot Hall, Princeton University, Princeton, NJ 08544, USA

ARTICLE INFO

Article history:
Received 25 April 2011
Received in revised form 10 January 2012
Accepted 19 January 2012
Available online xxxx

Editor: P. DeMenocal

Keywords:
paleoclimate
Milankovitch
spectral analysis
orbital tuning
astrochronology
geomagnetic polarity time scale

ABSTRACT

We address the problem of detecting quasi-periodic variability at orbital frequencies within pre-Pleistocene climate records using depth-derived and orbitally tuned chronologies. Many studies describing orbital variability in pre-Pleistocene sediment hosted isotope records employ climatic records that are set on orbitally tuned chronologies, without accounting for the bias in spectral power estimates introduced by orbital tuning. In this study we develop a method to quantify the effects of tuning upon spectral estimates and, in particular, to more properly determine the statistical significance of spectral peaks associated with orbital frequencies. We apply this method to two marine sediment $\delta^{18}O$ records spanning the Oligo-Miocene, from ODP cores 1090 and 1218. We find that using linear age–depth relationships reveals statistically significant spectral peaks matching eccentricity in core 1090, and obliquity and precession in core 1218, where the last appears most significant. Tuning the chronologies to the orbital solutions of Laskar et al. (2004) increases the statistical significance of the precession peak, whereas the obliquity and eccentricity peaks become indistinguishable from those expected from tuning noise. This result can be understood in that tuning records with high signal to noise ratios tends to lead to more significant spectral peaks, whereas a linear age–depth relationship is better suited for detecting peaks when signal to noise ratios are low. We also demonstrate this concept using synthetic records.

1. Introduction

The presence of a response to orbital variations, in so much as it can be objectively and confidently identified in paleoclimate records, provides valuable insight into the climate system. As a response to a known forcing, orbital variability in paleoclimate records is useful in understanding the sensitivity of the climate system to changes in radiative forcing. Additionally, since we are able to compute astronomical configurations and insolation curves over the entire Paleogene, correlation of sediment-hosted signals and orbital curves provides for the possibility of constructing accurate orbital chronologies (Hinnov, 2004; Pälike et al., 2006a; Westerhold et al., 2008), insomuch as the orbital solutions are themselves well constrained (Laskar et al., 2004).

Variability at orbital frequencies, also called Milankovitch frequencies, was first demonstrated in late-Pleistocene marine sediment cores in the obliquity and precession bands (Hays et al., 1976), and, later, in the obliquity band for the early Pleistocene (Ruddiman et al., 1986). The quality and quantity of Pleistocene Proxy records have permitted for cycle by cycle identification of the orbital imprint in climate records, and orbitally tuned chronologies have been developed by matching these signals against astronomical solutions calculated back in time (Lisiecki and Raymo, 2005). These astrochronologies have been shown to be in close agreement with ages obtained from depth-derived age models based on stacked records (Huybers, 2007; Lisiecki and Raymo, 2007), as well as radiometric measurements (Shackleton et al., 1990).

Pre-Pleistocene records also have been shown to contain significant orbital band variability, absent any orbital tuning. Records obtained during Ocean Drilling Program (ODP) Leg 154, on the Ceara Rise (Curry et al., 1995) have been shown to have significant obliquity and eccentricity band variability (Weedon et al., 1997; Zachos et al., 1997). Furthermore, Sexton et al. (2011) demonstrated significant variability at precession, obliquity and eccentricity frequencies in an Eocene aged magnetic susceptibility record from ODP site 1258.

However, many records have been evaluated for orbital behavior only after some amount of tuning is performed. Subsequent work on the Ceara Rise records, for example, used the age model of Shackleton et al. (1999), who calibrated magnetic susceptibility first to an orbital curve consisting mostly of obliquity with a small precession component, and then to the 405 kyr eccentricity. Based on this age model, the spectra for various oxygen and carbon isotope records from sites 926 and 929 were recomputed (Pälike et al., 2006a; Paul et al., 2000; Zachos et al., 2001), and peaks at orbital frequencies were found to be significant when compared against a red noise background (Mann and Lees, 1996). Although the possibility of excess spectral power resulting from tuning was noted (Paul et al., 2000), the influence of orbital tuning upon such spectral estimates has not been formally quantified.

* Corresponding author. Tel.: +1 617 384 5345.
E-mail address: cproist@fas.harvard.edu (C. Proistosescu).

0012-821X/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.epsl.2012.01.022

More recent studies have focused on records from ODP site 1090 on the Agulhas ridge in the south Atlantic (Billups et al., 2002, 2004), and from ODP site 1218 in the equatorial Pacific (Pälike et al., 2006b; Wade and Pälike, 2004). Both site 1090 (Channell et al., 2003) and site 1218 (Lanci et al., 2005) have a complete magnetostratigraphic record, containing all major magnetic reversals within the interval spanned by the record. Age models were obtained by assigning the magnetic reversal ages based on the polarity time scale and then refined by successive tuning of physical properties and stable isotopes, to an orbital curve consisting of eccentricity, obliquity and precession. Significance levels in the form of confidence intervals were computed for the spectral power estimates of tuned records (Pälike et al., 2006b), but again not accounting for the fact that the records have been tuned.

It is useful to distinguish between the problem of constraining the chronology of a record using orbital information and the problem of objectively identifying orbitally driven components in the variability. Here we focus only on the latter, using two approaches to this problem, that we expect to be broadly applicable to other orbital detection problems in geological records. The first approach is in the spirit of Hays et al. (1976): a purely depth-derived age model is used and the spectrum is evaluated for significant concentrations of energy in the orbital bands. The problem one quickly runs up against is that age uncertainty can easily disperse concentrations of spectral energy present at orbital frequency (Perron and Huybers, 2009). There is an open question whether, for such uncertain records, variability may be more readily detected using limited orbital assumptions when constructing an age model. In our second approach, we explore the use of orbital tuning in detecting orbital variability, through using a Monte Carlo method to account for the bias that tuning introduces in spectral estimates.

2. Age models

We analyze two marine sediment δ^{18}O records from ODP site 1090 (Billups et al., 2002) and ODP site 1218 (Wade and Pälike, 2004) as case studies for testing for Milankovitch signals in pre-Pleistocene climate records. We selected these records for several reasons. First, we required δ^{18}O records with high enough sampling resolution such that variability at orbital frequencies can be spectrally resolved, if present. For the ~40 kyr obliquity and ~20 kyr precession periods sampling intervals smaller than ~20 kyr and ~10 kyr, respectively, are required by the Nyquist theorem. Given distortions of high frequency terms down to about half the Nyquist frequency due to interpolation (Rhines and Huybers, 2011), even higher sampling resolutions are preferred. Site 1090 spans the late Oligocene to early Miocene (25–16.5 Ma) and is sampled on average every 5.8 kyr. The high resolution subset of site 1218, described in Wade and Pälike (2004), spans 30.0 to 26.5 Ma and has an average sampling rate of 4.3 kyr. Thus, both records have sufficiently high resolutions and are long enough to encompass several cycles of even the longer Milankovitch cycles, such as the 405 kyr eccentricity. Second we require cores from which magnetostratigraphic records are available, such that a purely depth-derived age model independent of orbital assumptions can be built based on correlation of magnetic polarity reversals to the geomagnetic polarity time scale. Both sites yielded complete magnetostratigraphic polarity records (Channell et al., 2003; Lanci et al., 2005), and the starting point for our age models is the Cenozoic geomagnetic polarity time scale.

Within the last two decades, the geomagnetic polarity time scale has seen several adjustments and revisions. However, in all its present incarnations it is based on the work of Cande and Kent (1992, 1995) who obtained a composite geomagnetic polarity sequence of normal and reversed magnetic polarity intervals across the South Atlantic ocean floor and computed the mean relative distances between reversals along an average direction of seafloor motion. A time scale was then obtained using nine absolute age calibration points, assuming quasi-uniform spreading rates in between each pair of calibration points, and performing a cubic spline interpolation against distance on the flow line to obtain a distance-to-age relationship for the magnetic reversals (Fig. 1).

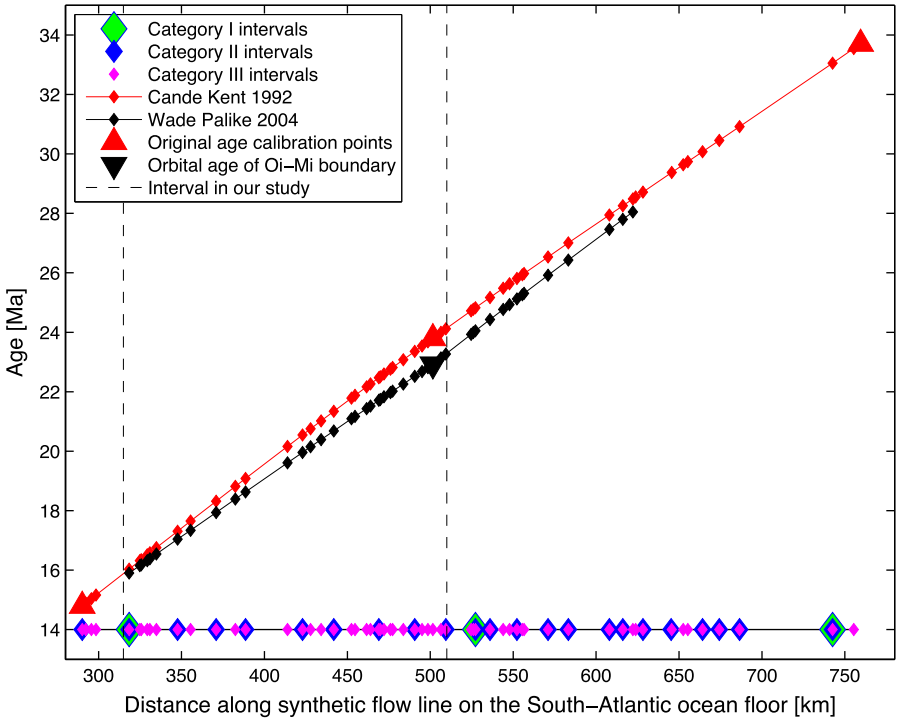

Fig. 1. Geomagnetic Polarity Time Scale (GPTS). The horizontal line displays the synthetic flow line in the South Atlantic. The average distances between reversals were calculated using a combination of finite rotation poles and stacked magnetic profiles, under the assumption that sea-floor spreading rates varied slowly in time and were normally distributed in space (Cande and Kent, 1992). The large triangles are the radiometric age control points. For comparison, the original GPTS of Cande and Kent (1992) is compared with the orbitally tuned version of Wade and Pälike (2004). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

In the original study, the calibration points were radiometric ages correlated with the magnetic anomaly sequence through the magnetostratigraphy of marine sediments overlying ocean crust (Cande and Kent, 1992). However, more recent studies (Billups et al., 2004; Pälike et al., 2006a; Wade and Pälike, 2004) use a calibration point at the Oligo-Miocene boundary that was astronomically derived (Shackleton et al., 1999). This age was chosen to maximize the similarity between the Oligo-Miocene climate signal and orbital parameters, making any studies using the astronomically calibrated geomagnetic polarity timescales somewhat biased towards showing an orbital relationship.

Starting from the ages of magnetic reversals in the untuned polarity time scale, we build purely depth-derived age models for the δ^{18}O records. Age calibration points are assigned to the magnetic reversals, with ages linearly interpolated with mean composite depth between these calibration points to obtain a depth-to-age curve for the entire length of the core. We asses the depth-to-age relationship for two versions of the age model, each based on a slightly different geomagnetic polarity timescale (Cande and Kent, 1992; Huestis and Acton, 1997). Both versions are independent of orbital assumptions and differ only in their estimation of the average position of magnetic reversals on the ocean floor.

Quantifying the uncertainties in the age model poses several difficulties. The calibration point at the Oligo-Miocene boundary has a chronogram estimated age of 23.8 Ma $\pm$ 1 Myr (Harland, 1990). Uncertainties in the lengths of the intervals between reversals on the sea floor profile are also significant (Cande and Kent, 1992). The larger intervals are estimated from fairly small sample sizes, ranging from 5 to 9 profiles, and with relative errors ranging between 4% and 17% at two standard deviations, assuming normality. The smaller sub-intervals often come from one or two measurements, with no estimated uncertainties. Thus, the errors in the timing of each magnetic reversal are a compound of errors in both the age calibration and the interval lengths, and they will accumulate for intervals in between two radiometric age calibration points. These magnetic reversals are in turn used as age calibration points for depth-derived age models, so the errors propagate through several layers of interpolation. Estimating all these compounded errors is beyond the scope of this paper, but we will explore the sensitivity of the results to the selection of age models.

3. A statistical test for orbital signals

3.1. Purely depth-derived age models

Using the age models based on untuned polarity time scales, we perform a statistical test for the presence of quasi-periodic variability at orbital frequencies within Oligo-Miocene δ^{18}O records. We start by using the Cande and Kent (1992) version, since it is the standard used in the literature. Our test is based on searching for statistically significant concentrations of spectral energy at the orbital frequencies. The records are re-interpolated on equally sampled grids, at the average time step, and normalized to zero mean and unit variance. Spectral power is computed using a periodogram estimator.

The statistical significance of spectral peaks is judged against the null hypothesis that the data simply represent random stochastic variability. This approach is similar to that of Hays et al. (1976), who tested the null hypothesis that the data represent a random signal having a general distribution of variance similar to the one observed in their low-resolution spectrum. We term the entire background variability represented by the broad-band continuum of spectral power as noise in order to distinguish it from potentially orbital signals. In other contexts of course this other variability may be considered as a signal as it includes unforced random climate variability due to internal processes, and possible nonlinear responses to orbital variability. The background variability is modeled here as an autoregressive process of order 1, or AR(1), which results in a spectrum that starts out as red noise at high frequency and tapers to white noise at low frequency, in a manner similar to the spectrum of the proxies. This type of spectrum is typical for a system with finite memory, where random disturbances are dissipated from the system over a finite time scale (Roe, 2009). In discrete time, an AR(1) process is described by,

$$X_t = \phi X_{t-\Delta t} + \in_t,$$

where ϕ is the autocorrelation coefficient and $\in$ represents a normally distributed random disturbance with mean zero. We fit the AR(1) process to the observed spectrum using the algorithm detailed in Neumaier and Schneider (2001) and Schneider and Neumaier (2001). For ODP 1090 the estimated autocorrelation coefficient is $\phi = 0.73$ and the variance of the $\in$ disturbances is 0.47 at a time step of 5.8 kyr. For ODP 1218 the estimated autocorrelation coefficient is $\phi = 0.87$ and the variance of the $\in$ disturbances is 0.41 at a time step of 4.3 kyr.

A Monte Carlo approach is then employed to numerically estimate confidence intervals for our hypothesis test. We generate 10,000 realizations of the AR(1) process used to model the background noise, and from these realizations we estimate the distribution of spectral energy at each frequency in the spectrum of the noise. The 95% confidence interval is set so that, at each frequency, 95% of the realizations of the AR(1) process have energy below this level. Thus, if a peak in the spectrum of the δ^{18}O record rises above the confidence interval, then the null hypothesis that the peak is attributable to background noise can be rejected with 95% confidence. We choose this Monte Carlo approach because we are interested in having individual realizations of the null hypothesis that will later allow us to directly quantify the bias that tuning introduces into the background.

Narrow band concentrations of spectral energy are observed at 1/445 kyr^{-1} in the ODP 1090 data (Fig. 2) and at 1/23 kyr^{-1} and 1/41 kyr^{-1} in the ODP 1218 data (Fig. 3). The 1/445 peak is near, but not exactly at, the Oligo-Miocene eccentricity frequency of 1/405 kyr^{-1} (Laskar et al., 2004), and the slight frequency offset could be a consequence of the age-uncertainties. By the frequency shift theorem (Bracewell, 2000), errors in the estimated duration of the record are inversely proportional to shifts in the frequency, $f_{estimated} = f_{true} \cdot T_{true}/T_{estimated}$. The uncertainty in the total length of the record is at least as large as the 1 Myr (1σ) uncertainty in the age of the radiometric calibration point at the Oligo-Miocene transition that is used in the calibration of both age models (Fig. 1). Given a $\pm$1 Myr error in the total length of the record, we would expect a 1/405 kyr^{-1} eccentricity peak to surface anywhere in a frequency band between 1/456 kyr^{-1} and 1/363 kyr^{-1}. The multitude of precession peaks near ~1/20 kyr^{-1} makes it difficult to ascertain whether a frequency offset is present, and the obliquity peak is consistent with no stretching or squeezing in ODP 1218.

The 445 kyr peak (ODP 1090, Fig. 2a) and both the 41 kyr and 23 kyr peaks (ODP 1218, Fig. 3a) are above the 95% confidence interval, but a few issues immediately arise in interpreting the implications of their significance. The first issue is that not all the peaks are exactly at the orbital frequency, which introduces a subtlety in the formulation of the null hypothesis. Confidence intervals are used to estimate the probability that a peak *protect at a particular frequency* is attributable to background noise. However, since we know that the peak might be shifted due to errors in the total interval length, we have searched for peaks in a rather broad frequency band, and there is a larger probability that background noise will produce at least one peak which is spuriously significant within this frequency interval as opposed to at a single frequency. For the 1/405 kyr^{-1} eccentricity peak, for example, we estimated that timing uncertainties could shift the peak to anywhere in the 1/456 kyr^{-1} and 1/363 kyr^{-1} interval. Thus, in our discrete spectra we have effectively

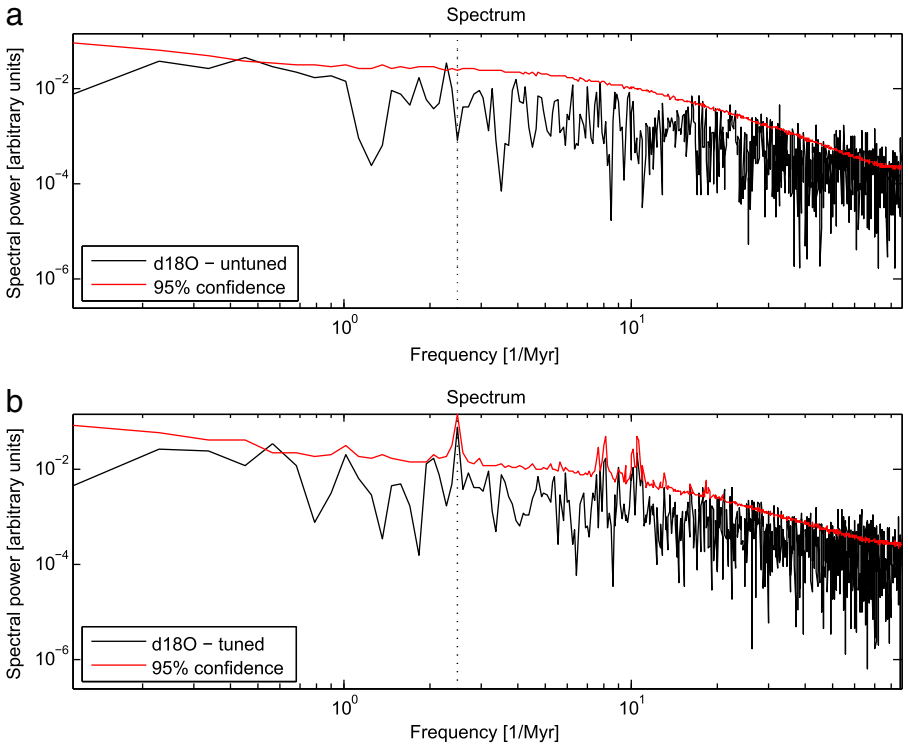

Fig. 2. Periodogram estimator of spectral power for the δ^{18}O record at ODP site 1090, for both untuned and tuned data. The red lines represent the 95% confidence interval, computed using a Monte Carlo method and assuming an AR(1) background noise. (*a*) Depth derived age model based on the Cande and Kent (1992) polarity time scale. Significant concentration of spectral energy is found near the 1/405 kyr^{-1} eccentricity frequency, denoted by the dotted line. (*b*) Orbital age model, with confidence intervals adjusted for the bias introduced through tuning the background noise to eccentricity. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

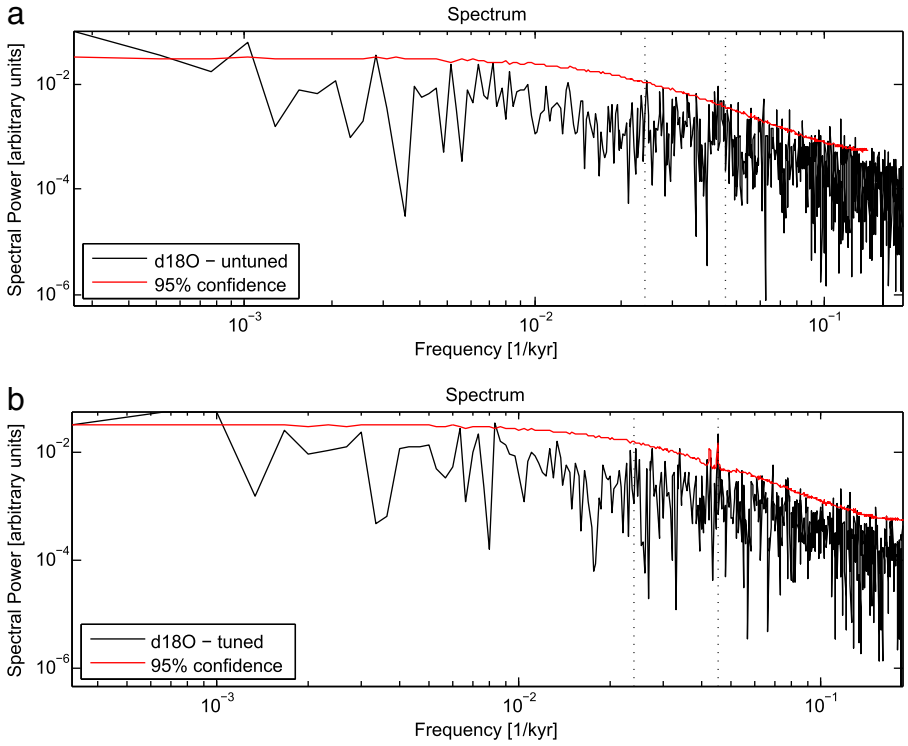

Fig. 3. Similar to Fig. 2, but for ODP 1218. (*a*) Depth derived age model based on the Cande and Kent (1992) polarity time scale. Significant concentrations of spectral energy are found near the 1/23 kyr^{-1} precession and 1/41 kyr^{-1} obliquity frequencies, denoted by the dotted lines. (*b*) Orbital age model, with confidence intervals adjusted for the bias introduced through tuning the background noise to precession. This shows that the precession peak is still significant even after tuning. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

extended our search to six different frequency bands. Reformulating the null hypothesis to account for the fact that we are searching for a peak among multiple frequencies, the significance levels of the 1/405 kyr^{-1} eccentricity peak and the 1/41 kyr^{-1} decrease to just under the 95% critical value, but the 1/23 kyr^{-1} precession remains significant at 95% confidence. Furthermore, the significance of the 445 kyr peak is sensitive to the choice of the age model. When we use an age model based on the slightly different polarity time scale of Huestis and Acton, (1997), the eccentricity band has no significant concentration of power, at any candidate frequency. The fact that the results of the statistical test are sensitive to the specific formulation of the null hypothesis and the choice of age models warrants the exploration of a different approach.

3.2. Tuned age models

We now asses the efficacy of using orbital tuning to test for orbital variability, which is interesting to explore two reasons. First, it may be that tuning allows for better identification of orbital behavior. Second, many other studies have tuned their records and then attempted to identify statistically significant peaks in the resulting spectral estimates, and we would like to better understand the effects of the tuning process.

Similar to previous studies, we start by assigning the records a linear age to depth relationship based on the polarity time scale, and then use an objective tuning algorithm to match the $\delta^{18}O$ record to each orbital curve in Laskar et al. (2004). Tuning is accomplished using a dynamical time warping algorithm (Lisiecki and Lisiecki, 2002) that chooses a depth to age relation for the proxy records that minimizes the sum of squared differences between the proxy and the orbital curve. The amount of age model adjustment the tuning algorithm is allowed to make can be controlled by introducing a parameter called the slope-weight, which penalizes deviations from the original age model (Pälike, 2002). The higher the slope-weight, the less tuning can deviate from the accumulation rate implied by the linear age–depth relationship.

Not surprisingly, after tuning, the periodograms of the $\delta^{18}O$ records show sharp peaks at the orbital frequencies (Fig. 2b, Fig. 3b). However, this result does not necessarily mean the peak is statistically significant. A record comprised solely of stochastic variability will exhibit sharp peaks in the spectral domain if subjected to the same tuning algorithm. Most realizations of the AR(1) process used to model the background noise spectrum of the $\delta^{18}O$ record, when tuned, exhibit sharp peaks that appear significant when contrasted to the spectral estimate of the untuned noise. Thus, revised critical values and 95% confidence intervals need to be estimated that account for the effects of tuning.

The revised null hypothesis is that peaks in a spectral estimate are due to broadband stochastic variability that has been tuned erroneously. Revised confidence intervals are then estimated via a Monte Carlo approach by taking realizations of the AR(1) process used to model background noise and subjecting them to the same tuning algorithm as the $\delta^{18}O$ record (Fig. 2b, Fig. 3b).

The behavior and results of this statistical test depend on the amount of adjustment the tuning algorithm is allowed to impose. With more adjustment (i.e., a lower slope-weight parameter) the orbital peak in the record becomes stronger, but so too becomes the probability of overtuning, or spuriously injecting a large amount of spectral power from the background variability (Muller and MacDonald, 2002). An optimal configuration of the tuning algorithm would ideally minimize the amount of spectral power that can be injected into a narrow band from the stochastic background continuum and, at the same time, maximize the amount of spectral power recovered from a distorted periodic signal.

For each orbital curve we tune to, we sweep across the full range of possible slope-weights and examine the p-value of the orbital

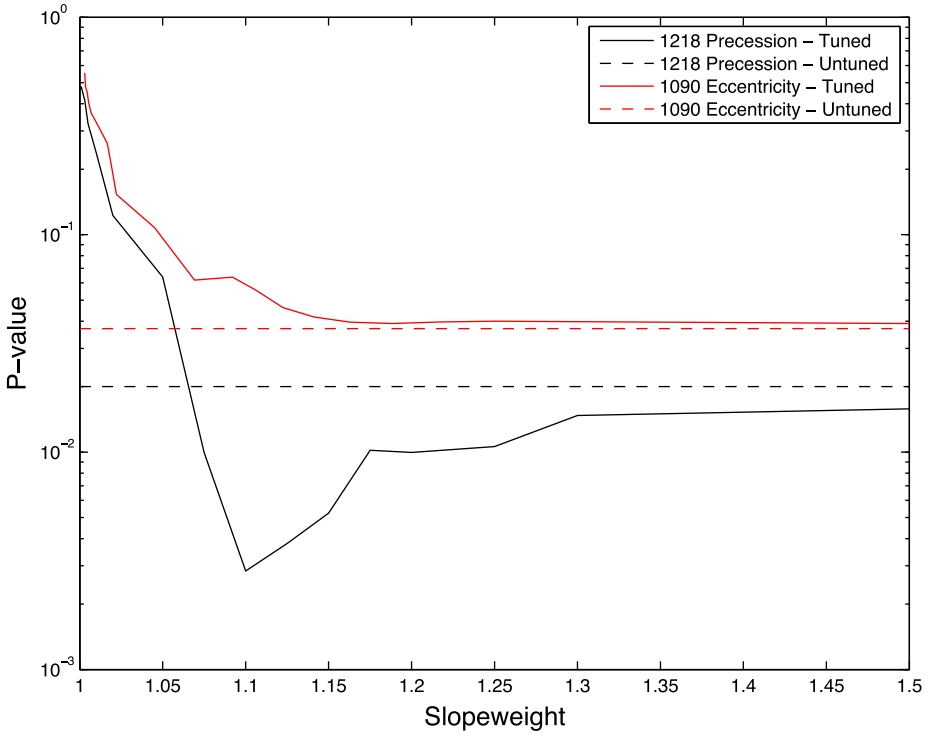

Fig. 4. p-Values for the spectral peaks at the 1/405 kyr^{-1} eccentricity frequency in ODP 1090 and 1/22 kyr^{-1} precession frequency in ODP 1218. The p-value of a spectral peak is equal to one minus its significance, and indicates the probability that a peak drawn from a sample of the null hypothesis distribution has spectral power equal or greater than the observed peak. The slope-weight of the tuning algorithm is a measure of the allowable age model adjustment. The higher the slope-weight, the smaller the adjustments that the tuning algorithm is allowed to make. As the slope-weight increases, the p-values will asymptote to the untuned case. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

peaks. Large values for the slope weight result in minimal tuning, while smaller values of the slope weight result in greater tuning. The p-value of a spectral peak is a measure of how high the tuned orbital peaks rise above the adjusted confidence intervals. By definition, it will represent the probability that a realization of the null hypothesis, tuned to the same orbital curve, will have equal or more spectral power than the tuned peak in the δ^{18}O record. As the amount of age model adjustment is minimized, the p-value of the peaks asymptotes to the untuned values (Fig. 4). We perform the tuning exercise for each of the three main orbital parameters, for both the ODP 1090 and ODP 1218 records. For the orbital frequencies where there was no significant concentration of spectral power using the untuned age model, no significant spectral peaks appear after tuning. The 1/405 kyr^{-1} and the 1/41 kyr^{-1} peaks have a similar behavior, in that for large to moderate amounts of tuning the spectral peaks are not significant. Apparently, the tuning algorithm injects more spurious power from the background noise than is recovered from the distorted orbital signal, assuming an orbital signal is present.

A different behavior is observed for the 1/23 kyr^{-1} peak where, for a moderate amount of tuning, the significance of the peak increases dramatically. The robustness of the results for precession across testing configuration lead us to confidently conclude that significant precession band variability is present in the ODP 1218 δ^{18}O record. This result is, to our knowledge, the first unbiased statistical test for orbital variability using orbitally tuned records.

3.3. Statistical power of the test

The foregoing method permits for a test of whether orbital variability is present using orbital tuning. However, the results of tuning are uneven, with the strongest results becoming more significant and others less so. This pattern may reflect the nature of the true signal or the characteristics of the test. To explore these possibilities, we turn to analyzing synthetic records to determine under what circumstances a tuned or untuned approach to testing for orbital variability is most powerful. The power of the test is a statistical measure of the separation of the null and alternate hypotheses. Specifically, it is the probability of rejecting the null hypothesis, H_0, when the alternate hypothesis, H_1, is true, and should not be confused with p, which is the probability that H_0 is wrong. As we will show, statistical power is also a useful metric for assessing if the significance of a spectral peak increases or decreases with tuning.

The same null hypothesis, H_0, and associated AR(1) process described in Section 3.1 are used here. For the alternate hypothesis, H_1, the record is modeled as a sum of AR(1) noise and an eccentricity curve (Laskar et al., 2004), placed on a corrupted age model. The parameters of H_1 are the signal-to-noise ratio, defined as the ratio of the variance of the orbital curve to the total variance of the AR(1) noise, and the jitter, defined as the variance growth rate of the age model uncertainty (Huybers and Wunsch, 2004). From Monte Carlo ensembles, we can estimate the probability density functions of spectral power in the orbital bands for H_0 and H_1. The power of the test is then directly computed as one minus the percentage of ensemble members of the H_1 that fall under the 95% confidence intervals derived from H_0 (Fig. 5a). When tuning is performed on realizations of H_0 and H_1, additional power is injected into the orbital bands, so both the tuned ensembles of H_0 and H_1 will have more spectral power. If the relative gain for H_0 is larger, then the power of the test decreases, and vice-versa (Fig. 5b).

The relative gain in spectral power of H_1 will depend on the signal-to-noise ratio and the amount of age model corruption. We perform a parameter space exploration over a reasonable range of jitter for

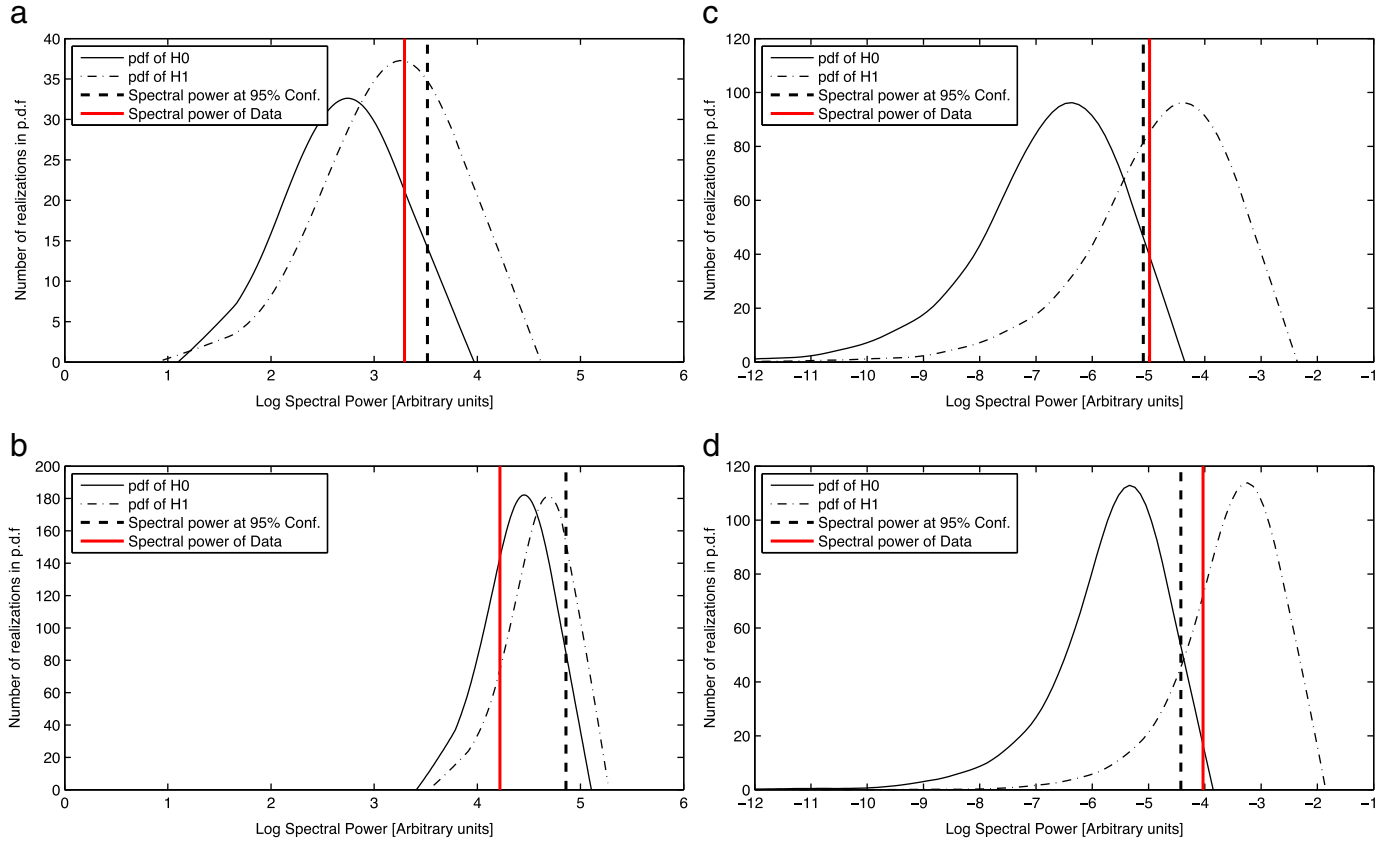

Fig. 5. Hypothesis test for the presence of orbital variability: Spectral power distribution of H_0 (solid), H_1 (dashed) and δ^{18}O (red). The parameters for the alternate hypothesis are a 0.5 signal-to-noise ratio and a 0.25 jitter. (*a*) ODP 1090, eccentricity, untuned H_0, H_1, δ^{18}O; (*b*) ODP 1090, eccentricity, tuned H_0, H_1, δ^{18}O; (*c*) ODP 1218, precession, untuned H_0, H_1, δ^{18}O; (*d*) ODP 1218, precession, tuned H_0, H_1, δ^{18}O. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

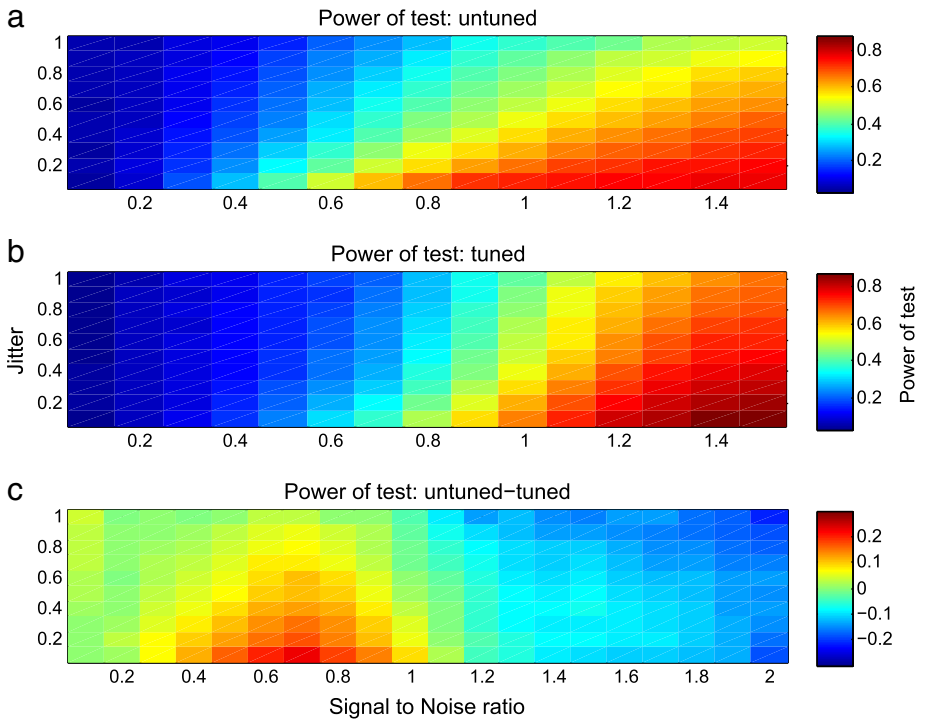

Fig. 6. The power of the test, i.e. one minus the probability of erroneously failing to reject the null. (*a*) For a test based on a purely depth-derived age model. (*b*) For a test based on an orbitally tuned age model. (*c*) The difference in statistical power between the two tests (untuned minus tuned). As expected, statistical power increases as the signal-to-noise ratio increases and as the amount of jitter decreases. The test based on a purely depth-derived age model performs better for signal-to-noise ratios below about one, while the test based on an orbitally tuned age model performs better for signal-to-noise ratios above one. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the sediment accumulation rate (Huybers and Wunsch, 2004), and the signal to noise ratios. Thus, for each of the two tests, statistical power is computed as a function of the signal-to-noise ratio and the amount of jitter (Fig. 6). We only examine what happens to an orbital signal with the 405 kyr eccentricity period, with an interval duration equal to the duration of the ODP 1090 record. For higher frequency signals, the power of the test will decrease more rapidly with jitter, as higher frequencies are more easily corrupted by age model uncertainties (Huybers and Wunsch, 2004; Perron and Huybers, 2009). As expected, there is a trade off between the signal-to-noise ratio and jitter: the power of the test increases as the signal-to-noise ratio increases or as the degree of jitter decreases.

The test for orbital forcing that uses a purely depth-derived age model has a higher statistical power if the signal to noise ratio is lower than about one, while the test based on an orbitally tuned age model has a higher statistical power for signal-to-noise ratios higher than about one, at least assuming a moderate degree of jitter. The suggestion is that orbital tuning can be used to sharpen periodic signals that are strong relative to the background variability, but tests based on orbitally tuned chronologies perform poorly in detecting signals in a predominantly noisy record, as evidenced by the increase in overlap between the distributions of the null and alternate hypothesis (Fig. 5). The behavior of the synthetic records is consistent with the results obtained for ODP 1090 and ODP 1218 where the only peak that increased in significance with adjusted confidence intervals, from a p-value of 0.02 to a p-value of 0.003, was the precession peak, which already had the lowest p-value of all three orbital peaks.

4. Conclusions

We have examined two different spectral domain tests for the presence of quasi-periodic orbital forcing in paleoclimate records inherently dominated by a broad-band distribution of variance, using two $\delta^{18}O$ records from the Oligocene and Miocene as case studies. The first test is based on a purely depth-derived age model, devoid of orbital forcing assumptions and employing a linear age-to-depth relation between magnetostratigraphic calibration points, while the second test is based on records that are orbitally tuned to orbital curves. Without accounting for the spectral energy introduced in the orbital bands from tuning the background, the null hypothesis would be erroneously rejected at the 95% significance level. The need to reformulate the null hypothesis and adjust confidence intervals when using astrochronologies is a major caveat that no previous study appears to have quantitatively accounted for. We have developed a Monte Carlo type methodology for estimating the statistical significance of spectral peaks in tuned records, and used it to perform an unbiased test for orbital variability in pre-Pleistocene climate records set on orbitally tuned chronologies. This methodology should be generally applicable to other orbitally tuned records.

Using untuned age models, significant spectral peaks were detected at eccentricity (ODP 1090) obliquity (ODP 1218) and precession (ODP 1218) frequencies. The use of orbitally tuned age models decreased the significance levels of the weaker eccentricity and obliquity peaks but increased the significance level of the stronger precession peak. Experiments on synthetic records are consistent with these results and suggest that tests using untuned age models are better at detecting periodic variability in noisier records, with a signal-to-noise ratio lower than about one, while orbital tuning may improve the test for distorted records that nonetheless have strong orbital components. Because the signal-to-noise ratio is difficult to estimate a priori, the use of both types of tests appears warranted.

Acknowledgments

The ODP 1090 data was provided by Katharina Billups. Frederik Simons and Alexander Stine provided useful comments, along with Peter deMenocal, Philip Sexton and two anonymous reviewers. ACM was founded by a Sloan Foundation Research Fellowship, PH

was supported by NASA grant 65P-1089456, and CP was supported by a grant from the Packard Foundation.

References

Billups, K., Channell, J.E.T., Zachos, J., 2002. Late Oligocene to early Miocene geochronology and paleoceanography from the subantarctic South Atlantic. Paleoceanography 17 (1) Jan., 010000-1.

Billups, K., Pälike, H., Channell, J.E.T., Zachos, J.C., Shackleton, N.J., 2004. Astronomic calibration of the late Oligocene through early Miocene geomagnetic polarity time scale. Earth Planet. Sci. Lett. 224, 33–44 Jul.

Bracewell, R., 2000. The Fourier Transform & Its Applications, 3rd Ed.

Cande, S.C., Kent, D.V., 1992. A new geomagnetic polarity time scale for the late Cretaceous and Cenozoic. J. Geophys. Res. 97 Sep., 13917–+.

Cande, S.C., Kent, D.V., 1995. Revised calibration of the geomagnetic polarity timescale for the Late Cretaceous and Cenozoic. J. Geophys. Res. 100, 6093–6095 Apr.

Channell, J., Galeotti, S., Martin, E., Billups, K., Scher, H., Stoner, J., 2003. Eocene to Miocene magnetostratigraphy, biostratigraphy, and chemostratigraphy at ODP Site 1090 (sub-Antarctic South Atlantic). Geol. Soc. Am. Bull. 115 (5), 607.

Curry, W.B., Shackleton, N.J., Richter, C., et al., 1995. Proc. ODP, Initial Reports, 154. Ocean Drilling Program, College Station, TX.

Harland, W., 1990. A Geologic Time Scale 1989. Cambridge Univ Press.

Hays, J.D., Imbrie, J., Shackleton, N.J., 1976. Variations in the earth's orbit: pacemaker of the ice ages. Science 194, 1121–1132 Dec.

Hinnov, L.A., 2004. Earth's orbital parameters and cycle stratigraphy. In: Ogg, J., Smith, A., Gradstein, F. (Eds.), A Geologic Time Scale 2004, pp. 55–62.

Huestis, S.P., Acton, G.D., 1997. On the construction of geomagnetic timescales from non-prejudicial treatment of magnetic anomaly data from multiple ridges. Geophys. J. Int. 129, 176–182 Apr.

Huybers, P., 2007. Glacial variability over the last two million years: an extended depth-derived age model, continuous obliquity pacing, and the Pleistocene progression. Quat. Sci. Rev. 26 (1–2), 37–55.

Huybers, P., Wunsch, C., 2004. A depth-derived Pleistocene age model: uncertainty estimates, sedimentation variability, and nonlinear climate change. Paleoceanography 19 1028–+.

Lanci, L., Parés, J., Channell, J., Kent, D., 2005. Oligocene magnetostratigraphy from equatorial Pacific sediments (ODP Sites 1218 and 1219, Leg 199). Earth Planet. Sci. Lett. 237 (3–4), 617–634.

Laskar, J., Robutel, P., Joutel, F., Gastineau, M., Correia, A., Levrard, B., 2004. A long-term numerical solution for the insolation quantities of the Earth. Astron. Astrophys. 428 (1), 261–285.

Lisiecki, L., Lisiecki, P., 2002. Application of dynamic programming to the correlation of paleoclimate records. Paleoceanography 17 (4), 1049.

Lisiecki, L.E., Raymo, M.E., 2005. A Pliocene–Pleistocene stack of 57 globally distributed benthic δ^{18}O records. Paleoceanography 20 (26) Jan., A261003 +.

Lisiecki, L., Raymo, M., 2007. Plio-Pleistocene climate evolution: trends and transitions in glacial cycle dynamics. Quat. Sci. Rev. 26 (1–2), 56–69.

Mann, M., Lees, J., 1996. Robust estimation of background noise and signal detection in climatic time series. Clim. Chang. 33 (3), 409–445.

Muller, R., MacDonald, G., 2002. Ice ages and Astronomical Causes: Data, Spectral Analysis and Mechanisms. Springer Verlag.

Neumaier, A., Schneider, T., 2001. Estimation of parameters and eigenmodes of multivariate autoregressive models. ACM Trans. Math. Softw. (TOMS) 27 (1), 57.

Pälike, H., 2002. Extending the geological calibration of the geological time scale. Ph.D. thesis, University of Cambridge.

Pälike, H., Frazier, J., Zachos, J.C., 2006a. Extended orbitally forced palaeoclimatic records from the equatorial Atlantic Ceara Rise. Quat. Sci. Rev. 25, 3138–3149 Dec.

Pälike, H., Norris, R.D., Herrle, J.O., Wilson, P.A., Coxall, H.K., Lear, C.H., Shackleton, N.J., Tripati, A.K., Wade, B.S., 2006b. The heartbeat of the Oligocene climate system. Science 314 Dec., 1894–.

Paul, H.A., Zachos, J.C., Flower, B.P., Tripati, A., 2000. Orbitally induced climate and geochemical variability across the Oligocene/Miocene boundary. Paleoceanography 15, 471–485.

Perron, J., Huybers, P., 2009. Is there an orbital signal in the polar layered deposits on Mars? Geology 37 (2), 155.

Rhines, A., Huybers, P., 2011. Estimation of spectral power laws in time uncertain series of data with application to the Greenland Ice Sheet Project 2 δ18O record. J. Geophys. Res. 116 (D1), D01103.

Roe, G., 2009. Feedbacks, timescales, and seeing red. Annu. Rev. Earth Planet. Sci. 37, 93–115.

Ruddiman, W.F., Raymo, M., McIntyre, A., 1986. Matuyama 41,000-year cycles: North Atlantic Ocean and northern hemisphere ice sheets. Earth Planet. Sci. Lett. 80, 117–129 Oct.

Schneider, T., Neumaier, A., 2001. Algorithm 808: ARfitA Matlab package for the estimation of parameters and eigenmodes of multivariate autoregressive models. ACM Trans. Math. Softw. (TOMS) 27 (1), 58–65.

Sexton, P., Norris, R., Wilson, P., Pälike, H., Westerhold, T., Röhl, U., Bolton, C., Gibbs, S., 2011. Eocene global warming events driven by ventilation of oceanic dissolved organic carbon. Nature 471 (7338), 349–352.

Shackleton, N., Berger, A., Peltier, W., 1990. An alternative astronomical calibration of the lower Pleistocene timescale based on ODP Site 677. Trans. R. Soc. Edinburgh Earth Sci. 81, 251–261.

Shackleton, N.J., Crowhurst, S.J., Weedon, G.P., Laskar, J., 1999. Astronomical calibration of Oligocene–Miocene time. Philos. Trans. R. Soc. London, Ser. A 357 Aug., 1907–+.

Wade, B.S., Pälike, H., 2004. Oligocene climate dynamics. Paleoceanography 19 (26) Dec., A264019 +.

Weedon, G., Shackleton, N., Pearson, P., 1997. The Oligocene time scale and cyclostratigraphy on the Ceara Rise, western equatorial Atlantic. 5 Proceedings of the Ocean Drilling Program: Scientific Results, Vol. 154, pp. 101–114.

Westerhold, T., Röhl, U., Raffi, I., Fornaciari, E., Monechi, S., Reale, V., Bowles, J., Evans, H., 2008. Astronomical calibration of the paleocene time. Palaeogeogr. Palaeoclimatol. Palaeoecol. 257 (4), 377–403.

Zachos, J.C., Flower, B.P., Paul, H., 1997. Orbitally paced climate oscillations across the Oligocene/Miocene boundary. Nature 388, 567–570 Aug.

Zachos, J.C., Shackleton, N.J., Revenaugh, J.S., Pälike, H., Flower, B.P., 2001. Climate response to orbital forcing across the Oligocene–Miocene boundary. Science 292, 274–278 Apr.