PALEOCEANOGRAPHY, VOL. 24, PA1204, doi:10.1029/2007PA001568, 2009

Change point method for detecting regime shifts in paleoclimatic time series: Application to $\delta^{18}O$ time series of the Plio-Pleistocene

Eric Ruggieri,[1] Tim Herbert,[2] Kira T. Lawrence,[3] and Charles E. Lawrence[1]

Received 24 October 2007; revised 24 October 2008; accepted 14 November 2008; published 28 January 2009.

[1] Although different paleoenvironmental time series resolve past climatic change at different time scales, nearly all share one characteristic: they are nonstationary over the length of the record sampled. We describe a recursive dynamic programming change point algorithm that is well suited to identify shifts in the Earth system's variability, as it represents a nonstationary time series as a series of regimes, each of which is homogeneous. The algorithm fits the data by minimizing squared errors not only over the parameters of the models for each subsequence but also over an arbitrary number of boundary points without restrictions on the lengths of regimes. The versatility of the algorithm is illustrated by an application to 5 Ma of Plio-Pleisotcene $\delta^{18}O$ variations. We seek to identify either the single dominant "Milankovitch" frequency or linear combinations of frequencies and consistently identify changes ~780 ka and ~2.7 Ma, among others, in each analysis done. Our applications also provide support to the recent hypothesis that obliquity-based Milankovitch terms can account for the circa 100 ka cycle that empirically dominates the most recent 1 million years.

Citation: Ruggieri, E., T. Herbert, K. T. Lawrence, and C. E. Lawrence (2009), Change point method for detecting regime shifts in paleoclimatic time series: Application to $\delta^{18}O$ time series of the Plio-Pleistocene, *Paleoceanography*, *24*, PA1204, doi:10.1029/2007PA001568.

1. Introduction

[2] A very large fraction of paleoclimatic research involves producing and analyzing time series. Instrumentation and laboratories have become progressively more efficient at generating long, well-resolved representations of the Earth's past climate. Examples include an effort to characterize global temperatures over the past millennium from a variety of proxies [*Briffa et al.*, 1995; *Esper et al.*, 2002; *Mann et al.*, 1998; *Jones and Mann*, 2004], millennium-long reconstructions of equatorial Pacific surface oceanography using stable isotopes from corals [*Cole et al.*, 1993; *Cobb et al.*, 2003; *Quinn et al.*, 1998], high-resolution stable isotope and trace gas records from polar ice cores [*Chappellaz et al.*, 1993; *Dansgaard et al.*, 1993; *Mayewski et al.*, 1993; *Petit et al.*, 1999] and records of glacial-interglacial climate cycles derived from ocean sediment cores [*Bloemendal and deMenocal*, 1989; *Imbrie et al.*, 1984, 1989; *Herbert and Mayer*, 1991; *Joyce et al.*, 1990; *Lisiecki and Raymo*, 2005; *Ruddiman et al.*, 1986].

[3] Although these different time series resolve past climatic change at different time scales, nearly all share one characteristic: they are nonstationary over the length of the record sampled. Changes routinely occur in the mean (trends), amplitude of variability, and in dominant spectral frequencies over the course of time [*Cobb et al.*, 2003; *Ghil et al.*, 2002; *Jones and Mann*, 2004; *Maasch*, 1988; *Mann*, 2004; *Paillard*, 1998; *Ravelo et al.*, 2004; *Rutherford et al.*, 2003; *Vautard and Ghil*, 1989]. One might then idealize paleoclimatic time series as a succession of a number of segments with internally homogeneous statistical properties, bounded by abrupt or gradual shifts to subsequent or antecedent regimes. Such a view might lead to progress in understanding climate evolution in a number of ways. It might help optimize sampling strategies, by allowing paleoclimatologists to improve the temporal resolution or spatial coverage of "model" segments of time, which represent much longer spans of Earth history adequately from a statistical point of view. Identifying the duration of "regimes" and determining the timing of their transitions might also lead to a more fundamental understanding of the underlying forces in the climate system that lead to the nonstationarity observed at many time scales. However, *Wunsch* [1999] cautions that purely random fluctuations of a stationary time series may yield a time series that appears nonstationary.

[4] Since we do not have theories that confidently predict when and how often the climate system went through transitions, the search at present is empirical. We introduce here the change point method to the study of one particular aspect of the Earth's past in which nonstationarity is already well known: the evolution of oxygen isotopes measured from benthic foraminifera, a proxy for high-latitude climate change (high-latitude ice sheet growth and decay, and variations in temperatures of deep water masses fed by the high-latitude surface oceans). We analyze a 5-Ma-long orbitally tuned, global stack of many individual records [*Lisiecki and Raymo*, 2005], as well as a shorter 2.5-Ma-long nonorbitally tuned stack [*Huybers*, 2007]. Much pre-

[1]Center for Computational Molecular Biology, Division of Applied Mathematics, Brown University, Providence, Rhode Island, USA.
[2]Department of Geological Sciences, Brown University, Providence, Rhode Island, USA.
[3]Department of Geology and Environmental Geosciences, Lafayette College, Easton, Pennsylvania, USA.

0883-8305/09/2007PA001568$12.00

vious work has established that the benthic δ^{18}O record contains a number of periodic components produced or paced by cyclic variations in the Earth's orbit. We also know from previous work that the oxygen isotope record contains at least two first-order transitions in behavior: the shift from earlier 41-ka-dominated glacial cycles to later 100-ka-dominated glacial cycles, the so-called "mid-Pleistocene transition", which occurred some time between 0.8 and 1.2 Ma [*Ruddiman et al.*, 1986; *Berger et al.*, 1993] and the earlier intensification of large-scale Northern Hemisphere glaciation at around 2.7 Ma [*Shackleton et al.*, 1984; *Raymo*, 1994]. We evaluate the strengths and weaknesses of the change point method in identifying both of these major changes in paleoclimatic behavior and in deducing more subtle aspects of climate evolution over the past 5 Ma. The results are sufficiently positive to suggest that the change point method may be applied to many other time series in which changes in behavior are suspected.

2. Background: Detecting Changes in Orbital Components in Plio-Pleistocene δ^{18}O Time Series

[5] Existing analytical methods stop far short of adequately modeling the complex, nonstationary story of environmental change. Most approaches have relied on modifications of conventional spectral analysis. The common procedure of parsing paleoclimatic time series into shorter time windows of fixed length, and analyzing these as a series of Fourier spectra [e.g., *Park and Herbert*, 1987; *Joyce et al.*, 1990; *Yiou et al.*, 1991; *Birchfield and Ghil*, 1993] is most directly relevant here. This method is sufficient to characterize intervals of geological time in terms of dominant frequencies, and to detect first-order changes in the relative sensitivity of past climate to the different orbital terms. However, because it uses moving windows of fixed size, its blurring of boundaries makes it inherently low resolution. Furthermore, sidebands of fundamental frequencies will appear during intervals of rapid (relative to the window length) spectral change, obscuring the underlying dynamics of the time series. Wavelet analysis [*Liu and Chau*, 1998; *Lau and Weng*, 1995; *Torrence and Compo*, 1998; *Bolton et al.*, 1995] has been applied to paleoclimatic time series as a means to better localize the spectral properties of nonstationary time series, but it too suffers from the appearance of sidebands during intervals of rapid change [*Bolton et al.*, 1995]. Its ability to illustrate contributions jointly in time and frequency domains is an asset, but it lacks an obvious procedure for objectively determining points of change in a time series for a single frequency or linear combinations of frequencies. Singular spectrum analysis (SSA), like principal component or empirical orthogonal function techniques, decomposes a time series into a series of orthogonal functions that can be ordered by the variance explained. SSA is especially useful for distilling a time series into trend (not necessarily linear), oscillatory, and noise components [*Ghil et al.*, 2002]. In common with wavelet analysis, SSA [*Vautard and Ghil*, 1989; *Vautard et al.*, 1992; *Ghil et al.*, 2002] allows the investigator to detect changes or breaks in the amplitude of oscillatory components. However, because the spectral resolution of oscillatory components is often rather broad [*Vautard and Ghil*, 1989], it is not especially suited for detecting regime changes in time series which are known or suspected to contain strong periodic components, as is the case for glacial-interglacial variations in orbitally resolved time series.

[6] Studies of the transition from 41 ka to 100 ka glacial cycles in the Pleistocene confirm the unsatisfactory state of break point determination at present: different approaches have variously put the onset at 1.5 Ma [*Rutherford and D'Hondt*, 2000], ~0.90 Ma [*Maasch*, 1988; *Raymo et al.*, 1997], or 0.64 Ma [*Mudelsee and Schulz*, 1997] and determined that it was both gradual [*Park and Maasch*, 1993] and abrupt [*Maasch*, 1988; *Mudelsee and Schulz*, 1997]. The unsettled debate may arise because the time series methods outlined above inadequately trade off frequency and temporal resolution in evolving time series. We analyze Plio-Pleistocene climate evolution using a flexible change point detection algorithm with sinusoidal models. The approach has aspects in common with *Tome and Miranda* [2004], who described a least squares fitting procedure for fitting linear models to paleoclimate time series. Tome and Miranda automate the creation of a matrix of over determined linear equations and consecutively solve for every combination of possible solutions that satisfies their constraints; in the end, they choose the solution that minimizes the sum of squared residuals. However, their work is limited to piecewise linear fits and the time complexity of the algorithm grows exponentially with the length of the time series, thus becoming computationally intensive with long series.

[7] In what follows, we describe an algorithm that removes all restrictions on window size. The spacing between breaks in behavior and the length of segments depends only on the underlying behavior of the time series and the model chosen for fitting the data. Specifically, we propose an algorithm that optimally partitions a time series at k "change points", subdividing a time series into (k + 1) subintervals, each of which is characterized by a specified function using parameter values that minimize the sum of square residuals. This least squares algorithm works in the time domain and optimizes over both the parameters of functions describing the subintervals and the locations of the change points, locations in the time series where significant changes in the amplitude, frequency, or phase occur. The algorithm will return the globally optimum solution (in terms of least squares) and has no possibility of returning merely locally optimum solutions. If the global optimum is unique, then the unique global optimum will be returned; if there are several equivalent (in the least squares sense) global optima, then only one of the equivalent global optima will be returned (the algorithm could easily be adjusted to return all equivalent optima). In the present analysis, our algorithm returns the optimal location of change points in a time series for a specified number of change points, but does not estimate the uncertainty of the location in time of change points nor in the values of the parameters. Uncertainty estimates can be obtained if the problem is instead framed probabilistically and Bayesian statistical procedures are employed [*Liu and Lawrence*, 1999].

[8] This algorithm approaches a problem that is a generalization of the problem addressed by *Tome and Miranda*

[2004], but differs from it in two important ways: (1) Dynamic programming [*Bellman*, 1957] is employed for optimization and (2) each subinterval is described by linear combinations of sinusoidal functions. For this particular implementation of the change point method, subintervals are characterized by linear combinations of sinusoidal functions of specified orbitally related frequencies with the amplitude and phase parameters for all frequencies chosen to minimize the sum of squared residual error. (The change point algorithm could easily be adapted to models other than sinusoidal, if appropriate.) We chose to focus on two principle ''Milankovitch'' frequencies, which emerge from the classic Milankovitch hypothesis that summer insolation at 65°N paces ice age cycles and whose presence in the data was confirmed by Fourier analysis. Cycles at periods of 23 and 41 ka represent the most important terms of climatic precession and obliquity, respectively. In addition, we wish to detect changes in the importance of the circa 100 ka cycle that dictates the timing of the large ice ages of the late Pleistocene [*Imbrie et al.*, 1984, 1992]. This period may derive in some nonlinear fashion from eccentricity modulation of the precessional amplitude, as bundles of 4 or 5 precessional cycles [*Ridgwell et al.*, 1999] or it may result from nonlinear responses to the obliquity cycle [*Tziperman and Gildor*, 2003; *Huybers and Wunsch*, 2005] as groups of two or three 41 ka cycles. Because the origin of the 100 ka cycle is so unclear, we treat its detection in several different ways in this study.

[9] We approached the problem of detecting change points in the 5-Ma-long δ^{18}O time series in steps of ascending complexity. In the first analysis, we defined break points by changes in the dominance or amplitude of the single best ''Milankovitch'' frequency that explains the most variance in each subinterval. We varied the number of change points from 2 (corresponding to well-known major shifts in the behavior of the oxygen isotope record at circa 0.9 and 2.7 Ma) to a larger number (up to 8) to explore how well the method fares against conventional analyses, and to explore more nuanced changes. Second, we increased the complexity of the model so that regimes are characterized by optimal linear combinations of the 23, 41, and 100 ka frequencies. In this setting, we compare results obtained from both orbitally tuned and nonorbitally tuned data sets. Last, we compared the ability of a model that represents the 100 ka glacial cycle as a combination of 82 and 123 ka integer multiples (subharmonics) of the obliquity cycle to capture circa 100 ka energy in the oxygen isotope time series, to the fit obtained with the classic assumption that the 100 ka component represents a fixed response to the eccentricity envelope to climatic precession. We interpret the timing and nature of segments defined by the change point method in the context of decades of work in studying Plio-Pleistocene glaciations, and point the way to future improvements in applying the change point method to paleoclimatic problems.

3. Methods

[10] Since we want to consider all possible locations of change points and lengths of intervening segments (i.e., climate regimes), change point analysis of the δ^{18}O time series presents a huge number of possible solutions. Dynamic programming provides a practical path through the forest of possibilities [*Bellman*, 1957]. It guarantees globally optimal solutions of problems that can be divided into progressively smaller sets of subproblems, the smallest of which can easily be solved. The method then builds solutions to a progressively growing set of subproblems until the original full problem has been solved. Dynamic programming has become a well established means of solving problems that can be partitioned in this way. It is described in many textbooks in operations research [*Hillier and Lieberman*, 2005, chapter 10; *Wagner*, 1975, chapters 8 and 10; *Winston*, 2003, chapter 18; *Lew and Mauch*, 2007]. This methodology has been widely and successfully applied in many fields, including extensive application in the emerging fields of genomics and bioinformatics [*Durbin et al.*, 1998, chapter 2; *Jones and Pevzner*, 2004, chapter 6; *Konopka and Crabbe*, 2004]. Dynamic programming was first applied to the change point problem by *Hawkins* [1976]. Change point analysis was first applied to biopolymer sequences by *Auger and Lawrence* [1989], and *Liu and Lawrence* [1999] show how Bayesian inference procedures allow for assessment of uncertainties in the unknown parameters in a change point analysis.

[11] In this paper we address the following problem: given a time series, in this case a benthic δ^{18}O time series, and a proposed set of models, we seek to minimize the sum of squared error over all possible sets of parameters and over all possible locations of change points. We impose the constraint that no two change points are within a distance equal to length of the longest cycle of each other. (This constraint arises so as to avoid aliasing of frequencies due to sparse data points in a time series.) For an initial choice of model (here, sinusoidal) and number of subintervals (change points), the algorithm produces the best possible fit to the entirety of the data. It allows us to relax the requirement of a fixed window that has been used in past paleoclimatic analyses and replace it with the ability to use windows of any size. The use of sinusoidal functions is not new. Fourier based methods depend upon the sinusoid and these functions have been shown to give a reasonable fit to the data.

[12] The algorithm has three steps: (1) parameter fitting, for each possible model (in this case, sinusoids) in each of the N(N + 1)/2 subintervals of a time series of length N; (2) the forward step, to find the minimum sum of square residuals considering all possible positions of k change points for the full time series and for all subintervals that begin at time zero; and (3) the backtrace step, to identify the solution (parameter set or argument) minimizing the sum of square residuals, i.e., the optimal set of change points. We begin with a series of time points $t_n = 1,2,\ldots N$ that are not necessarily equally spaced. At each of these points, the measured value of the proxy of interest Y_{t_n} is available. We also assume that we have M different models that are proposed to fit the data. The following is a summary of the computer code, the so-called pseudo code, used in each of these steps.

[13] Step 1 is parameter fitting for M periodic functions with given frequency components:

for i = 1,. . .,N

for j = i + d,. . .,N

for m = 1,. . .,M

$$\boldsymbol{SS}_{i,j}^{m} = \underset{\substack{\alpha_0,\ldots,\alpha_p \\ \beta_1,\ldots\beta_p}}{\mathbf{Min}} \cdot \left\{ \sum_{t=t_i}^{t_j} \left(Y_t - \alpha_0 - \sum_{p=1}^{p} \alpha_p \sin(\omega_p t) + \beta_p \cos(\omega_p t) \right)^2 \right\}$$

end for

$$\boldsymbol{SS}_{i,j} = \mathrm{Min}\{SS_{i,j}^{1}, \ldots SS_{i,j}^{M}\}$$

end for

end for

where ω_p is a specified frequency, $SS_{i,j}$ is the sum of square residuals in the subinterval beginning at position t_i and ending at position t_j, and $d \geq 0$ is the size of the smallest permitted window. In cases where we want to consider M different, distinct models, the inner loop will allow the algorithm to find the minimum sum of square residuals for each model individually before choosing the minimum among all competing models.

[14] While our focus here is on Milankovitch cycles, this procedure is applicable to a broad spectrum of fitting functions. For least squares fit to a general fitting function $f(\cdot|\Theta)$ with parameters Θ step 1 becomes:

for i = 1,. . .,N

for j = i +d,. . .,N

for m = 1,. . .,M

$$SS_{i,j}^{m} = \underset{\Theta}{\mathbf{Min}} \left\{ \sum_{t=t_i}^{t_j} \left(Y_t - f\left(X_{i,t}, \ldots X_{p,t}|\Theta\right)^2 \right) \right\}$$

end for

$$SS_{i,j} = \mathrm{Min}\left\{ SS_{i,j}^{1}, \ldots SS_{i,j}^{M} \right\}$$

end for

end for

where $X_{i,t}$ are a covariates, or "predictive variables". As indicated next, the remaining two steps of this algorithm can be completed for any function for which the necessary minimizations can be completed to obtain $SS_{i,j}$, $i = 1,\ldots,N$ and $j > i$. Thus, $f(\cdot|\Theta)$may be nonlinear. When a linear function is employed, the well known explicit solutions of linear regression equations can be employed to obtain the required minima.

[15] The step just described fits the parameters and finds the minimum sum of square residuals for every possible subinterval. To find the optimal k change points, we find the overall minimum sum of square residuals by optimally piecing together (k + 1) subintervals that include each of the N data points only once. There are a large number, $O(MN^K)$, of ways to piece together subintervals (for each value of k), so except for a short time series, this optimization cannot be completed by enumerating all of the possibilities [A total of $\sum_{k=1}^{K\max} MN^k$ solutions exist]. Instead, dynamic programming, a recursive algorithm, addresses this combinatoric explosion by breaking the problem into a progressive set of smaller problems. In change point analysis, this progressive set includes all substrings of the sequence that begin with the first data point and end at all time points from time 2 to time N. Optimal values, here the minimum squared errors, are first obtained for the simplest of these, the one change point problem. The stored values from this first pass are then used to find the minimum squared errors for all solutions with two change points and these in turn for three change point models, and so on, until the maximum number of change points, K_{max}, has been included. Upon completion of the forward step, minimum squared errors for all subproblems have been obtained and stored. Let $f_k(v)$ be the minimum sum of square deviations for the subinterval from 1 to v with k change points. Note that $f_1(v) = SS_{1,v}$.

[16] Step 2 is the forward step:

for k = 2,. . .K

for n = k,. . .,N

$$f_k(t_n) = \min_{v<t_n} \left\{ f_{k-1}(v) + SS_{v+1,t_n} \right\}$$

end for

end for

The forward step finds the minimum sum of square residuals with up to K_{max} change points, for the entire time series and all possible subintervals. To find the locations of the change points, we work backward. On the backtrace step, the results from the forward step are employed to find the optimal solution. The location of the Kth change point is found by using the minimum squared error for the set of subproblems with K – 1 change points. The remaining change points are found recursively by stepping backward in a similar manner. This means that we begin at the end of the time series, find the location of the last change point, and then recursively repeat this process with smaller values of k and progressively shorter sequences using the stored values from the forward procedure at each stage of the

backward procedure. Let c_k be the location of the kth change point.

[17] Step 3 is the backtrace step:

$$c_{K+1} = t_N$$

for k = K, K−1, ...1

$$c_k = \underset{v \in [k-1,\, c_{k+1})}{\arg\min} \left\{ f_{k-1}(v) + SS_{v+1,c_{k+1}} \right\}$$

end for

These steps reduce the time requirements for this computation to $O(MN^2)$ for the parameter fitting step and O(KN) for the forward and backward steps.

4. Applications and Results

[18] The change point algorithm was applied to two different types of data sets. One data set is the orbitally tuned 5-Ma-long isotope stack of *Lisiecki and Raymo* [2005] (hereinafter referred to as LR05). Because orbital tuning may induce a bias in the locations of the change points, a second analysis was also done on a 2.5-Ma-long data set that is free from orbital tuning [*Huybers*, 2007] (hereinafter referred to as H07), an extension of *Huybers and Wunsch* [2005]. The H07 approach pays the price of its freedom from orbital tuning by adding considerable age uncertainties in regions distant from absolute age control points. To address this uncertainty, several realizations of plausible age models would have to be simulated and independently tested by the change point algorithm to ensure that results are not due to stochastic fluctuations. This type of analysis is beyond the scope of this paper and we chose not to weigh in on the merits of one type of data set over the other.

[19] The two data sets are quite similar. As *Huybers* [2007] points out, the differences between the ages of the two models has a standard deviation of only 6 ka, which is less than the expected uncertainty for the depth-derived ages. There are however, two important differences. Prior to 600 ka, the sampling densities of the two data sets are different. H07 has data points corresponding to every 1 ka, while LR05 has data points spaced every 2 ka from 600 ka to 1500 ka, and then every 2.5 ka from 1500 ka to the end of the H07 data set. Second, a circa 100 ka pulse is seen in the H07 data at 2 Ma that is not found in LR05. Accordingly, the change point algorithm will bound this interval with a pair of change points when the 100 ka sinusoid is utilized in a model. A comparison of the results from the algorithm on the two published data sets is given below in section 4.2. This type of comparison provides insight into the most important changes of the climate system that are visible across different models and across different data sets.

[20] The first step for either data set is to remove the overall trend from the data via an exponential function so that the change points represent changes in the Earth's response to orbital inputs rather than the cooling and long-term Plio-Pleistocene trend toward increased ice volume (see Figures 1a and 1b).

[21] All analyses described here include change points associated not only with changes in the dominant frequencies, but also changes in the amplitude and phase parameters for a given frequency. The phase parameter takes values between −180° and +180° and gives the phase shift of the sinusoidal functions relative to time 0 (present day). This choice was arbitrary and has no bearing on the final results. The results are identical if we choose t_1 to be the most recent or the most distant time point. The amplitudes are in δ^{18}O units. Here and in the remainder of this study, we constrained the minimum interval length to be at least as large as the longest wavelength included in the regression model. Because of the nature of the algorithm, all change points will represent an abrupt or discrete change in the system when considered in isolation. However, gradual changes can be realized when the change points and resulting model parameters are viewed in context. Several nearby change points that show slowly varying or stepwise changes in parameter values are indicative of a gradual change in the system.

[22] There is a marked decrease in the variability prior to about 2500 ka (Figure 1b). Accordingly, change points before 2500 ka are less likely since there is less squared variation to account for. Therefore, in addition to an analysis directly on the detrended data shown in Figure 1b, we conducted a "weighted" least squares analysis. After completing step 1 on each subinterval, we weighted subintervals by 1/variance around the mean of that interval prior to application of steps 2 and 3 (Figure 1c). In general, we found that this weighting made little difference except between ∼2400 and ∼2700 ka, where an additional change point was located in this region in the weighted analysis compared to the unweighted analysis, and within the "transition interval" ∼800 to ∼1200 ka discussed below, where we saw a shift in the location of the change points. Here, we present the weighted analysis while referring the reader to the unweighted analysis located in the auxiliary material.[1]

4.1. Dominant Frequency Analysis

[23] We begin our analysis by examining the LR05 time series and addressing the traditional problem of identification of the single dominant orbital frequency in each subinterval, selecting in each subinterval only one of the three sinusoidal functions: "eccentricity" (100 ka), obliquity (41 ka), or precession (23 ka). This is an example of the above pseudocode with M = 3. Strong support for the existence of these frequencies within at least some region of the data already exists through conventional Fourier analysis. Other sinusoidal functions could also have been chosen. The choice of 100 ka to represent eccentricity was based on the empirical power of a circa 100 ka cycle during the late Pleistocene. This model is a first attempt to explain the data and serves primarily as a comparison to spectral analysis techniques that have been used in paleoclimatology.

[1] Auxiliary materials are available in the HTML. doi:10.1029/2007PA001568.

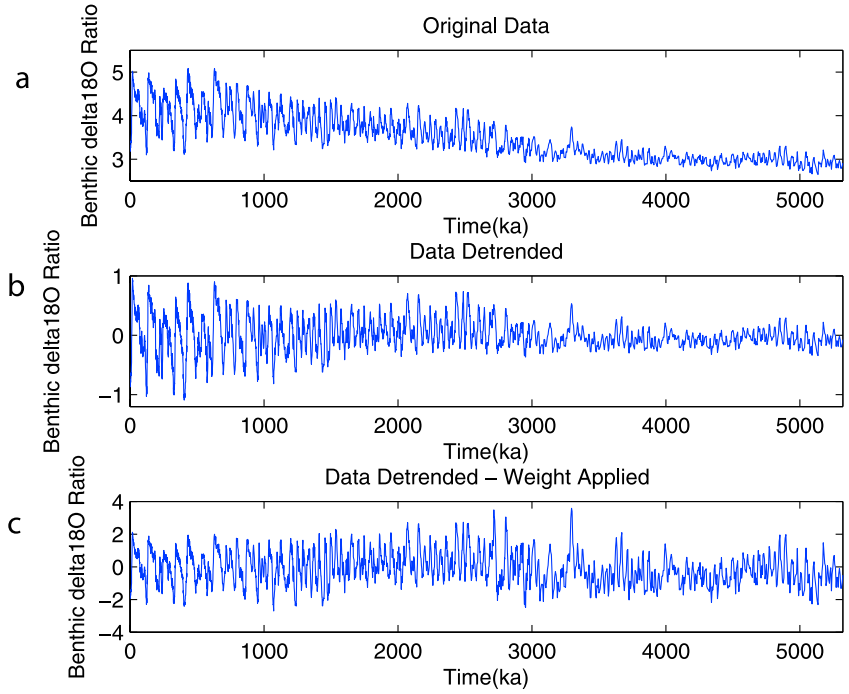

Figure 1. Benthic δ^{18}O data. (a) The original δ^{18}O data as obtained from *Lisiecki and Raymo* [2005]. (b) The δ^{18}O data after the overall trend was removed using an exponential function. (c) The detrended data after a weight was applied to each subinterval. The subintervals were selected by the change point algorithm when the input was the linear combination of the Milankovitch frequencies (actual change points given in Table 2). The weight used was 1/(standard deviation) of that subinterval.

[24] There is no statistically rigorous method to determine the number of change points required in a least squares analysis. As an alternative, we examined changes in squared error as a function of the number of change points in an effort to identify the "break in the curve", beyond which adding more change points yields little improvement in fitting the data. For the analysis of dominant frequencies, we used four change points (Figure 2a).

[25] Our dominant frequency analysis indicates that the most recent 1 Ma exhibit a more complex pattern requiring more change points than the previous 4 Ma, even in the weighted analysis. Thus, as shown in Table 1, when four change points are selected, three are between 778 ka and the present (113 ka, 424 ka and 778 ka) and the regimes bounded by these change points are dominated by the 100 ka cycle with changes in the amplitude and phase parameters. Figure 3 shows the observed data and the fitted functions. We find a change point near the minimum permitted length for the most recent interval, here at 113 ka. When we relaxed the minimum length constraint, we consistently found that the most recent 80 ka includes a change point that seeks to capture an unusual pattern of the recent past. The change point at ~780 ka persists throughout the subsequent analyses (Tables 2–4). The two intervals prior to ~780 ka (780–2713 ka and 2715–5320 ka) are both best fit by the obliquity forcing function with a threefold change in the amplitude of the obliquity response at about 2713 ka (Table 1). Only the influences of the 41 ka (obliquity) and 100 ka (eccentricity?) forcing functions are inferred over the entire time series. The unweighted analysis with four change points yields similar results to the weighted analysis, with change points at 113, 425, 786, and 2703 ka, and an R^2 value that is less than half a percent greater than the weighted least squares, 0.465. Complete details of the unweighted analysis are presented in the auxiliary material.

[26] Over the full 5320 ka interval, less than half [0.462] of the squared variation (R^2) in the ratio of oxygen isotope values is accounted for by the best fit of a single dominant "Milankovitch" frequency in 5 subintervals (Table 1). Furthermore, in the interval from 2715 ka to 5320 ka only 18% of the squared variation is accounted for by the best fitting of these functions (Table 1). Table 1 illustrates the ability of the algorithm to identify dominant frequencies without the required restrictions on window size. Increasing the number of change points provides a better fit, but doubling the number of change points (to 8) only increases the total R^2 value to 0.521. Reducing the number of change points to 2, corresponding to the two well-recognized changes in the behavior of Plio-Pleistocene glaciation (the mid-Pleistocene transition (~900 ka) and the intensification of Northern Hemisphere glaciation ~2.7 Ma) returns changes at 692 ka and 2713 ka, but an R^2 value of only 0.357.

[27] In the two change point analysis, the mid-Pleistocene change occurs at 692 ka in comparison to its location at 778 ka in the four change point analysis. The movement of the location of change points when additional change points are added to the analysis is not unexpected. For example, suppose that the data contained more "changes" than the

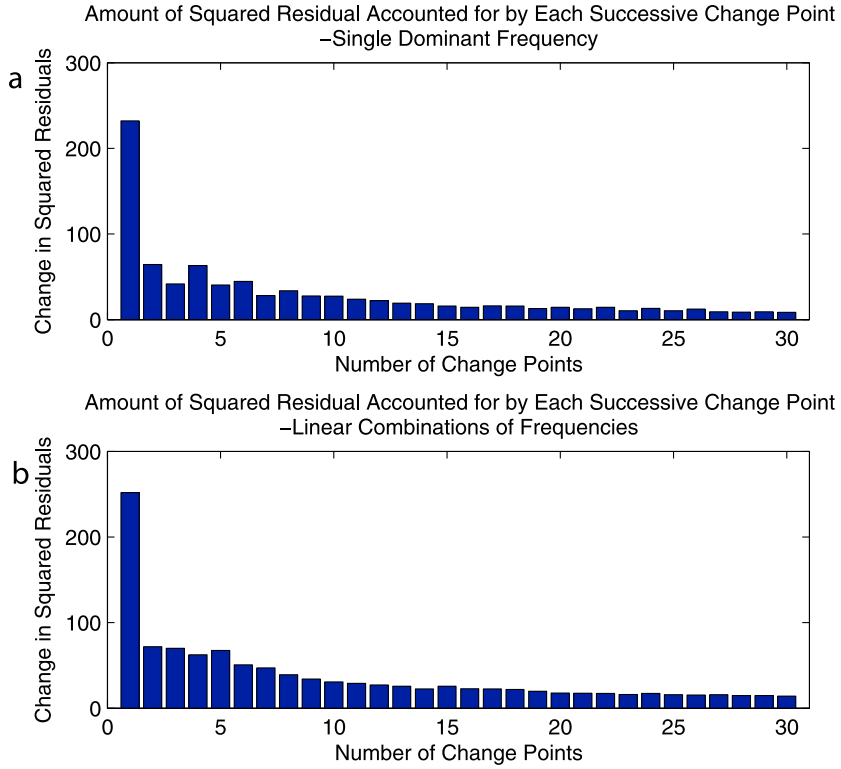

Figure 2. Amount of squared residual accounted for by each successive change point. The slope of the squared error curve as a function of the number of change points: (a) the change in the residual sum of squares with the addition of each successive change point for the weighted analysis of the single dominant frequency and (b) the change in the residual sum of squares with the addition of each successive change point for the weighted analysis of linear combinations of the Milankovitch frequencies. The ''break in the curve'' corresponds to a change in the slope of the squared error curve.

number of change points the algorithm was asked to find. If this were the case, a single subinterval may have to account for several inhomogeneous regions within its bounds and would thus average the parameter values across these inhomogeneous regions. With an appropriate number of change points, each subinterval truly has homogeneous parameter values within its bounds.

[28] We were also concerned that our results might be sensitive to the precise frequency chosen for the circa 100 ka glacial cycle. To address this concern, the single dominant frequency analysis was also separately run using 95 ka and 104 ka cycles instead of the 100 ka cycle. When the 95 ka cycle was used instead of the 100 ka cycle, the change points were located at time positions 112, 788, 916, and 2713 ka, with a total R^2 value of 0.449, a 1.5% reduction. As was true for the analysis using a 100 ka sinusoid, the circa 100 ka cycle, in this case the 95 ka term, dominates (i.e., best fits) the three most recent intervals, while the 41 ka cycle best fits the other two intervals. When using the 104 ka cycle instead of the 100 ka cycle, the change points occur at time positions 423, 792, 1216, and 2715, with a total R^2 value of 0.422, a 4.3% reduction. Again, the circa 100 ka cycle dominates the most recent three intervals (0–1216 ka), while the 41 ka cycle dominates the other two (1218–

Table 1. Single Dominant Frequency[a]

Subintervals	Dominant Period (ka)	Amplitude (δ^{18}O)	Phase	Constant (α)	R^2
1 to 113	100	0.4584	−150.78	0.1435	0.5368
114 to 424	100	0.5049	141.47	−0.0804	0.566
425 to 778	100	0.412	−160.20	0.021	0.5218
780 to 2713	41	0.2188	158.78	0.037	0.2895
2715 to 5320	41	0.0896	−156.12	−0.0468	0.1828

[a]The single dominant frequency (LR05) in each of five subintervals using only the three Milankovitch frequencies: 23, 41, and 100 ka, with the constraint that the minimal subinterval length be at least as long as the longest wave in the analysis (100 ka). The amplitude is given in δ^{18}O units, the phase is given in degrees from −180 to 180, and the R^2 values are the percent of the squared variation that a single sinusoid is able to account for on its own. Total R^2 = 0.4615.

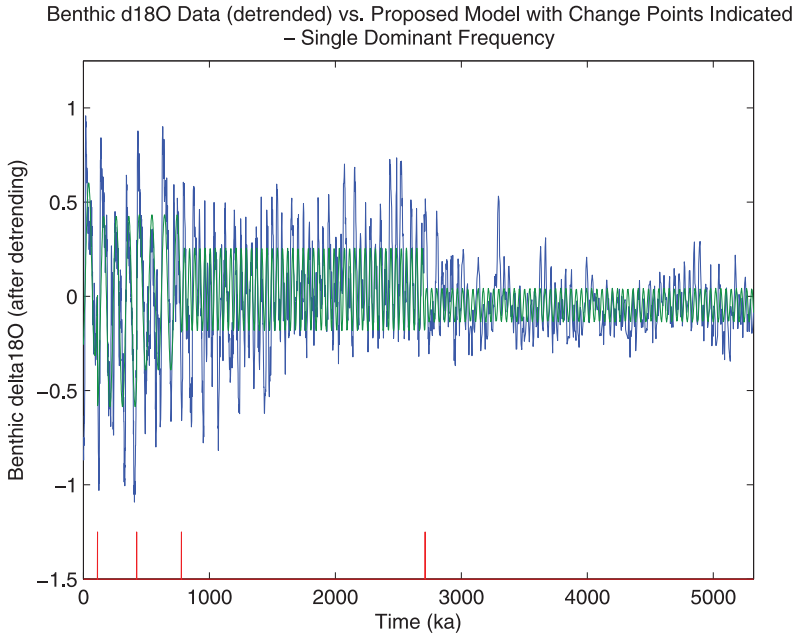

Figure 3. Benthic δ^{18}O data plotted versus proposed model with change points indicated: single dominant frequency (LR05). The least squares fit for the single dominant Milankovitch based cycle in each subinterval. The data are plotted in blue, the proposed model is plotted in green, and the change points are plotted as red spikes at the bottom. The data shown are the original data after the trend was removed. The change points were selected using the weighted least squares approach but are plotted against the unweighted data.

5320 ka). Importantly, throughout this range of candidate eccentricity frequencies, the change points at ~790 ka and ~2715 ka remain unchanged. Our modeling of circa 100 ka behavior in glaciation over the past 5 Ma is therefore quite reliable, whatever our limitations on its interpretation may be.

4.2. Linear Combinations Analysis

[29] Next we consider models with linear combinations of sinusoidal functions of the three classic Milankovitch frequencies (100, 41, and 23 ka). This is a special case of the above pseudocode with M = 1. In this section, we analyze both the 5 Ma orbitally tuned benthic δ^{18}O LR05 data set as well as the 2.5 Ma nonorbitally tuned benthic δ^{18}O H07 data set. The goal of this section is not only to compare the results of the change point algorithm on the two data sets, but to be able to draw inferences that are independent of the age models used. We begin our discussion with an analysis of the LR05 data set.

[30] In this analysis, we increased the number of change points to seven to show a more nuanced result (Figure 4 and Table 2). Change points occur at 102 ka, 380 ka, 786 ka, 1030 ka, 1198 ka, 2418 ka, and 2713 ka. Five of the seven change points occur between the early onset of the mid-Pleistocene transition identified in the literature ~1500 ka [*Rutherford and D'Hondt*, 2000] and the present, while the remaining two change points occur near the intensification of Northern Hemisphere glaciation ~2.7 Ma. As shown in the auxiliary material, adding up to 10 change points does

Table 2. Linear Combinations of Milankovitch Frequencies[a]

	23 ka		41 ka		100 ka		
Subintervals	Amplitude	Phase	Amplitude	Phase	Amplitude	Phase	Constant (α)
1 to 102	0.2043	71.68	0.2986	146.18	0.376	−147.34	0.1797
103 to 380	0.1962	39.51	0.2916	−161.88	0.4354	133.52	−0.0209
381 to 786	0.0344	−158.14	0.2569	−164.03	0.4675	−169.64	−0.0303
788 to 1030	0.1431	28.07	0.2002	178.88	0.2761	13.61	−0.0492
1032 to 1198	0.0925	14.63	0.1281	139.93	0.3148	−82.71	−0.0154
1200 to 2418	0.0292	−129.26	0.225	152.78	0.0553	102.22	0.0402
2420 to 2713	0.0273	29.12	0.2588	162.55	0.184	−65.01	0.177
2715 to 5320	0.012	−172.05	0.0893	−156.22	0.0293	−15.29	−0.0467

[a]Linear combinations of Milankovitch frequencies (LR05): 23, 41, and 100 ka. The seven change points selected by the algorithm yield eight subintervals with corresponding amplitude and phase parameters for each of the three input sinusoids in each subinterval. The amplitudes are given in δ^{18}O units, and the phase is given in degrees from −180 to 180. Once again, a constraint was imposed that the minimum subinterval length be at least as long as the longest wave in the analysis (100 ka). Total R^2 = 0.6641.

Table 3. Percent of Squared Residual Accounted for by Each Subset of Inputs[a]

Subintervals	23	41	100	23, 41	41, 100	23, 100	23, 41, 100
1 to 102	0.1384	0.3589	0.5207*	0.5013	0.7222*	0.6171	0.8233
103 to 380	0.0862	0.198	0.4908*	0.291	0.6952*	0.5818	0.7923
381 to 786	0.0044	0.1546	0.5367*	0.16	0.7*	0.5389	0.7029
788 to 1030	0.0934	0.2313	0.3557*	0.31	0.5356*	0.4548	0.6206
1032 to 1198	0.0561	0.0764	0.5055*	0.1405	0.5876*	0.5451	0.6356
1200 to 2418	0.0063	0.3739*	0.0211	0.3804	0.3969*	0.0273	0.4032
2420 to 2713	0.0086	0.4348*	0.1905	0.4392	0.6636*	0.2006	0.6688
2715 to 5320	0.0036	0.1828*	0.0204	0.1861	0.2023*	0.024	0.2056

[a]Given the location of the seven change points from the linear combination of Milankovitch frequencies (Table 2), we look at the amount of squared variation (R^2) that each subset of sinusoids is able to account for (LR05). For a given number of sinusoids included in the model, the subset that accounts for the most squared variation is indicated with an asterisk.

not relocate six of the seven change points and moves the seventh change point by only 2 ka.

[31] To better understand the contributions of each of the Milankovitch terms to each subinterval (i.e., climate regime), we examine the contributions of all subsets of Milankovitch terms to the proportion of squared variation (R^2) accounted for in each of the predicted subintervals, with the seven change point locations fixed to those in Table 2 (Table 3). An asterisk indicates the subset of Milankovitch sinusoids that accounts for the most variance in each subinterval. The contribution of the precession-based term is small compared to the obliquity and eccentricity terms (Table 3).

[32] The amplitude of the obliquity-driven glacial cycle has a roughly constant magnitude from 2715 ka to the present, even though it is not the dominant frequency for the past 1.2 Ma (Table 2). Prior to 1.2 Ma, the obliquity term dominates while the other two terms add little to the overall fit. This result is robust against the selection of the number of change points, as opting for fewer (5) or more (10) change points yields similar results. While the obliquity response dominates in the interval 2420 ka to 2713 ka, the amplitude of the eccentricity based function increases relative to adjoining intervals. It is perhaps not coincidental that this ~300 ka interval of mixed spectral behavior follows immediately after the significant increase in the amplitude of glacial cycles at about 2713 ka. In all cases, the amplitude of the obliquity based function was nearly threefold lower before 2715 ka ago, consistent with the absence of large Northern Hemisphere glaciers and their associated feedbacks during this interval [cf. *Shackleton et al.*, 1984]. In this interval, we see that all three sinusoids together are only able to account for just over 20% of the squared variation of the δ^{18}O data. We concede that our statistics concerning the dominance of obliquity forcing may be biased by the obliquity-based tuning procedure used by *Lisiecki and Raymo* [2005] to date the δ^{18}O record. However, despite the fact that *Lisiecki and Raymo* [2005] also included precession in their tuning procedure, the 23 ka cycle never plays an important role in the change points we identified, and never explains a large fraction of the variance in our models.

[33] In the last 1200 ka, the eccentricity-based term is the "dominant" cycle because it is able to account for the most squared variation in each of these subintervals. The 100 ka based cycle shows a marked increase in amplitude beginning 1198 ka ago (Table 2). The amplitude further increases at 786 ka. Thus, the time between roughly 1200 ka and 800 ka could be considered a "transition interval". When we increased the number of change points in the single dominant frequency analysis from four to seven, the 100 ka cycle dominates from 1200 ka to the present. Thus, the extension of the interval in which eccentricity is dominant, stems from the inclusion of three additional change points, again suggesting that the interval between ~780 ka and ~1200 ka is a transition interval.

[34] In the most recent 5 subintervals (0–1198 ka), the addition of the obliquity term to the 100 ka term improves

Table 4. Parameters for Linear Combinations of Milankovitch Frequencies[a]

	23 ka		41 ka		100 ka		
Subintervals	Amplitude	Phase	Amplitude	Phase	Amplitude	Phase	Constant (α)
1 to 791	0.0722	7.71	0.1467	50.38	0.434	13.50	−0.0106
792 to 1195	0.0584	72.26	0.1782	19.57	0.1793	2.96	0.0339
1196 to 1700	0.0291	−26.85	0.2521	39.84	0.0431	−52.44	−0.0264
1701 to 1990	0.0273	34.89	0.2003	−36.07	0.0445	37.32	0.0544
1991 to 2103	0.0805	26.96	0.0774	89.58	0.2362	51.11	0.0719
2104 to 2331	0.0146	45.11	0.176	29.45	0.0206	−77.50	−0.0877
2332 to 2580	0.0367	−59.36	0.2101	−6.77	0.1574	76.38	0.0189

[a]Linear combinations of Milankovitch frequencies (H07): 23, 41, and 100 ka. The six change points selected by the algorithm yield seven subintervals with corresponding amplitude and phase parameters for each of the three input sinusoids in each subinterval. The amplitudes are given in δ^{18}O units, and the phase is in degrees from −180 to 180. A constraint was imposed that the minimum subinterval length be at least as long as the longest wave in the analysis (100 ka). Total R^2 = 0.5680.

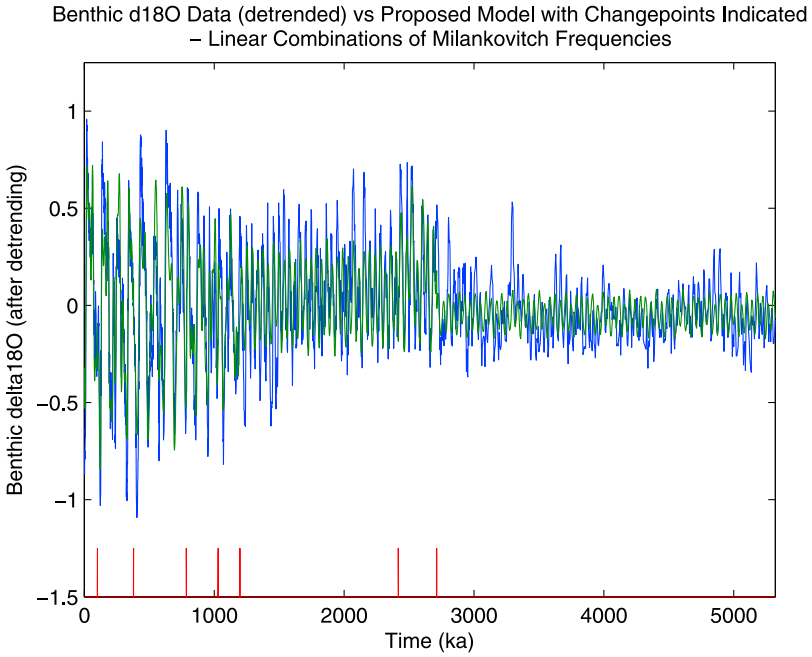

Figure 4. Benthic δ^{18}O data plotted versus proposed model with change points indicated: linear combinations of Milankovitch frequencies (LR05). The least squares fit for the linear combination of the three Milankovitch based cycles and a constant term with seven change points. The actual data are in blue, the proposed model is in green, and the red spikes at the bottom represent the optimal locations of change points selected by the algorithm.

the fit by over 16% in all subintervals with the exception of 1032–1198 ka where the fit is improved by 8%. Also, we note that the contribution to squared error is nearly additive (i.e., adding together the R^2 values for the 41 ka and 100 ka individually, gives the R^2 for the combination 41 ka and 100 ka subset), suggesting that the sinusoidal waves are nearly orthogonal [*Bretthorst*, 1988]. When the frequencies are different, all sinusoids of infinite lengths will be orthogonal; in this case, we are dealing with sometimes short segments so that orthogonality cannot be assumed.

[35] For 6 of the 8 intervals, the constant term is close to zero as expected after detrending, but not so for the other two intervals. For the most recent 102 ka, the constant is 0.179. This is consistent with higher δ^{18}O levels associated with the well reported greater glacial maximum in this interval. Also, in the interval 2420 ka to 2713 ka, the constant has a value of 0.177, perhaps indicating greater average ice volume in this interval relative to adjoining intervals.

[36] In total, 66% of the total unweighted squared variation in these data can be fit by the three sinusoidal functions in the indicated eight subintervals. As there is clear evidence that the climate behavior at the orbital periods is not linear [*Ashkenazy and Tziperman*, 2004; *King*, 1996], our simple model undoubtedly underestimates the total fraction of climate variance paced by orbital forcing through the Pliocene and Pleistocene. At the same time, some of the variation explained by the model may be byproduct of the orbital tuning of the data set.

[37] The nonorbitally tuned H07 data set provides us with a different set of change points than does the LR05 data set (Table 4). To aid the comparison, we sought the six optimal change points as the seventh change point in the LR05 data set falls outside of the timeframe of the H07 data set. In total, the change point algorithm with six change points for the linear combination of Milankovitch frequencies was able to account for 56.8% of the squared variation in the data, nearly 10% less than in the LR05 data set. This difference may be attributable to orbital tuning and the inclusion of a 41 ka sinusoid in the model being tested. Two of the six change points are at nearly identical spots, ~788 ka and ~1200 ka. The fact that these two change points appear in both data sets strongly indicates that major changes occurred at these times. The biggest difference between the change points locations in the two data sets was in their clustering. In the H07 data set, the majority of the six change points are located prior to 1200 ka (Figure 5); in the LR05 data set, a majority of the change points are located after 1200 ka. We caution that the two data sets have different sampling densities at 2 Ma and that we cannot rule this out as one possible cause of this discrepancy.

[38] As with the LR05 data set, in H07, we see a stepwise increase in the amplitude of the 100 ka cycle at 1200 ka and again at ~790 ka. In the H07 data set, after 1200 ka the 41 ka cycle remains an important component to the system, albeit subordinate to the 100 ka cycle. Because the data set is not orbitally tuned, we no longer see the phase locking of the 41 ka cycle that was present in the LR05 data set. However, the change in phase remains small (less than 30°) except

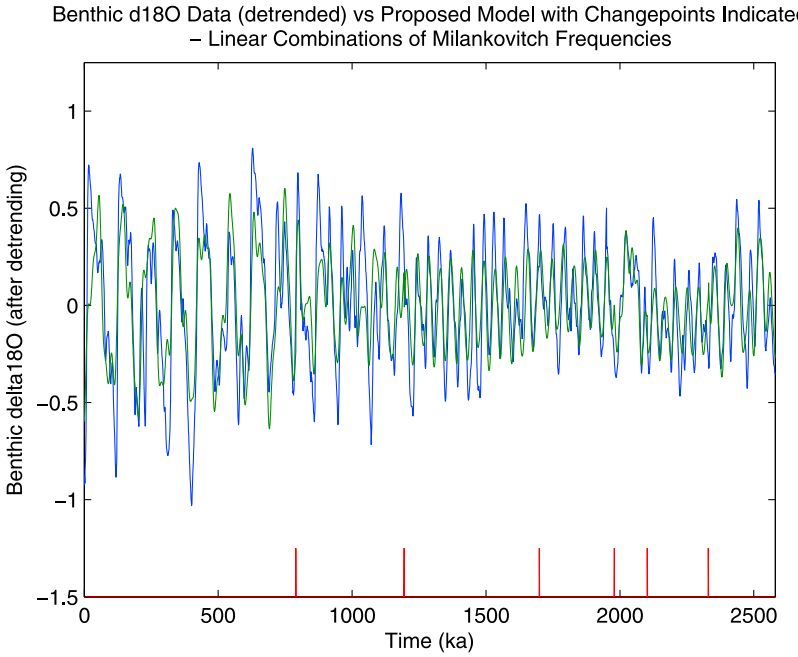

Figure 5. Benthic δ^{18}O data plotted versus proposed model with change points indicated: linear combinations of Milankovitch frequencies (H07). The least squares fit for the linear combination of the three Milankovitch-based cycles and a constant term with six change points. The actual data are in blue, the proposed model is in green, and the red spikes at the bottom represent the optimal locations of change points selected by the algorithm.

between 1700 and 2103 ka. The change point at 1700 ka deserves special attention. It appears that this change point is a result of a phase shift rather than an amplitude change of any of the three model components. It is plausible that this change point arises from age model uncertainty. The interval from 1991 to 2103 ka is an anomaly compared to the subintervals that surround it. In this interval, we see the relative weakness of the 41 ka signal and the emergence of a longer-period signal, a combination that is absent from the LR05 stack at this same time period. Accordingly, the change point algorithm picks out this interval as distinct from its neighbors. Finally, as with the LR05 data set, the precessional component is only a minor contributor throughout the time series.

[39] Looking more closely at the change points obtained from the H07 data set, we see that half of them fall near magnetic reversals, boundaries which were used to define the H07 age model: 780*, 990, 1070, 1770*, 1950*, and 2580 ka. The only change point in the LR05 time series that coincides with a magnetic reversal is the timing of the onset of the 100 ka regime at ~790 ka. It thus seems likely that in the case of the H07 data set, the change point method detects shifts in amplitude or phase of glacial cycles which are at least partly an artifact of how the time scale was derived.

4.3. The "100 ka" Cycle Analysis

[40] A recent study [*Huybers and Wunsch*, 2005] suggests that the 100 ka cycle conventionally attributed to orbital eccentricity (in some form) may stem instead from nonlinear climate system responses to the obliquity forcing. Nonlinear interactions of an obliquity response with strong feedbacks, such as the adjustment of large ice sheets and/or the extent of sea ice, might generate periodicities in glacial cycles near the observed 100 ka period. In particular, the transition to circa 100 ka glacial cycles might resemble the "Devil's staircase" transition in nonlinear systems [*Jin et al.*, 1994], where subharmonics (rational number multiples of the fundamental period) may dominate the overall response of the system [*Liu et al.*, 2008].

[41] To examine this question, we compare the ability of three different models to fit both the LR05 and H07 data sets: (1) "classic Milankovitch", 23, 41, and 100 ka; (2) "subharmonics of obliquity", 41, 82, and 123 ka; and (3) "alternate 100 ka", 41, 95, and 124 ka. The "classic Milankovitch" model is the same model that was just discussed. The 95 ka and 124 ka components of the "alternate 100 ka" model were chosen because they represent two of the major spectral components of the eccentricity function [*Berger and Loutre*, 1991]. The hypothesis that we wish to examine is whether or not the "classic Milankovitch" model fits the data as well or better than two alternative models for ice volume. In order to keep the model complexities the same, the number of change points was held fixed at seven and only three sinusoidal components were included in each model. As a result, the precession sinusoid was left out of both the "subharmonics of obliquity" and the "alternate 100 ka" models. Our aim is twofold: (1) to determine which model provides the best overall fit to the data and (2) to look for similarities in the change point locations across the three models in order to

discover changes in the time series that are independent of the functions used to model the data.

[42] We first compared the proportion of the variance (R^2) accounted for by each of the competing models on the LR05 data set. When seven change points were used, the "subharmonics of obliquity" model was able to account for 66.9% of the squared variation in the data, while the "classic Milankovitch" and "alternate 100 ka" models were able to account for 66.4% and 65.6% of the squared variation, respectively. In essence, all fit the data nearly equally well. From a practical standpoint, these models would appear to be nearly equivalent in their ability to represent the δ^{18}O data, even though two of the three models leave out precession, a minor but still important component. Inclusion of the precession term in the model increases the R^2 values for the "subharmonics of obliquity" and "alternate 100 ka" to 0.699 and 0.684, respectively. One must caution reading too much into the improvement in fit. These models now have two extra parameters (amplitude and phase of the 4th sinusoidal input) with which to fit the data.

[43] We also examined these three models and their fit to the H07 data. As before, in order to keep the model complexities the same, the number of change points was held fixed at six as the seventh change point in LR05 falls outside of the temporal range of the H07 data set. When six change points were used, the "subharmonics of obliquity" model was able to account for 64.2% of the squared variation in the data. The "classic Milankovitch" and "alternate 100 ka" models were able to account for 56.8% and 55.9% of the squared variation, respectively. In contrast to the LR05 data set, it appears that one of the models, "subharmonics of obliquity", substantially outperforms the other two models. Our analysis therefore supports the recent suggestion [*Huybers and Wunsch*, 2005; *Huybers*, 2007] that obliquity related terms describe the quasiperiodic components of glacial-interglacial change since the intensification of Northern Hemisphere glaciation at least as well as the traditional model.

[44] The change point locations for each of the three models on the LR05 data set are as follows: (1) subharmonics of obliquity, 263, 586, 1030, 1240, 1950, 2272, and 2713 ka; (2) classic Milankovitch, 102, 380, 786, 1030, 1198, 2418, and 2713 ka; and (3) alternate 100 ka, 118, 790, 1030, 1202, 1698, 2485, and 2878 ka. One of the most interesting details about the locations of the change points in the three models is the commonalities that exist between the three sets. There is one change point that appears exactly in all three models, 1030 ka, and others that are similar in at least two of the three models: ~788, ~1200, ~2450, and ~2713 ka. The importance of these times is highlighted by their inclusion in the optimal change point sets across multiple models. Interestingly, these common change point locations occur during the mid-Pleistocene transition, between 0.8 Ma and 1.2 Ma, and at the onset of the intensification of Northern Hemisphere glaciation around 2.7 Ma. Complete details of these two models with corresponding parameter values can be found in the auxiliary material.

[45] As previously mentioned, the change point locations for each of the three models on the H07 data set were somewhat different from their corresponding models on the LR05 data set: (1) subharmonics of obliquity, 262, 532, 1104, 1224, 2090, and 2303 ka; (2) classic Milankovitch, 791, 1195, 1700, 1990, 2103, and 233l ka; and (3) alternate 100 ka, 794, 1026, 1195, 1699, 1977, and 2337 ka. The change points for the "classic Milankovitch" and the "alternate 100 ka" models are very similar to each other on this data set (five of the six change points are within 13 ka of each other), but different from those of the "subharmonics of obliquity" model and different from their counterparts in the LR05 data set. On the other hand, the change points for the "subharmonics of obliquity" are similar in both data sets. There are two similarities across the three models on the H07 data set, namely the change points at ~1200 ka and ~2330 ka. Also of note is the change point found ~790 ka in two of the three models. This change point was also found both in the analysis on the LR05 data set and in the analysis on the single dominant frequency.

5. Evaluation

[46] Using the well-established "principle of optimality" [*Bellman*, 1957], we describe a general change point algorithm for the characterization of paleoclimatic time series and illustrate its application to the analysis of δ^{18}O data for the identification of the Earth system's response to Milankovitch forcing. The algorithm's novel characteristics are (1) globally optimal least squares fitting to a time series with windows of arbitrary size for a specified number of change points; (2) fitting of time series with combinations of predictor variables; and (3) applicability with any fitting function, linear or nonlinear. We presented an application of the change point method tailored to evaluating climate change on Ma time scales, but note that the method may be equally viable in the detection of significant changes in climate behavior on other time scales, including those induced by human activity.

[47] The strongest evidence for the utility of this algorithm in paleoclimatology comes from its confirmation of previously identified first-order spectral changes in Plio-Pleistocene δ^{18}O records, identification of the consistent importance of obliquity forcing throughout the Plio-Pleistocene, and demonstration that obliquity and its subharmonics as well as an alternate representation of the empirically derived 100 ka cycle fit the δ^{18}O record as well if not better than the traditional Milankovitch model [cf. *Hays et al.*, 1976]. This last result suggests that the use of the empirical 100 ka cycle may no longer be required. Using the LR05 data set as an example, the most important change points correspond to the intensification of Northern Hemisphere glaciation at about 2.7 Ma [*Shackleton et al.*, 1984] and to the mid-Pleistocene onset of large 100 ka glacial cycles (Tables 1–3). The 2.7 Ma change point comes not from a switch in the dominant periodicity (the 41 ka cycle dominates the orbital component of δ^{18}O change for more than 2 Ma before the change point and for nearly 2 Ma after), but from a significant increase in the amplitude of the 41 ka component in the δ^{18}O time series after 2.7 Ma and is associated with a large increase in the amount of variance explained by an orbital model for δ^{18}O change. Because the time series was weighted by the variance of each subinter-

val, the 2.7 Ma change point arises largely from the increased signal-to-noise ratio of the 41 ka component of benthic δ^{18}O after 2.7 Ma. Also, it appears that the 23 ka precessional component of ice volume strengthened significantly after 2.7 Ma, although it is always subordinate to the 41 ka obliquity cycle and accounts for only a very small portion of the variance throughout the time series (Table 3).

[48] Idealized as the transition from one dominant periodicity of glacial cycles to the next, the mid-Pleistocene transition occurred at about 0.788 Ma in both the LR05 and H07 analyses (Tables 1, 2, and 4), on the young side of most published analyses [*Park and Maasch*, 1993; *Rutherford and D'Hondt*, 2000]. However, a more complex representation of the δ^{18}O time series as sums of orbital terms (Table 3) displays interesting behavior during the transition from the "41 ka" world to the "100 ka" world. Two short (circa 200 ka) intervals between 0.788 and 1.2 Ma emerge from the LR05 data. These relatively short intervals may correspond to the apparently stepped transition from 41 ka to 100 ka dominated cycles detected by *Mudelsee and Schulz* [1997], in which the authors found that an increase in average ice volume preceded the spectral shift to 100 ka glacial cycles by more than 200 ka. In our analysis of the earlier (1.2 Ma) break point, the amplitude of the 41 ka cycle drops significantly at the same time that the strength of the 100 ka component of glaciation rises steeply. This dip in the obliquity component exists only for a single subinterval, from 1032 ka to 1198 ka, before returning to its previous level (Table 2). The younger bound for the mid-Pleistocene transition, 0.788 Ma, corresponds to the time after which the strength of the 100 ka cycle is nearly constant, a result echoed by the H07 data set. The mid-Pleistocene transition is also associated with a stepwise increase in the 23 ka precessional component (Tables 2 and 4). However, the 23 ka precessional sinusoid remains the least important of the three model components in terms of its amplitude (in both data sets) throughout the entire time period analyzed (Tables 2 and 4).

[49] Change point analysis of LR05 also suggests that climatic break points become more frequent toward the present. If the LR05 age model is robust, our analysis suggests that the climate system lurched more often as the intensity of glaciations increased in the late Pleistocene. The increasingly unstable pattern of glacial-interglacial cycles, evident by the large number of change points after the mid-Pleistocene transition in the LR05 data set, may indicate that the climate has drifted into a progressively more nonlinear state over time, and/or that new feedbacks have arisen in the late Pleistocene that significantly alter the frequency and amplitude behavior of the δ^{18}O time series. Paradoxically, the amount of variance explained by a simple sum of "Milankovitch" sinusoids increases dramatically toward the present (Table 3). This observation suggests that the larger amplitude glacial cycles of the late Pleistocene are better fit by finite length sinusoids than their early Pleistocene and Pliocene counterparts. The tendency of internal feedbacks to resonate at or near orbital periodicities may therefore have grown over time (see *Ashkenazy and Tziperman* [2004] for examples of how ice age cycle could become phase locked to orbital forcing).

[50] On the other hand, change point analysis of H07 suggests that the time around 2 Ma was more rugged. However, climate related inferences based on the H07 data set are limited because the model did not fit the data as well as it did the LR05 data set and because several of the change points may be artifacts of how the age model was determined (Tables 2 and 4). When comparing the results obtained from the orbitally tuned LR05 data set to the nonorbitally tuned H07 data set, two change points (788 ka and 1198 ka) appear in both data sets. Further comparisons between the two data sets are hindered by the differing sampling density around 2 Ma and the appearance of numerous changes in the H07 around magnetic reversals.

6. Caveats and Future Work

[51] The approach described here does have some important limitations. Thus, some caveats are in order. All of the models that were used in this analysis were assumed to be sinusoidal functions on the basis of the idea that climate change is the result of external forcing (i.e., solar insolation and Milankovitch Theory). However, the use of sinusoidal functions is not new as it is used for Fourier based methods as well. As discussed earlier, the change point algorithm can be adapted to more complex model functions. With respect to the near 100 ka cycle, it may be that this phenomenon originates from internal feedbacks in the climate system rather than external eccentricity forcing. The 100 ka term was included in the analysis because of its empirical significance in the δ^{18}O record.

[52] The change point algorithm in its current form is an example of interrupted regression [*Marsh and Cormier*, 2001] and thus does not have a continuity constraint. Therefore, the algorithm allows for both gradual and abrupt changes to take place. If continuity is desired, spline regression techniques can instead be used, but the problem then becomes nonlinear in its parameters and approximation techniques must instead be used [*Marsh and Cormier*, 2001]. Typically, the optimization is a very hard problem [*Lee*, 2002] and dynamic programming is not available to deal with this circumstance.

[53] One of the most important limitations of this algorithm is the lack of a rigorous method to infer the number of change points in a time series, leading to a need to examine the robustness of a study's conclusions against variations in this number. We find our major conclusions are insensitive to small variations in the number of change points (2–3 fewer/more). Drastic differences in this number would inevitably alter the results. Another important limitation of the present analysis is the absence of information concerning uncertainly in change point locations. With the current algorithm, it is not possible to distinguish between change points whose specific location is strongly supported by the available data, from those in which there may be considerable uncertainly about the exact timing of a change. It is possible that another set or several other sets of change points are almost as good as the optimal set returned by the algorithm. As a consequence, the ability of the algorithm to explore the rapidity of change is limited. Fortunately, a Bayesian approach, similar to that employed by *Liu and Lawrence* [1999], is available to

address this limitation, and to draw inferences from multivariate proxy data. By creating a probability distribution on the data and the parameters, we can draw sample solutions in proportion to their probability. Instead of returning the unique global optimum, the Bayesian approach returns an ensemble of solutions from which to draw inferences. The variability of these sampled solutions will give insight into the uncertainty in the number of change points, their locations, and the values of the parameters in the model. In the future, we intend to explore the rapidity of Plio-Pleistocene regime shifts using the Bayesian approach.

[54] Finally, it is clear that the results of the change point method will depend upon the choice of an initial model for ice age behavior; the present analysis is a point of departure, not a solution of the many enigmas of the ice ages. Our method does allow for a more formal and flexible study of transitions or breaks in time series than previously attempted, with the innovation of flexibility in window sizes and in subsequent modeling functions.

[55] **Acknowledgments.** The authors would like to thank Lorraine Lisiecki and Peter Huybers for the use of their data sets. We also would like to thank Warren Prell and Steve Clemens as well as the reviewers who provided valuable feedback and thoughtful comments on this manuscript.

References

Ashkenazy, Y., and E. Tziperman (2004), Are the 41 ka glacial oscillations a linear response to Milankovitch forcing?, *Quat. Sci. Rev.*, *23*, 1879–1890, doi:10.1016/j.quascirev.2004.04.008.

Auger, I. E., and C. E. Lawrence (1989), Algorithms for the optimal identification of segment neighborhoods, *Bull. Math. Biol.*, *51*, 39–54.

Bellman, R. (1957), *Dynamic Programming*, Princeton Univ. Press, Princeton, N.J.

Berger, A., and M. F. Loutre (1991), Insolation values for the climate of the last 10 million years, *Quat. Sci. Rev.*, *10*, 297–317, doi:10.1016/0277-3791(91)90033-Q.

Berger, W. H., T. Bickert, H. Schmidt, and G. Wefer (1993), Quaternary oxygen isotope record of pelagic foraminifers: Site 806, Ontong Java Plateau, *Proc. Ocean Drill. Program Sci. Results*, *130*, 381–395.

Birchfield, G. E., and M. Ghil (1993), Climate evolution in the Pliocene and Pleistocene from marine-sediment records and simulations: Internal variability versus orbital forcing, *J. Geophys. Res.*, *98*, 10,385–10,399, doi:10.1029/93JD00200.

Bloemendal, J., and J. deMenocal (1989), Evidence for a change in the periodicity of tropical climate cycles at 2.4 Ma from whole-core magnetic susceptibility measurements, *Nature*, *342*, 897–900, doi:10.1038/342897a0.

Bolton, E. W., K. A. Maasch, and J. M. Lilly (1995), A wavelet analysis of Plio-Pleistocene climate indicators: A new view of periodicity evolution, *Geophys. Res. Lett.*, *22*, 2753–2756, doi:10.1029/95GL02799.

Bretthorst, L. (1988), *Bayesian Spectrum Analysis and Parameter Estimation*, Springer, New York.

Briffa, K. R., P. D. Jones, F. H. Scheingruber, S. G. Shiyatov, and E. R. Cook (1995), Unusual twentieth-century summer warmth in a 1,000-year temperature record from Siberia, *Nature*, *376*, 156–159, doi:10.1038/376156a0.

Chappellaz, J., T. Blunier, D. Raynaud, J. M. Barnola, J. Schwander, and B. Stauffer (1993), Synchronous changes in atmospheric CH4 and Greenland climate between 40 and 8 ka bp, *Nature*, *366*, 443–445, doi:10.1038/366443a0.

Cobb, K. M., C. D. Charles, H. Cheng, and R. L. Edwards (2003), El Nino/Southern Oscillation and tropical Pacific climate during the last millennium, *Nature*, *424*, 271–276, doi:10.1038/nature01779.

Cole, J. E., R. G. Fairbanks, and G. T. Shen (1993), The spectrum of recent variability in the Southern Oscillation: Results from a Tarawa Atoll coral, *Science*, *260*, 1790–1793, doi:10.1126/science.260.5115.1790.

Dansgaard, W., et al. (1993), Evidence for general instability of past climate from a 250-kyr ice-core record, *Nature*, *364*, 218–220, doi:10.1038/364218a0.

Durbin, R., S. R. Eddy, A. Krogh, and G. Mitchinson (1998), *Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids*, Cambridge Univ Press, Cambridge, U.K.

Esper, J., E. R. Cook, and F. H. Schweingruber (2002), Low-frequency signals in long tree-ring chronologies for reconstructing past temperature variability, *Science*, *295*, 2250–2253, doi:10.1126/science.1066208.

Ghil, M., et al. (2002), Advanced spectral methods for climatic time series, *Rev. Geophys.*, *40*(1), 1003, doi:10.1029/2000RG000092.

Hawkins, D. M. (1976), Point estimation of the parameters of piecewise regression models, *Appl. Stat.*, *25*(1), 51–57, doi:10.2307/2346519.

Hays, J. D., J. Imbrie, and N. J. Shackleton (1976), Variations in the Earths orbit: Pacemaker of the ice ages, *Science*, *194*, 1121–1132, doi:10.1126/science.194.4270.1121.

Herbert, T. D., and L. Mayer (1991), Long climatic time series from DSDP/ODP physical property measurements, *J. Sediment. Petrol.*, *61*, 1089–1108.

Hillier, F. S., and G. J. Lieberman (2005), *Introduction to Operations Research*, 8th ed., McGraw-Hill, Boston, Mass.

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, doi:10.1016/j.quascirev.2006.07.013.

Huybers, P., and C. Wunch (2005), Obliquity pacing of glacial cycles, *Nature*, *434*, 491–494, doi:10.1038/nature03401.

Imbrie, J., J. D. Hays, D. G. Martinson, A. McIntyre, A. C. Mix, J. J. Morley, N. G. Pisias, and N. J. Shackleton (1984), The orbital theory of Pleistocene climate: Support from a revised chronology of the marine ^{18}O record, in *Milankovitch and Climate*, edited by A. Berger et al., pp. 269–305, D. Reidel, Hingham, Mass.

Imbrie, J., A. McIntyre, and A. Mix (1989), Oceanic response to orbital forcing in the late Quaternary: Observational and experimental strategies, in *Climate and Geo-Sciences*, edited by A. Berger et al., pp. 121–164, Kluwer, Dordrecht, Netherlands.

Imbrie, J., et al. (1992), On the structure of major glaciation cycles: 1. Linear responses to Milankovitch forcing, *Paleoceanography*, *7*, 701–738, doi:10.1029/92PA02253.

Jin, F.-F., J. D. Neelin, and M. Ghil (1994), El Niño on the devils staircase: Annual subharmonic steps to chaos, *Science*, *264*, 70–72, doi:10.1126/science.264.5155.70.

Jones, N., and P. Pevzner (2004), *An Introduction to Bioinformatics Algorithms*, MIT Press, Cambridge, Mass.

Jones, P. D., and M. E. Mann (2004), Climate over past millennia, *Rev. Geophys.*, *42*, RG2002, doi:10.1029/2003RG000143.

Joyce, J. E., L. R. C. Tjalsma, and J. M. Prutzman (1990), High resolution planktonic stable isotopic record and spectral analysis for the past 5.35 My: Ocean Drilling Program Site 625 Northeast Gulf of Mexico, *Paleoceanography*, *5*, 507–529, doi:10.1029/PA005i004p00507.

King, T. (1996), Quantifying nonlinearity and geometry in time series of climate, *Quat. Sci. Rev.*, *15*, 247–266, doi:10.1016/0277-3791(95)00060-7.

Konopka, A. K., and M. J. C. Crabbe (2004), *Compact Handbook of Computational Biology*, Marcel Dekker, New York.

Lau, K. M., and H. Weng (1995), Climate signal detection using wavelet transform: How to make a time series sing, *Bull. Am. Meteorol. Soc.*, *76*, 2391–2402, doi:10.1175/1520-0477(1995)076<2391:CSDUWT>2.0.CO;2.

Lee, T. C. M. (2002), On algorithms for ordinary least squares regression spline fitting: A comparative study, *J. Stat. Comput. Simul.*, *72*, 647–663, doi:10.1080/00949650213743.

Lew, A., and H. Mauch (2007), *Dynamic Programming: A Computational Tool*, Springer, New York.

Lisiecki, L. E., and M. E. Raymo (2005), A Pliocene-Pleistocene stack of 57 globally distributed benthic δ^{18}O records, *Paleoceanography*, *20*, PA1003, doi:10.1029/2004PA001071.

Liu, H. S., and B. F. Chau (1998), Wavelet spectral analysis of the Earths orbital variations and paleoclimatic cycles, *J. Atmos. Sci.*, *55*, 227–236, doi:10.1175/1520-0469(1998)055<0227:WSAOTE>2.0.CO;2.

Liu, J. S., and C. E. Lawrence (1999), Bayesian inference on Biopolymer models, *Bioinformatics*, *15*, 38–52, doi:10.1093/bioinformatics/15.1.38.

Liu, Z., L. C. Cleaveland, and T. D. Herbert (2008), Early onset and origin of 100-kyr cycles in Pleistocene tropical SST records, *Earth Planet. Sci. Lett.*, *265*, 703–715, doi:10.1016/j.epsl.2007.11.016.

Maasch, K. A. (1988), Statistical detection of the mid-Pleistocene transition, *Clim. Dyn.*, *2*, 133–143, doi:10.1007/BF01053471.

Mann, M. E. (2004), On smoothing potentially non-stationary climate time series, *Geophys. Res. Lett.*, *31*, L07214, doi:10.1029/2004GL019569.

Mann, M. E., R. S. Bandley, and M. K. Hughes (1998), Global-scale temperature patterns and climate forcing over the past six centuries, *Nature*, *392*, 779–783, doi:10.1038/33859.

Marsh, L. C., and D. R. Cornier (2001), Spline regression models, in *Sage Univ. Pap. Ser. on Quant. Appl. in the Soc. Sci., 07-137*, Sage, Thousand Oaks, Calif.

Mayewski, P. A., L. D. Meeker, M. C. Morrison, M. S. Twickler, S. Whitlow, K. K. Ferland, D. A. Meese, M. R. Legrand, and J. P. Steffensen (1993), Greenland ice core "signal" characteristics: An expanded view of climate change, *J. Geophys. Res.*, *98*, 12,839–12,847, doi:10.1029/93JD01085.

Mudelsee, M., and M. Schulz (1997), The Mid-Pleistocene climate transition: Onset of 100 ka cycle lags ice volume build-up by 280 ka, *Earth Planet. Sci. Lett.*, *151*, 117–123, doi:10.1016/S0012-821X(97)00114-3.

Paillard, D. (1998), The timing of Pleistocene glaciations from a simple multiple-state climate model, *Nature*, *391*, 378–381, doi:10.1038/34891.

Park, J., and T. D. Herbert (1987), Hunting for paleoclimatic periodicities in a sedimentary series with uncertain time scale, *J. Geophys. Res.*, *92*(B13), 14,027–14,040, doi:10.1029/JB092iB13p14027.

Park, J., and K. Maasch (1993), Plio-Pleistocene time evolution of the 100- ka cycle in marine paleoclimate records, *J. Geophys. Res.*, *98*, 447–461, doi:10.1029/92JB01815.

Petit, J. R., et al. (1999), Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica, *Nature*, *399*, 429–436, doi:10.1038/20859.

Quinn, T. M., T. M. Crowley, F. W. Taylor, C. Henin, P. Joannot, and Y. Join (1998), A multicentury stable isotope record from a New Caledonia coral: Interannual and decadal SST variability in the southwest Pacific since 1657, *Paleoceanography*, *13*, 412–426, doi:10.1029/98PA00401.

Ravelo, A. C., et al. (2004), Regional climate shifts caused by gradual global cooling in the Pliocene epoch, *Nature*, *429*(6989), 263–267, doi:10.1038/nature02567.

Raymo, M. E. (1994), The initiation of Northern Hemisphere glaciation, *Annu. Rev. Earth Planet. Sci.*, *22*, 353–383, doi:10.1146/annurev.ea.22.050194.002033.

Raymo, M. E., D. W. Oppo, and W. Curry (1997), The mid-Pleistocene climate transition: a deep-sea carbon perspective, *Paleoceanography*, *12*, 546–559.

Ridgwell, A. J., A. J. Watson, and M. E. Raymo (1999), Is the spectral signature of the 100 ka glacial cycle consistent with a Milankovitch origin?, *Paleoceanography*, *14*, 437–440, doi:10.1029/1999PA900018.

Ruddiman, W. F., M. Raymo, and A. McIntyre (1986), Matuyama 41000 year cycles: North Atlantic Ocean and Northern Hemisphere ice sheets, *Earth Planet. Sci. Lett.*, *80*, 117–129, doi:10.1016/0012-821X(86)90024-5.

Rutherford, S., and S. D'Hondt (2000), Early onset and tropical forcing of 100,000-year Pleistocene glacial cycles, *Nature*, *408*, 72–75, doi:10.1038/35040533.

Rutherford, S., M. E. Mann, T. L. Delworth, and R. J. Stouffer (2003), Climate field reconstruction under stationary and nonstationary forcing, *J. Clim.*, *16*(3), 462–479, doi:10.1175/1520-0442(2003)016<0462:CFRUSA>2.0.CO;2.

Shackleton, N. J., et al. (1984), Oxygen isotope calibration of the onset of ice-rafting and history of glaciation in the North Atlantic region, *Nature*, *307*, 620–623, doi:10.1038/307620a0.

Tome, A. R., and P. M. A. Miranda (2004), Piecewise linear fitting and trend changing points of climate parameters, *Geophys. Res. Lett.*, *31*, L02207, doi:10.1029/2003GL019100.

Torrence, C., and G. C. Compo (1998), A practical guide to wavelet analysis, *Bull. Am. Meteorol. Soc.*, *79*, 61–78, doi:10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2.

Tziperman, E., and H. Gildor (2003), On the mid-Pleistocene transition to 100-ka glacial cycles and the asymmetry between glaciation and deglaciation times, *Paleoceanography*, *18*(1), 1001, doi:10.1029/2001PA000627.

Vautard, R., and M. Ghil (1989), Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series, *Physica D*, *35*, 395–424, doi:10.1016/0167-2789(89)90077-8.

Vautard, R., P. Yiou, and M. Ghil (1992), Singular spectrum analysis: A toolkit for short noisy chaotic signals, *Physica D*, *58*, 95–126, doi:10.1016/0167-2789(92)90103-T.

Wagner, H. M. (1975), *Principles of Operations Research*, Prentice Hall, Englewood Cliffs, N.J.

Winston, W. L. (2003), *Operations Research: Applications and Algorithms*, 4th ed., Duxbury, Boston, Mass.

Wunsch, C. (1999), The interpretation of short climate records with comments on the North Atlantic and Southern Oscillations, *Bull. Am. Meteorol. Soc.*, *80*, 245–255, doi:10.1175/1520-0477(1999)080<0245:TIOSCR>2.0.CO;2.

Yiou, P., J. Genthon, J. Jouzel, M. Ghil, H. Le Treut, J. M. Barnola, C. Lorius, and Y. N. Korotkevich (1991), High-frequency paleovariability in climate and in CO_2 levels from Vostok ice-core records, *J. Geophys. Res.*, *96*, 20,365–20,738.

T. Herbert, Department of Geological Sciences, Brown University, Providence, RI 02912, USA. (timothy_herbert@brown.edu)

C. E. Lawrence and E. Ruggieri, Center for Computational Molecular Biology, Division of Applied Mathematics, Brown University, Providence, RI 02912, USA. (charles_lawrence@brown.edu)

K. T. Lawrence, Department of Geology and Environmental Geosciences, Lafayette College, Easton, PA 18042, USA.