Reconciling discrepancies between Uk37 and Mg/Ca reconstructions of Holocene marine temperature variability

Thomas Laepple[1] and Peter Huybers[2]

[1]Alfred Wegener Institute for Polar and Marine Research
Bremerhaven, Germany

[2]Department of Earth and Planetary Sciences, Harvard University
Cambridge, USA

Abstract

Significant discrepancies exist between the detrended variability of late-Holocene marine temperatures inferred from Mg/Ca and Uk37 proxies, with the former showing substantially more centennial-scale variation than the latter. Discrepancies exceed that attributable to differences in location and persist across various calibrations, indicating that they are intrinsic to the proxy measurement. We demonstrate that these discrepancies can be reconciled using a statistical model that accounts for the effects of bioturbation, sampling and measurement noise, and aliasing of seasonal variability. The smaller number of individual samples incorporated into Mg/Ca measurements relative to Uk37 measurements leads to greater aliasing and generally accounts for the differences in the magnitude and distribution of variability. An inverse application of the statistical model is also developed and applied in order to estimate the spectrum of marine temperature variability after correcting for proxy distortions. The correction method is tested on surrogate data and shown to reliably estimate the spectrum of temperature variance when using high-resolution records. Applying this inverse method to the actual Mg/Ca and Uk37 data results in estimates of the spectrum of temperature variance that are consistent. This approach provides a basis by which to accurately estimate the distribution of intrinsic marine temperature variability from marine proxy records.

Keywords: Holocene climate variability, signal and noise in proxies, multiproxy comparison, spectral analysis, SST variability

---

1. Introduction

Although Mg/Ca and Uk37 proxies are used to infer the same physical attribute of near-surface marine temperature, the recorded temperature signals reflect disparate life cycles and biophysical functioning of the proxy producing organisms and disparate incorporation and preservation of the signal in sediments. Potential contributions include non-temperature influences on the incorporation of Mg/Ca into foraminiferal shells (Arbuszewski et al., 2010), re-suspension and redeposition of Uk37 markers (Ohkouchi et al., 2002), and other possible post-depositional effects on Mg/Ca (Regenberg et al., 2006) and Uk37 (Hoefs et al., 1998; Gong and Hollander, 1999). These consideration make it important to infer temperatures from multiple sources and evaluate their consistency (e.g. De Vernal et al., 2006).

Previous studies noted that the temperature variability inferred for the last millenium differ according to the proxy type used (Richey et al., 2011). Proxy dependence has also been noted for the temporal patterns of deglacial warming (Steinke et al., 2008; Mix, 2006) and mid- to late-Holocene temperature trends (Leduc et al., 2010; Lohmann et al., 2012). Differences between Mg/Ca and Uk37 derived temperatures have been suggested to arise from differences in seasonal recording (Leduc et al., 2010; Schneider et al., 2010). How orbital variations manifest in proxy records sensitively depends on how the seasonal cycle is recorded (e.g., Huybers and Wunsch, 2003; Laepple et al., 2011), and this effect might explain diverging multi-millenial signals between proxies, or at least some fraction of the differences over the Holocene (Lohmann et al., 2012). Importantly, however, the ten-thousand-year timescale orbital variations are not expected to explain the differences in millennial and higher-frequency variability that are focused on in this study.

Here, we explore the discrepancies between Mg/Ca and Uk37 proxies of sea surface temperature at centennial to millennial timescales, identify a physical-statistical model for their origin, and present a method to correct for the associated biases when estimating temperature variance. Although it would be possible to interpret each individual proxy—or record or even single sample—as a unique perspective on past temperature, the emphasis here is to statistically account for distinctions between proxy measurements for the purposes of facilitating synthesis between records and comparison with instrumental observations and model simulations of temperature. In contrast to typical synthesis efforts that focus on reconstructing the time-history of temperature, we seek to estimate the magnitude of temperature variability as a function of timescale or, more precisely, the spectral distribution of sea surface temperature variability at centennial to millennial frequencies. Beyond holding intrinsic interest, quantitative estimates of temperature variability prior to the anthropogenic era and at frequencies lower than those afforded by instrumental records are generally needed when seeking to interpret specific changes in temperature and attribute them to a set of causes (e.g., Barnett et al., 1999).

2. Data and Methods

We focus our analysis on the two most prominent proxies of near-surface marine temperature, Mg/Ca ratios from planktic foraminifera (Lea et al., 1999) and the Uk37 ratio of different long-chain ketones (Brassell et al., 1986). Both proxies are recovered from sediment cores and are affected by bioturbation. An important distinction, however, is that each Mg/Ca measurement is typically made using a small number of crushed planktic foraminifera, usually about 30, whereas Uk37 is an organic proxy that is sampled from millions of molecules.

2.1. Proxy and instrumental data

The proxy dataset assembled for this study aims to be comprehensive in the sense of including all sufficiently long and well-resolved sediment records that cover the mid- to late-Holocene. Most Mg/Ca and Uk37 records are from the GHOST database (Leduc et al., 2010), though also included are two recently published high-resolution Mg/Ca records: MV99-GC41/PC14 (Marchitto et al., 2010) and MD99-2203 (Cleroux et al., 2012). Specifically, we include 6 planktonic Mg/Ca records from G.ruber and G.bulloides and 16 Uk37 records, all of which are dated by radiocarbon and have an average sampling rate of 100 years or less (Figure 1). Lower resolution records are excluded in the analysis because it is then difficult to accurately correct for sampling effects, as is later demonstrated.

All proxy records of a given type are recalibrated in a uniform manner to faciliate intercomparison. Uk'37 records are calibrated using 0.033 Uk37/°C, Uk37 records using 0.035 Uk37/°C, and Mg/Ca records using 9.35% Mg/Ca per °C.. These choices are the mean of all author calibrations of the analysed datasets but also agree with the standard calibrations given by Mueller et al. (1998) (0.033 Uk'37/°C) and Dekens et al. (2001) (9% Mg/Ca per °C).). See Table 1 and Figure 1 for more details regarding individual records. Instrumental observations that we later use to model the proxy recording process are from the HADSST3 compilation of sea surface temperatures (SST) (Kennedy et al., 2011b,a).

2.2. Spectral estimation

Spectral estimates are used to quantify timescale dependent variability. Although techniques exist to estimate spectra from unevenly sampled data (Lomb, 1976), our experimentation with synthetic signals indicates that more accurate results are obtained by first interpolating to a uniform sampling rate and then employing state-of-the-art spectral estimation techniques. Linear interpolation of an unevenly sampled record tends both to reduce the energy at the highest frequencies of a spectral estimate and to alias variability into lower frequencies (Rhines and Huybers, 2011).

To minimize the influence of high-frequency damping, we determine the finest interpolation resolution for which the frequency spectrum is largely unbiased. For an evenly sampled record, the optimal interpolation resolution would equal the sampling resolution, but for unequal time steps the optimal interpolation resolution is no longer obvious, and we employ a numerical method to determine an appropriate value. This process involves generating random numbers that follow a power-law processes with $\beta = 1$; subsampling these synthetic time series according to the sampling sequence of a given proxy record, and then interpolating to a resolution equal to the finest sampling time step of the original record. The spectral estimate of the resampled stochastic process is divided by the theoretical spectra, and the highest reliable frequency is determined by when this ratio crosses a value of 0.7. The selected interpolation resolution is then set to resolve the identified frequency, with the selected value rounded to the nearest 50 years and referred to as the optimal interpolation resolution. The optimal interpolation resolution depends on the evenness of the sampling. For example, core D13882 has a 53 year mean sampling resolution but contains some 140 year gaps, and the optimal interpolation resolution is 200 years, whereas other cores with a similar mean sampling rate have a 100 year optimal resolution.

To minimize issues associated with aliasing, data are first linearly interpolated to ten times the optimal resolution, lowpass filtered using a finite response filter with a cut-off frequency of 1.2 divided by the target time step, and then resampled at the optimal resolution.

Spectra are estimated using Thomson's multitaper method (Percival and Walden, 1993) with three windows. Time series are detrended prior to analysis, as is standard for spectral estimation. The multitaper approach introduces a small bias at the lowest frequencies and we omit the two lowest frequencies in all figures. For visual display purposes, power spectral estimates are also smoothed using a Gaussian kernel with constant width in logarithmic frequency space (Kirchner, 2005), when using logarithmic axes. When the smoothing kernel extends outside of the frequency range resolved by a record, it is truncated at both the low- and high-frequency ends of the kernel to maintain its symmetry and to avoid biasing estimates.

Our focus will be on the average power spectral estimate for each proxy type because this gives an improved signal-to-noise ratio and facilitates intercomparison between proxy types. This average power spectrum for each proxy type necessarily contains samples from regions with differing variability and that cover different frequency intervals. To avoid discontinuities across frequencies where the number of available estimates change, proxy spectra are scaled to an average value in the largest common frequency interval. Note that it is the spectral estimates that are averaged together, giving an estimate of the spectral energy, and that this is distinct from averaging together records in the time-domain, which would give an estimate of mean temperature.

Records are not intercompared in the time domain because timing errors generally destroy covariance and coherence. Spectral estimates, however, are largely insensitive to timing errors when the underlying process follows a power-law (Rhines and Huybers, 2011), which appears a good approximation for the proxy records considered here. Thus, intercomparison of spectral estimates derived from proxy records with uncertain timing is feasible. Power laws are estimated by a least-squares fit to logarithmic frequency and logarithmic power-density estimates (Huybers and Curry, 2006). To more uniformly weight the estimate, spectra are binned into equally spaced log-frequency intervals and averaged before fitting. Power-laws are only estimated in a frequency range common to all proxy records, 1/2000yr to 1/400yr.

3. Discrepancies between Mg/Ca and Uk37

On average, the Mg/Ca reconstructions of temperature have 2.2 times greater variance than Uk37 reconstructions, when making comparisons at 150 year resolution. This discrepancy in variance is visually apparent (Figure 1). There are several possibilities for the differences in variance. One is that the actual temperature variability at the Mg/Ca sites is greater than that at the Uk37 sites. Instrumental records from HadSST3 show the opposite, however, that Uk37 sites have between 1.2-1.5 times more variance than the Mg/Ca sites, depending on the season that is considered. Furthermore, a latitudinal comparison shows that Mg/Ca is the more variable within given regions (Figure 2), and similar results hold when sectors are defined according to longitude and latitude.

Another possibility has to do with differences or uncertainties in the calibration of proxies to temperature. Choosing the most sensitive calibration for Mg/Ca (0.107 (Mg/Ca)/°C (Mashiotta et al., 1999)) and the least sensitive calibration for Uk37 (0.023 Uk37'/°C,(Sonzogni et al., 1997)) gives an average variance that is similar for the two proxy types, but applying such calibrations globally is almost certainly inappropriate. The low sensitivity Uk37 calibration was developed for a region near the upper temperature limit of this proxy (24-29°C), and Sonzogni et al. (1997) actually suggest a global temperature calibration for Uk37 (0.031 Uk37'/°C) similar to the mean calibration used in this study. Moreover, rescaling the variability employing any single calibration would not resolve discrepancies in the relative distribution of fast and slow variability, which is discussed in more detail later.

A final consideration is whether the proxies record different seasons and, therefore, show systematically different amounts of variability. Such a mechanism would work for monthly to interannual time-scales where, for example, extratropical winter sea surface temperatures are generally more variable than summer or annual mean temperatures because of greater storm activity (Wallace et al., 1990). Analysis of a state-of-the-art coupled climate model, MPI-ESM (Jungclaus et al., 2010), shows this seasonal distinction in variability at short timescales, but that the distinction in variability is no longer identifiable at the positions of the Mg/Ca and Uk37 records at frequencies below 1/100 years. These model results indicate that seasonal differences in when temperatures are recorded to be an inadequate explanation of the discrepancy.

A more detailed picture of the discrepancy in variability between the two proxies can be obtained through spectral analysis. The spectra of each proxy record in the compilation is estimated using a multitaper procedure. Though consistent results are found evaluating single records, these comparisons are noisy, and we instead focus on the average spectra across each proxy type (Figure 2). Both the Mg/Ca and Uk37 spectral averages show increased energy toward lower frequencies, but the magnitude and detailed shape of these estimates are incommensurate in that Mg/Ca records have proportionately more high- than low-frequency variability than Uk37. The energy found at millennial (1-3kyr) relative to centennial variability (200yr-500yr) is 3.6 for Uk37 and 2.0 for Mg/Ca, a relative difference that is unaffected by calibration choice.

More generally, the spectra can be described using a powerlaw scaling, $f^{-\beta}$, where f is frequency in cycle per year and β is the power-law exponent.

Spectral energy increases less steeply with decreasing frequency in Mg/Ca, which has $\beta_{Mg/Ca} = 0.58$,than in Uk37, which has $\beta_{Uk37} = 0.98$.

The discrepancies between Uk37 and Mg/Ca variance and distributions of spectral energy generally exceed a factor of two, and it appears necessary to resolve these discrepancies prior to being able to infer temperature variability to within this factor. An accurate estimate of the temperature spectrum is important because it would directly indicate the range of natural temperature variation expected over a given timescale, indicate the physics that controls temperature variations, and serve as the basis for a test of whether climate models adequately represent climate variability (Hasselmann, 1976; Barnett et al., 1996; Pelletier, 1998; Huybers and Curry, 2006).

4. Proxy correction technique

We posit that the difference in variability between proxy types arises from processes that can be grouped into three categories: (1.) errors in temperature estimates arising from measurement noise, vital effects, and changes in depth habitats (Schiffelbein and Hills, 1984), (2.) irregular and/or infrequent sampling times that cause aliasing of seasonal and other high-frequency variability (Kirchner, 2005; Laepple et al., 2011), and (3.) bioturbation that mixes samples across time horizons (Berger and Heath, 1968). A number of other sources of uncertainty are also present but we assume and later confirm that they do not have first-order implications for the recorded variability. Building on existing models for these noise sources, we attempt to quantify their aggregate influence upon the spectra of each proxy using a single statistical model. We will also show that such a representation can be inverted to better estimate the frequency spectrum of temperature from proxy records.

4.1. Basics of the approach

Any given temperature record that we consider, y, has a spectral estimate, $S_{\hat{y}}$, that is corrupted by noise. Here we seek a best estimate of the true power spectrum, $S_{\hat{x}}$, using a correction that relies upon the biophysical characteristics of the proxy sampling process. For this correction we apply the spectral filtering approach of Kirchner et al. (2005) wherein a statistical model is constructed of the sampling process—including bioturbation, measurement, and other intrasample noise—and a filter is designed from the output of the model for the purposes of optimally estimating the true power spectrum.

Given a perfect model of the true spectrum, S_{xm}, and the sampled spectrum, S_{ym}, an optimal estimate of the true temperature spectrum, $S_{\hat{x}}$, can be obtained,

$$S_{\hat{x}} = S_{\hat{y}} \frac{S_{xm}}{S_{ym}}, \tag{1}$$

where the fractional term involving the model spectra is equivalent to a filter. We use a piecewise model of the true temperature spectrum that calls on observed instrumental temperature at high frequencies and a power-law at low frequencies, as has been found adequate for describing the spectral scaling of many other proxy records (Huybers and Curry, 2006),

$$S_{xm} = \begin{Bmatrix} cf^{-\beta}, & f \leq f_c \\ S_{\hat{x}i}, & f_c > f \end{Bmatrix}. \tag{2}$$

The frequency range extends from the lowest frequency sampled by the observed record to once per two years. Higher frequency variability is separately treated in the sampling process. $S_{\hat{x}i}$ is the spectral estimate from the observed SST and c is chosen to ensure that both pieces of the spectra meet at the cutoff frequency, f_c. The cutoff frequency is set at 1/50yr, the lowest frequency constrained by the instrumental record. The power-law, β, is an adjustable parameter constrained to take on values between zero and two. A value of β equal to zero corresponds to white noise, whereas values in excess of one, but not greater than two, have previously been found in proxy records that span glacial-interglacial variability (Huybers and Curry, 2006).

4.2. Bioturbation and sampling

Bioturbation is represented assuming a well-mixed sediment layer whose thickness is taken as an adjustable parameter (Berger and Heath, 1968). This gives an impulse response function, g, that fully describes the mixing response over the thickness of the bioturbation layer, δ (Guinasso and Schink, 1975). A $\delta = 10$cm bioturbational layer is typical of marine sediments (Boudreau, 1998; Guinasso and Schink, 1975) and is our default parameter, but we also examine the robustness of our results using 2 and 20cm layers. Because sediment cores have different mean accumulation rates, a, the timescale associated with bioturbational smoothing varies. No bioturbation is imposed for cores MV99-GC41/PC14 (Marchitto et al., 2010) and SO90-39KG/56KA (Doose-Rolinski et al., 2001) because they are laminated.

Uk37 samples comprise very large numbers of organic molecules, and we approximate such sample as continuous. The sampling can be described as a convolution of the temperature time series with the bioturbation impulse response function in the time domain, but for the purposes of describing the influence at a particular time horizon we cast the response as a sum across annual time steps,

$$y(t_i) = \sum_j x(t_{i+j})g(j) + \eta(t_i). \tag{3}$$

Although the sum is nominally over the entire depth of the core, in practice, we sum from $3\delta/a$ above to $1\delta/a$ below the time horizon of interest. This time interval of four times the bioturbational layer divided by the accumulation rate contains 99% of the weight in the impulse response, g. The noise component, $\eta(t_i)$, represents the measurement error as well as other intra-test variations, such as those caused by variations in depth habitat, and is assumed independent between samples and normally distributed.

Mg/Ca samples comprise a discrete sample of foraminifera tests, usually ranging between values of 20 to 60 for planktic samples. The sampling process is divided into interannual and subannual components for purposes of computational efficiency. The interannual component is selected as the annual average temperature, $x(t+\epsilon)$, where ϵ represents timing offsets introduced by bioturbation and is randomly selected according to the probability distribution defined by g. An additional noise term is then added to represent subannual variability, giving $x(t+\epsilon)+\psi(m)$. The value of $\psi(m)$ is selected as the monthly temperature anomaly from the climatological seasonal cycle, where the month is randomly chosen according to the modern lifecycle of the specific foraminiferal species at that core site simulated by a dynamic population model, PLAFOM (Fraile et al., 2008),

$$y(t_i) = \frac{1}{N}\sum_{j=1}^{N} [x(t_i+\epsilon_j)+\psi(m_j)] + \eta(t_i). \tag{4}$$

Thus, as opposed to the case of Uk37 samples, the sum is across each of the N foraminifera comprising a sample.

4.3. Detailed example

The foregoing technique is described in detail with respect to a single Mg/Ca record in order to provide greater insight into the implications of the sampling and bioturbation model. We focus on a Mg/Ca record from G. ruber tests in core MD03-2707 (Weldeab et al., 2007) as being broadly representative of our approach and discuss how this analysis compares with that of Uk37 records. MD03-2707 was taken from the Gulf of Guinea, is associated with a mean sedimentation rate of 55cm/kyr, and is sampled at a mean resolution of 37 years. The PLAFOM model (Fraile et al., 2008) indicates a seasonality in G. ruber population in the Gulf of Guinea that peaks between June and October. SSTs exhibit an annual cycle of 3.4°C amplitude at this location (Rayner et al., 2006) and are coolest between June and October. Therefore, the uneven sampling of these SSTs leads to a bias toward cooler temperatures (Fig. 3a,b).

Importantly, the sampling of the seasonal cycle in SST by foraminferal tests contains a significant stochastic component, depending on the individual lifecycle of the approximately 30 samples that are crushed and collected together for each Mg/Ca measurement. This random component of how the seasonal cycle is sampled gives, in this case, a standard deviation between samples of 0.24°C (Fig. 3c,d). Although the number of Mg/Ca tests averaged together would be sufficient to resolve the seasonal variability, the nonuniform distribution leads to a stochastic aliasing of the seasonal variability.

The magnitude of aliased noise is different for every core, depending on the seasonality of SST and foraminifera populations as well as the number of tests averaged together for each Mg/Ca sample, thus necessitating that we model this process independently for each record. A positive correlation (R=0.49, p>0.1) between the variance of the Mg/Ca records and the modern seasonal range of SST at the position of the cores (Table 1) indicates the importance of this process. Note that there is one outlier among the Mg/Ca samples from core MD99-2203 (Cleroux et al., 2012) whose omission would raise the cross-correlation to R=0.85. In contrast, the Uk37 variance is not expected to show such seasonal aliasing and is uncorrelated to the seasonal range. An interesting feature of this effect is that insomuch as Mg/Ca samples are evenly distributed over the year—a feature usually considered to be advantageous—greater aliasing generally occurs because the finite number of samples are distributed over a larger range of seasonal variability.

Bioturbation is the other process in our model that significantly influences recorded variability. For Mg/Ca the influence of bioturbation is a random process that depends on what portions of the seasonal cycle happen to be sampled. To illustrate the effects of bioturbation, we generate synthetic proxy records consistent with the characteristics of the Gulf of Guinea site MD03-2707 following the piecewise spectral representation given in Eq. 2. To generate a record whose variability is consistent with that observed in instrumental SSTs nearest the core site, the Fourier transform of white noise is multiplied by the instrumental SST spectra and then transformed back into a time series. Lower frequency variability that is not covered by the instrumental records are initially parameterized to follow a power-law of $\beta = 1$ (Fig. 4). The correct value of β is uncertain and a search is made over a range of plausible values.

Bioturbation is assumed to extend down $\delta = 10$cm into the sediment, which equates to 182 years in this core, given the average accumulation rate. In the case of Uk37, where the number of samples are essentially infinite, bioturbation leads to a smoothed and time lagged version of the SST record. But in the case of Mg/Ca the discrete sampling discussed in the foregoing paragraph adds variability that the bioturbation only partially reduces. For core MD03-2707 we find that aliasing contributes more variance than bioturbation suppresses, such that the resulting Mg/Ca record is expected to have more variability than the actual SST record. This increase in temperature variance inferred from Mg/Ca records is found to generally hold across the records in our collection.

Finally, measurement noise and other sources of intratest variability are represented by addition of white noise, η. The noise standard deviation must be estimated from the proxy record, along with the value β, and in the case of MD03-2707 has a standard deviation of 0.5°C. The results of aliasing, bioturbatoin, and measurement noise are illustrated in Fig. 5 where a synthetic time series is sampled in accord with that of Uk37 and Mg/Ca samples. The resulting smoothing and aliasing are clearly evident. Note that increasing the number of individual foraminfera in a Mg/Ca sample, N, leads to a more stable value of $\frac{1}{N}\sum \psi(m_j)$ and results that are more consistent with that of Uk37. For $N = \infty$, the Mg/Ca and Uk37 results are identical in our models, excepting a possible mean offset associated with disproportionate sampling of the climatological seasonal cycle.

5. Application of the correction filter

Determining the most suitable correction filter (Eq. 1) for each record requires estimating the two adjustable parameters that define the background variability: the spectral slope β and the standard deviation associated with η. We perform an exhaustive search over the values of $\beta = \{0, 0.1, ...1.9, 2.0\}$ and $\mathrm{STD}(\eta) = \{0, 0.05, ...1.95, 2\}$, searching for the pair of values that minimize the mean square deviation between the logarithm of the observed spectra and the logarithm of the model spectra.

Sea surface temperature time series are generated in accord with each combination of the adjustable parameters, after which the bioturbation and sampling models are applied to produce a synthetic proxy record. This process is repeated 1000 times to approximate the distribution of possible results, with the spectra of the uncorrupted and corrupted version of the synthetic time series being recorded in each instance. Prior to performing the spectral analysis, records are interpolated to a uniform spacing in direct correspondence with the actual proxy record being represented. The average spectral estimate associated with the uncorrupted time series, $S_{\hat{x}m}$, is then divided by the average spectral estimates of the corrupted time series, $S_{\hat{y}m}$, to yield a filter. Following Eq. 1, multiplication of each filter times the corresponding spectral estimate associated with a given proxy record yields our best estimate of the spectrum of SST variability at that site.

5.1. *Test of the filtering approach on synthetic data and estimation of confidence intervals*

Before applying the proxy correction technique described above to the data, we first test its performance on surrogate time series. Surrogate time series are generated in accord with Eq. 2. High frequency variability is realized to be consistent with that observed in instrumental SSTs nearest each site, whereas lower frequencies follow a power-law of $\beta = 1$. For both Uk37 and Mg/Ca, the sample spacing from each core is applied, a 10cm bioturbation depth is assumed, and a 0.25 and 0.45 standard deviation of η is prescribed for Uk37 and Mg/Ca respectively. Additional parameters are also prescribed for each Mg/Ca record comprising the population seasonality from PLAFOM, instrumental SST seasonality, and the reported number of foraminifera tests in each sample (see Fig. 6). To test the effect of the sampling resolution on our method we also include two lower resolution cores in this analysis (MD01-2378 (Xu et al., 2008) and MD95-2043 (Cacho et al., 2001)) which are not used in the remaining part of the study.

The correction algorithm yields more accurate results given more highly resolved records. In particular, β is only well constrained when the sampling interval averages less than 100 years, especially for Mg/Ca records where aliasing of the seasonally cycle is of particular concern. Synthetic records with a larger average sampling interval also show biases in their associated estimates, and we therefore restrict the data used in this study to records with a mean sampling resolution of less than 100 years. We also find that power-law estimates begin to show bias for processes having a true power-law less than 0.1 for Uk37 and less than 0.7 for Mg/Ca records (Fig. 7). As we will show later, application of the spectral correction algorithm to the Holocene proxy records gives values of β near one. Therefore, the synthetic experiments indicate that the application of the spectral correction algorithm will yield accurate results when applied to the data in our collection.

We also use this surrogate approach to estimate the uncertainties associated with spectral estimation and the filtering process. Specifically, we simulate surrogate time series using the estimated β scaling relationship, then corrupt the records according to the properties associated with each actual record, apply the correction algorithm, and estimate the resulting average spectra. This algorithm is repeated one-thousand times, and a chi-square distribution is fit to the ensemble of results at each frequency using moment matching. The reported uncertainty estimates thus include the effects of the proxy correction technique along with the usual uncertainties associated with making a spectral estimate of a noisy and finite process.

5.2. Application to the actual data

Application of the correction filter to the individual Uk37 records leads to a 35% overall reduction in variance or, equivalently, spectral energy (Fig. 8a). The initial power-law associated with the average spectra of 0.98 only changes to 0.96 after correction, indicating that the overall shape of the spectra is only slightly altered. Application of the correction filter to the Mg/Ca records results in a 60% reduction in variance (Figure 8b). This large decrease in energy in the corrected estimates can be traced to the relatively small number of individual foraminfera combined together for each Mg/Ca estimate and the resulting aliasing and intrasample variability. Furthermore, the strongest relative reduction of variance occurs at the highest frequencies, causing β to change from 0.58 to 1.05 for the average Mg/Ca spectral estimate.

Upon applying our correction algorithm, the Uk37 and Mg/Ca power-law scaling coefficients become consistent with one another to within uncertainty, with values of $\beta_{Uk37} = 0.96 \pm 0.07$ and $\beta_{Mg/Ca} = 1.05 \pm 0.07$) (Figure 8c). The variance of Mg/Ca is decreased by the correction algorithm so that, on average, these records are only 20% more variable than the Uk37 records, a discrepancy that is well within the uncertainty in the calibrations for Mg/Ca and Uk37. Note that calibration uncertainty does not influence the power-law estimates, making the Uk37 and Mg/Ca power-law consistency the more stringent indicator of the correction algorithm's adequacy.

Independent information from the reported measurement and replicate measurements of Mg/Ca can also be used to evaluate the correction method. The correction filter has two parameters: β which describes the scaling behavior of the underlying temperature signal and η which describes the standard deviation of the random variations introduced by measurement error and all other processes except those associated with sampling and bioturbation. Values of β and η are determined from a two-dimensional parameter search for minimum misfit between the modeled and observed spectral estimate, and the contours representing this misfit (Fig. 9a-b) indicate that the standard deviation of η is constrained near 0.25°C and 0.5°C for Uk37 and Mg/Ca, respectively. The mean reported replicate error for Uk37 measurements is 0.23°C, corresponding in magnitude to the estimates made here. A close relationship also exists between the inferred and reported errors for Mg/Ca records across the four records for which replicate results are available (Figure 9c). The fact that the correction algorithm gives results that independently agree with the reported replicate error statistics further indicates that it yields accurate results.

6. Summary and conclusion

The differing temperature variability indicated by Uk37 and Mg/Ca records can be reconciled through correcting for the effects of aliasing, bioturbation, and other noise sources. The correction brings the overall variance or, equivalently, the average spectral energy between the Uk37 and Mg/Ca record into greater agreement, reducing the 100% greater Mg/Ca variance to having only 20% more variance. The residual difference can be accounted for by uncertainties in the temperature calibrations applied to either or both of the proxy types. The correction also brings the power-law scaling associated with each proxy into consistency within relatively small uncertainties.

Mg/Ca temperature estimates are strongly affected by aliasing of seasonal and interannual temperature variability due to the limited number of foraminiferal tests used in a given measurement, with additional variability contributed by measurement error, intra-sample variations (e.g. Sadekov et al., 2008), and issues associated with the cleaning processes (Barker, 2003). In contrast, Uk37 temperature estimates comprise a large number of molecules and do not admit seasonal and interannual aliasing. Accordingly, the estimated noise term for Uk37 measurements is about half that of the Mg/Ca proxy. Bioturbation is of secondary importance in this collection of records because they all are associated with high-accumulation rates.

Holocene sea surface temperature variability is found to follow a power-law scaling close to one at timescales between century and millennia. Earlier marine proxy studies found larger-magnitude scaling coefficients (Pelletier, 1998; Shackleton and Imbrie, 1990; Huybers and Curry, 2006), though this is not surprising because they examined variability over glacial-interglacial time scales. Glacial climates and the transition from glacial to inter-glacial climates show different frequency scaling behavior than the Holocene interval (Ditlevsen et al., 1996). We also note that previous studies made no corrections to their spectral estimates, and that such correction could increase the discrepancy insomuch as aliasing contributes energy at high frequencies, but might also increase consistency because bioturbation is expected to have a larger influence on records associated with lower accumulation rates. Regardless, the effect of those corrections on the spectral scaling of glacial-interglacial temperature evolution would likely be smaller than found in our analysis of Holocene record because of much larger amplitude temperature variability and, presumably, a higher signal-to-noise ratio. It will be of interest in future studies to examine how climate spectra vary as a function of background climate, and such an analysis is now more feasible because the present method should, at least in principle, also permit for correcting for artifacts associated with changes in signal-to-noise ratios.

Although seasonal differences in the abundance of proxy indicators are usually regarded as a disadvantage in climate reconstructions because it biases the estimate away from the annual mean (Wunsch, 2009; Laepple et al., 2011), it can be an advantage when one aims to reconstruct the amplitude of climate variability. A site with an equal foraminiferal flux over the year, especially at a site with strong seasonality in temperature, is more prone to aliasing of the seasonal cycle into the recorded signal as the small amount of samples are distributed over the whole range on the seasonal cycle. A further concern does arise, however, that the seasonal distribution of the foraminfera flux is most likely a function of the background climate itself and such nonstationarities have not been accounted for in the present analysis.

There are a number of other processes that might also corrupt proxy records. Uk37 markers might be affected by advection and redistribution (Ohkouchi et al., 2002) or preferential degradation of either one of the long chain alkenones (Hoefs et al., 1998; Gong and Hollander, 1999) or the Mg/Ca rich calcite in foraminifera (Regenberg et al., 2006). At least for degradation influences, we expect that these will mainly act on the trends and not strongly distort the continuum spectra of variability. As noted earlier, all records analyzed here have been detrended prior to making spectral estimates. Given that there is no expectation for the above mentioned sources of error to affect Mg/Ca and Uk37 records equally, the result that both Uk37 and Mg/Ca records show a similar spectrum of variability after correction suggests that the major sources of corruption in the temperature signal have been accounted for. Agreement between the estimated and observed replicate noise values further indicates that no major contributions to error have been overlooked.

These results also highlight the utility of smoothing Mg/Ca records, as is often all ready done in practice (e.g. Marchitto et al., 2010). The reconstructed temperature spectra show strong autocorrelation whereas the noise component is close to being uncorrelated. Therefore, smoothing is expected to more completely suppress noise variance relative to that of the temperature signal and, thereby, to increase the signal-to-noise ratio. It should be possible to design a filter that would optimally increase signal-to-noise ratios in a given Mg/Ca time series or other proxy records. We conclude that the correction algorithm presented here provides a basis by which to more accurately estimate marine temperature variance and its spectral distribution and should provide further insight into how to optimally control for noise sources.

Acknowledgments

We thank Tom Marchitto, Peter DeMenocal, Caroline Cleroux, Peter Stott, and Elizabeth Farmer for information regarding replicates; Jung-Hyun Kim and Guillaume Leduc for providing the GHOST dataset, Igaratza Fraile for providing the PLAFOM results and Stephen Barker, Gerrit Lohmann, Cristian Proistosescu, Andy Rhines, Stephan Dietrich, and Alexander Stine for helpful discussion. We also thank Julian Sachs and an anonymous reviewer for their constructive comments. TL was supported by the Alexander von Humboldt foundation and the Daimler and Benz foundation and PH received support from NSF award 0960787.

References

Arbuszewski, J., deMenocal, P., Kaplan, A., Farmer, E. C., Dec. 2010. On the fidelity of shell-derived 18Oseawater estimates. Earth and Planetary Science Letters 300 (3-4), 185–196.

Barker, S., 2003. A study of cleaning procedures used for foraminiferal Mg/Ca paleothermometry. Geochemistry Geophysics Geosystems 4 (9).

Barnett, T., Hasselmann, K., Chelliah, M., Delworth, T., Hegerl, G., Jones, P., Rasmusson, E., Roeckner, E., Ropelewski, C., Santer, B., et al., 1999. Detection and attribution of recent climate change: A status report. Bulletin of the American Meteorological Society 80 (12), 26312660.

Barnett, T., Santer, B., Jones, P., Bradley, R., Briffa, K., Jan. 1996. Estimates of low frequency natural variability in near-surface air temperature. The Holocene 6 (3), 255–263.

Bendle, J., Rosell-Mele, A., 2007. High-resolution alkenone sea surface temperature variability on the North Icelandic Shelf: implications for Nordic Seas palaeoclimatic development during the Holocene. The Holocene 17 (1), 9.

Berger, W., Heath, G., 1968. Vertical mixing in pelagic sediments. J. mar. Res 26 (2), 134143.

Boudreau, B., 1998. Mean mixed depth of sediments: The wherefore and the why. Limnology and Oceanography, 524526.

Brassell, S. C., Eglinton, G., Marlowe, I. T., Pflaumann, U., Sarnthein, M., Mar. 1986. Molecular stratigraphy: a new tool for climatic assessment. Nature 320 (6058), 129.

Cacho, I., Grimalt, J. O., Canals, M., Sbaffi, L., Shackleton, N. J., Schoenfeld, J., Zahn, R., 2001. Variability of the western Mediterranean Sea surface temperature during the last 25,000 years and its connection with the Northern Hemisphere climatic changes. Paleoceanography 16 (1), PP. 40–52.

Calvo, E., Grimalt, J., Jansen, E., Jul. 2002. High resolution U37K sea surface temperature reconstruction in the Norwegian Sea during the Holocene. Quaternary Science Reviews 21 (1213), 1385–1394.

Cleroux, C., Debret, M., Cortijo, E., Duplessy, J., Dewilde, F., Reijmer, J., Massei, N., Feb. 2012. High-resolution sea surface reconstructions off Cape Hatteras over the last 10ka. Paleoceanography 27, 14 PP.

De Vernal, A., Rosell-Mel, A., Kucera, M., Hillaire-Marcel, C., Eynaud, F., Weinelt, M., Dokken, T., Kageyama, M., 2006. Comparing proxies for the reconstruction of LGM sea-surface conditions in the northern north atlantic. Quaternary Science Reviews 25 (21-22), 28202834.

Dekens, P., Lea, D. W., Pak, D. K., Spero, H. J., 2001. Core top calibration of Mg/Ca in tropical foraminifera: refining paleo-temperature estimation. University of California, Santa Barbara.

Ditlevsen, P. D., Svensmark, H., Johnsen, S., 1996. Contrasting atmospheric and climate dynamics of the last-glacial and Holocene periods. Nature 379 (6568), 810–812.

Doose-Rolinski, H., Rogalla, U., Scheeder, G., Lueckge, A., Rad, U. v., Aug. 2001. High-resolution temperature and evaporation changes during the Late Holocene in the northeastern Arabian Sea. Paleoceanography 16 (4), P. 358.

Emeis, K., Struck, U., Blanz, T., Kohly, A., Vo, M., Apr. 2003. Salinity changes in the central Baltic Sea (NW Europe) over the last 10000 years. The Holocene 13 (3), 411 –421.

Farmer, E. J., Chapman, M. R., Andrews, J. E., Dec. 2008. Centennial-scale Holocene North Atlantic surface temperatures from Mg/Ca ratios in Globigerina bulloides. Geochemistry Geophysics Geosystems 9.

Fraile, I., Schulz, M., Mulitza, S., Kucera, M., 2008. Predicting the global distribution of planktonic foraminifera using a dynamic ecosystem model. Biogeosciences 5 (3), 891911.

Gong, C., Hollander, D. J., Feb. 1999. Evidence for differential degradation of alkenones under contrasting bottom water oxygen conditions: implication for paleotemperature reconstruction. Geochimica et Cosmochimica Acta 63 (3-4), 405–411.

Guinasso, N. L. G., Schink, D. R., 1975. Quantitative Estimates of Biological Mixing Rates in Abyssal Sediments. Journal of Geophysical Research 80 (21), PP. 3032–3043.

Harada, N., Ahagon, N., Sakamoto, T., Uchida, M., Ikehara, M., Shibata, Y., Aug. 2006. Rapid fluctuation of alkenone temperature in the southwestern Okhotsk Sea during the past 120 ky. Global and Planetary Change 53 (1-2), 29–46.

Hasselmann, K., Dec. 1976. Stochastic climate models Part I. Theory. Tellus 28 (6), 473–485.

Hoefs, M. J. L., Versteegh, G. J. M., Rijpstra, W. I. C., Leeuw, J. W. d., Damst, J. S. S., 1998. Postdepositional oxic degradation of alkenones: Implications for the measurement of palaeo sea surface temperatures. Paleoceanography 13 (1), 42–49.

Huybers, P., Curry, W., May 2006. Links between annual, Milankovitch and continuum temperature variability. Nature 441 (7091), 329–332.

Huybers, P., Wunsch, C., 2003. Rectification and precession signals in the climate system. Geophysical Research Letters 30 (19), n/an/a.

Isono, D., Yamamoto, M., Irino, T., Oba, T., Murayama, M., Nakamura, T., Kawahata, H., Jul. 2009. The 1500-year climate oscillation in the midlatitude North Pacific during the Holocene. Geology 37 (7), 591 –594.

Jungclaus, J. H., Lorenz, S. J., Timmreck, C., Reick, C. H., Brovkin, V., Six, K., Segschneider, J., Giorgetta, M. A., Crowley, T. J., Pongratz, J., Krivova, N. A., Vieira, L. E., Solanki, S. K., Klocke, D., Botzet, M., Esch, M., Gayler, V., Haak, H., Raddatz, T. J., Roeckner, E., Schnur, R., Widmann, H., Claussen, M., Stevens, B., Marotzke, J., 2010. Climate and carbon-cycle variability over the last millennium. Clim. Past 6 (5), 723–737.

Kennedy, J., Rayner, N., Smith, R., Saunby, M., Parker, D., 2011a. Reassessing biases and other uncertainties in sea-surface temperature observations measured in situ since 1850, part 2: biases and homogenisation. JGR Atmospheres.

Kennedy, J. J., Rayner, N. A., Smith, R. O., Parker, D. E., Saunby, M., Jul. 2011b. Reassessing biases and other uncertainties in sea surface temperature observations measured in situ since 1850: 1. Measurement and sampling uncertainties. Journal of Geophysical Research 116, 13 PP.

Kim, J., Meggers, H., Rimbu, N., Lohmann, G., Freudenthal, T., Mller, P. J., Schneider, R. R., May 2007. Impacts of the North Atlantic gyre circulation on Holocene climate off northwest Africa. Geology 35 (5), 387 –390.

Kim, J., Rimbu, N., Lorenz, S. J., Lohmann, G., Nam, S., Schouten, S., Rhlemann, C., Schneider, R. R., Nov. 2004. North Pacific and North Atlantic sea-surface temperature variability during the Holocene. Quaternary Science Reviews 23 (20-22), 2141–2154.

Kirchner, J., Jun. 2005. Aliasing in $1/f^{\alpha}$ noise spectra: Origins, consequences, and remedies. Physical Review E 71 (6).

Laepple, T., Werner, M., Lohmann, G., 2011. Synchronicity of Antarctic temperatures and local solar insolation on orbital timescales. Nature 471 (7336), 91–94.

Lamy, F., Rhlemann, C., Hebbeln, D., Wefer, G., Jun. 2002. High- and low-latitude climate control on the position of the southern Peru-Chile Current during the Holocene. Paleoceanography 17, 10 PP.

Lea, D. W., Mashiotta, T. A., Spero, H. J., Aug. 1999. Controls on magnesium and strontium uptake in planktonic foraminifera determined by live culturing. Geochimica et Cosmochimica Acta 63 (16), 2369–2379.

Leduc, G., Schneider, R., Kim, J. H., Lohmann, G., 2010. Holocene and Eemian sea surface temperature trends as revealed by alkenone and Mg/Ca paleothermometry. Quaternary Science Reviews 29 (7-8), 9891004.

Lohmann, G., Pfeiffer, M., Laepple, T., Leduc, G., Kim, J. H., 2012. A model-data comparison of the Holocene global sea surface temperature evolution. Climate of the Past Discussions 8, 10051056.

Lomb, N. R., Feb. 1976. Least-squares frequency analysis of unequally spaced data. Astrophysics and Space Science 39 (2), 447–462.

Lueckge, A., Mohtadi, M., Rhlemann, C., Scheeder, G., Vink, A., Reinhardt, L., Wiedicke, M., Feb. 2009. Monsoon versus ocean circulation controls on paleoenvironmental conditions off southern Sumatra during the past 300,000 years. Paleoceanography 24, 14 PP.

Marchitto, T. M., Muscheler, R., Ortiz, J. D., Carriquiry, J. D., van Geen, A., Dec. 2010. Dynamical Response of the Tropical Pacific Ocean to Solar Forcing During the Early Holocene. Science 330 (6009), 1378–1381.

Mashiotta, T. A., Lea, D. W., Spero, H. J., 1999. Glacialinterglacial changes in subantarctic sea surface temperature and δ< sup> 18</sup> o-water using foraminiferal mg. Earth and Planetary Science Letters 170 (4), 417432.

Mix, A. C., Jun. 2006. Running hot and cold in the eastern equatorial pacific. Quaternary Science Reviews 25 (1112), 1147–1149.

Mueller, P. J., Kirst, G., Ruhland, G., Von Storch, I., Rosell-Mel, A., 1998. Calibration of the alkenone paleotemperature index U37K'based on core-tops from the eastern South Atlantic and the global ocean (60 N-60 S). Geochimica et Cosmochimica Acta 62 (10), 17571772.

Ohkouchi, N., Eglinton, T. I., Keigwin, L. D., Hayes, J. M., Nov. 2002.

Spatial and Temporal Offsets Between Proxy Records in a Sediment Drift. Science 298 (5596), 1224 –1227.

Pelletier, J. D., 1998. The power spectral density of atmospheric temperature from time scales of 10-2 to 106 yr. Earth and Planetary Science Letters 158 (3-4), 157164.

Percival, D. B., Walden, A. T., 1993. Spectral analysis for physical applications: multitaper and conventional univariate techniques. Cambridge Univ Pr.

Rayner, N., Brohan, P., Parker, D., Folland, C., Kennedy, J., Vanicek, M., Ansell, T., Tett, S., 2006. Improved analyses of changes and uncertainties in sea surface temperature measured in situ since the mid-nineteenth century: the HadSST2 dataset. Journal of Climate 19 (3), 446469.

Regenberg, M., Nrnberg, D., Steph, S., Groeneveld, J., Garbe-Schnberg, D., Tiedemann, R., Dullo, W.-C., Jul. 2006. Assessing the effect of dissolution on planktonic foraminiferal Mg/Ca ratios: Evidence from caribbean core tops. Geochemistry Geophysics Geosystems 7, 23 PP.

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 d18O record. Journal of Geophysical Research 116 (D1), D01103.

Richey, J., Hollander, D., Flower, B., Eglinton, T., 2011. Merging late holocene molecular organic and foraminiferal-based geochemical records of sea surface temperature in the gulf of mexico. Paleoceanography 26 (1), PA1209.

Rodrigues, T., Grimalt, J. O., Abrantes, F. G., Flores, J. A., Lebreiro, S. M., Jul. 2009. Holocene interdependences of changes in sea surface temperature, productivity, and fluvial inputs in the Iberian continental shelf (Tagus mud patch). Geochemistry Geophysics Geosystems 10 (7).

Sachs, J., 2007. Cooling of Northwest Atlantic slope waters during the Holocene. Geophysical research letters 34 (3), L03609.

Sadekov, A., Eggins, S. M., Deckker, P. D., Kroon, D., Mar. 2008. Uncertainties in seawater thermometry deriving from intratest and intertest Mg/Ca variability in globigerinoides ruber. Paleoceanography 23, 12 PP.

Schiffelbein, P., Hills, S., 1984. Direct assessment of stable isotope variability in planktonic foraminifera populations. Palaeogeography, palaeoclimatology, palaeoecology 48 (2-4), 197213.

Schneider, B., Leduc, G., Park, W., 2010. Disentangling seasonal signals in holocene climate trends by satellite-model-proxy integration. Paleoceanography 25 (4), PA4217.

Shackleton, N. J., Imbrie, J., 1990. The d18O spectrum of oceanic deep water over a five-decade band. Climatic change 16 (2), 217230.

Sonzogni, C., Bard, E., Rostek, F., Dollfus, D., Rosell-Mel, A., Eglinton, G., May 1997. Temperature and salinity effects on alkenone ratios measured in surface sediments from the indian ocean. Quaternary Research 47 (3), 344–355.

Steinke, S., Kienast, M., Groeneveld, J., Lin, L. C., Chen, M. T., Rendle-Bhring, R., 2008. Proxy dependence of the temporal pattern of deglacial warming in the tropical south china sea: toward resolving seasonality. Quaternary Science Reviews 27 (7), 688700.

Stott, L., Cannariato, K., Thunell, R., Haug, G. H., Koutavas, A., Lund, S., 2004. Decline of surface temperature and salinity in the western tropical Pacific Ocean in the Holocene epoch. Nature 431 (7004), 56–59.

Wallace, J. M., Smith, C., Jiang, Q., Sep. 1990. Spatial patterns of atmosphere-ocean interaction in the northern winter. Journal of Climate 3 (9), 990–998.

Weldeab, S., Lea, D. W., Schneider, R. R., Andersen, N., Jun. 2007. 155,000 Years of West African Monsoon and Ocean Thermal Evolution. Science 316 (5829), 1303–1307.

Wunsch, C., 2009. A perpetually running enso in the pliocene? Journal of Climate 22 (12), 3506–3510.

Xu, J., Holbourn, A., Kuhnt, W., Jian, Z., Kawamura, H., Aug. 2008. Changes in the thermocline structure of the Indonesian outflow during Terminations I and II. Earth and Planetary Science Letters 273 (1-2), 152–162.

Zhao, M., Huang, C., Wang, C., Wei, G., Jun. 2006. A millennial-scale U37K sea-surface temperature record from the South China Sea (8N) over the last 150 kyr: Monsoon and sea-level influence. Palaeogeography, Palaeoclimatology, Palaeoecology 236 (12), 39–55.

Table 1: Proxy data used in the main study

Nr.	Name	Ref.	Lat.	Lon.	sed. rate	mean Δt (yr)	interp. Δt (yr)	duration	proxy	seasonal range
1	MD03-2707	(Weldeab et al., 2007)	2.5	9.4	55	36	100	6600	Mg/Ca G.ruber (pink)	3.4
2	MD98-2176	(Stott et al., 2004)	-5.0	133.4	50	43	100	6818	Mg/Ca G.ruber (white)	3.3
3	MD99-2155	(Farmer et al., 2008)	57.4	-27.9	166	52	100	6440	Mg/Ca G.bulloides	4.2
4	MD99-2203	(Cleroux et al., 2012)	35.0	-75.2	53	19	100	6374	Mg/Ca G.ruber (white)	7.4
5	MD98-2181	(Stott et al., 2004)	6.3	125.8	80	43	150	6969	Mg/Ca G.ruber (white)	1.7
6	MV99-GC41/PC14	(Marchitto et al., 2010)	25.2	247.3	79	90	150	6091	Mg/Ca G.bulloides	7.5
1	SO90-39KG/56KA	(Doose-Rolinski et al., 2001)	24.8	65.9	123	20	100	4880	Uk37	5.4
2	GeoB6007	(Kim et al., 2007)	30.9	-10.3	65	31	100	6750	Uk37	5.2
3	MD97-2151	(Zhao et al., 2006)	8.7	109.9	39	49	100	6020	Uk37	3.9
4	SSDP-102	(Kim et al., 2004)	35.0	128.9	216	61	150	6870	Uk37	12.8
5	IOW 225514	(Emeis et al., 2003)	57.8	8.7	66	72	150	5980	Uk37	11.7
6	CH07-98-GGC19	(Sachs, 2007)	36.9	-74.6	27	68	150	6470	Uk37	12.2
7	OCE326-GGC30	(Sachs, 2007)	43.9	-62.8	30	72	150	6940	Uk37	13.7
8	KR02-06	(Isono et al., 2009)	36.0	141.8	30	51	150	7013	Uk37	12.6
9	MD952011	(Calvo et al., 2002)	67.0	7.6	74	58	200	6450	Uk37	5.3
10	MD01-2412	(Harada et al., 2006)	44.5	145.0	95	73	200	6920	Uk37	14.1
11	D13882	(Rodrigues et al., 2009)	38.6	-9.5	55	53	200	6530	Uk37	5.6
12	IOW 225517	(Emeis et al., 2003)	57.7	7.1	52	94	200	5170	Uk37	11.7
13	JR51-GC35	(Bendle and Rosell-Mele, 2007)	67.0	-18.0	48	98	200	6880	Uk37	4.6
14	GeoB 3313-1	(Lamy et al., 2002)	-41.0	-74.5	107	90	200	6930	Uk37	4.3
15	GeoB 5901-2	(Kim et al., 2007)	36.4	-7.1	13	80	200	5840	Uk37	5.6
16	SO139-74KL	(Lueckge et al., 2009)	-6.5	103.8	106	78	200	5870	Uk37	1.7

Listed are the core number, core name, reference, latitude in °N, longitude in °E, sedimentation rates in cm/ky, the mean sampling interval in yr, the interpolated resolution in years, the duration of the record in years, and the seasonal range of the modern SST in °C. Values regarding the proxy records are computed over the last 7 ky BP, where BP is with respect to 1950 AD.

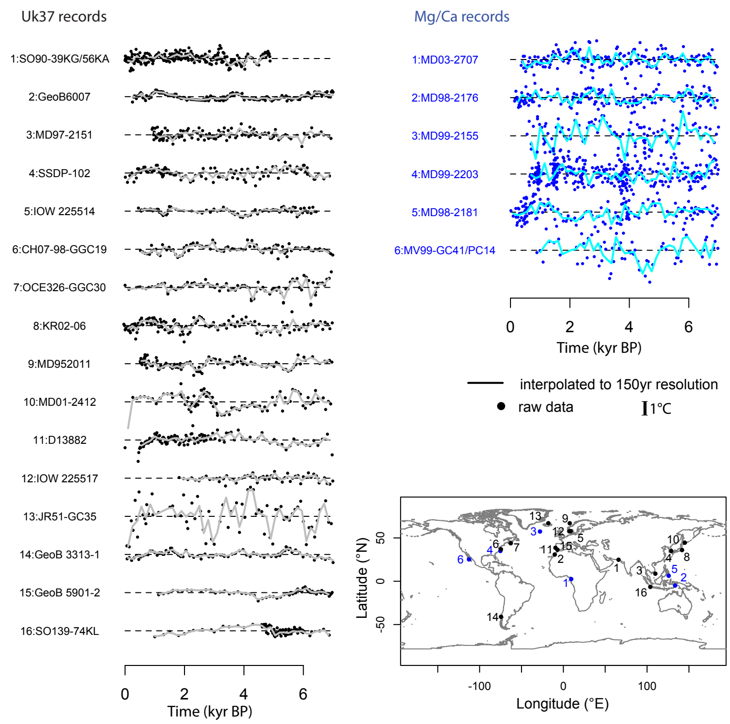

Figure 1: Map of proxy locations and proxy time series of Uk37 (black) and Mg/Ca (blue). All records are linearly detrended. The original data (dots), and the data interpolated to a common 150 year resolution (lines) are shown to facilitate visual intercomparison. The common y-axis scale for all records is show in the legend.

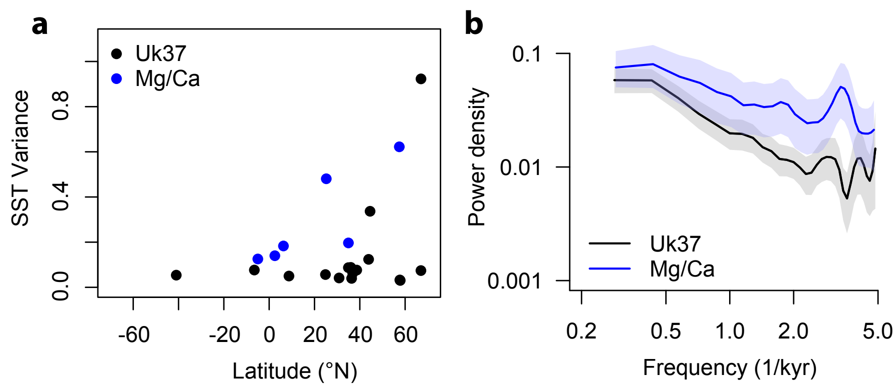

Figure 2: Spatial and spectral comparison of Mg/Ca and Uk37 derived SST variability **(a)** Variance of the proxy time series against latitude for Uk37 and Mg/Ca. All time series were interpolated to 150 yr prior to the variance calculation to minimize influences of the sampling interval on the variance estimate. **(b)** Spectral estimates of Uk37 and Mg/Ca SST records. The mean of the spectral estimates of the globally distributed single records is shown. 95% confidence intervals are indicated by shading

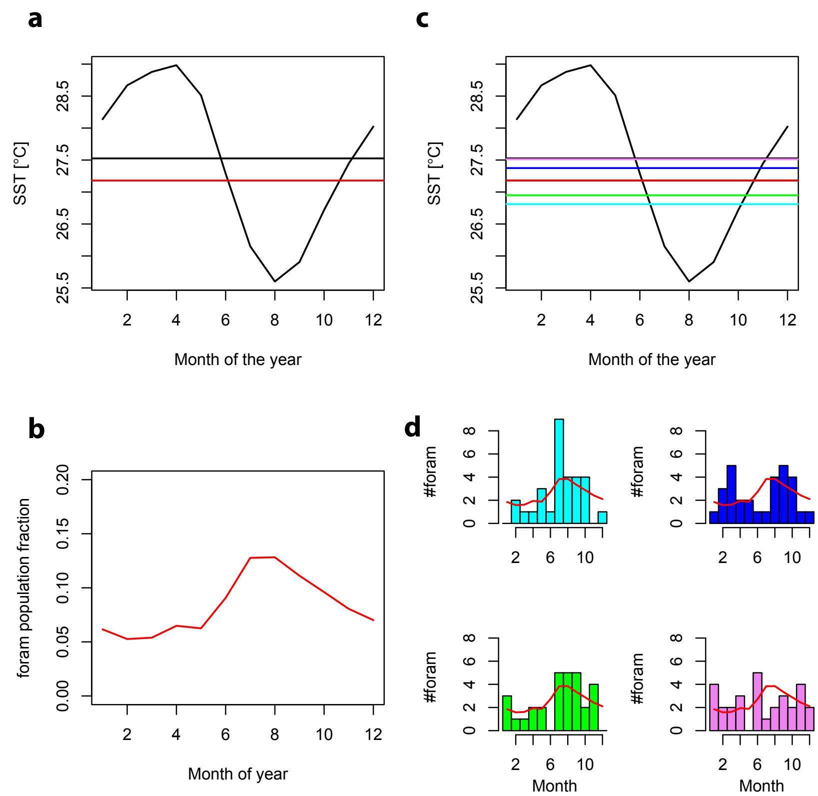

Figure 3: Example of aliasing of the seasonal cycle for the Mg/Ca record from core MD03-2707 (Weldeab et al., 2007). **(a)** The climatological seasonal cycle of sea surface temperature (SST) at the core site from instrumental records (black) and the annual mean SST (horizontal black line). **(b)** The seasonal cycle of G. ruber population at the core site from the PLAFOM model (Fraile et al., 2008). Weighting SST seasonality by population seasonality gives a bias toward cooler temperature (red horizontal line in **(a)**). **(c,d)** Because a typical Mg/Ca sample only consists of 30 foraminifera tests, the sampled seasonal cycle has a substantial stochastic component, indicated by the histograms in **(d)**, that leads to variations in recorded temperature that are indicated by the correspondingly colored horizontal lines in **(c)**.

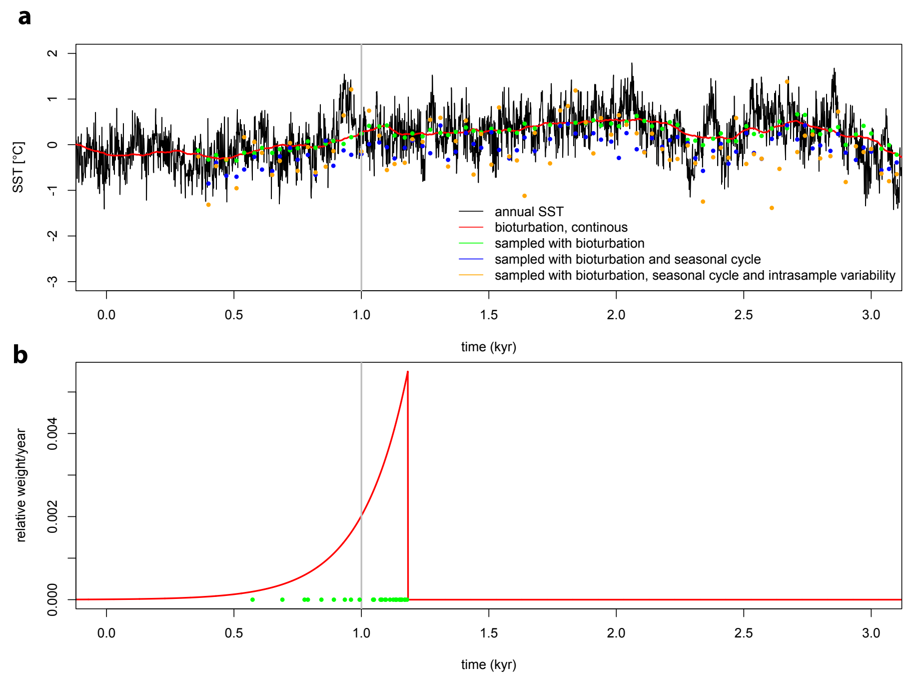

Figure 4: Example of bioturbation and aliasing at core MD03-2707 (Weldeab et al., 2007). **(a)** A synthetic SST time series appropriate for this core site (black) is sampled in various manners and subject to bioturbation. **(b)** The impulse response to bioturbation at this core site (red line) for a sample at 1 kyr (vertical grey line), assuming $\delta = 10$cm. In the case of continuous sampling, bioturbation leads to a smoothed and lagged version of the SST record (red line in (a)). However, the Mg/Ca measurements consist of a limited number of foraminiferal tests (green dots in **(b)**), leading to additional variability in the sampled record (green dots in **(a)**) caused by aliasing of the interannual variability. In addition, aliasing of the seasonal cycle leads to further variability and an offset in the mean (blue dots in **(a)**, c.f. Fig. 3). Finally, intratest and measurement noise contribute additional noise that is estimated to have a standard deviation of 0.5°C for this record (orange dots in **(a)**).

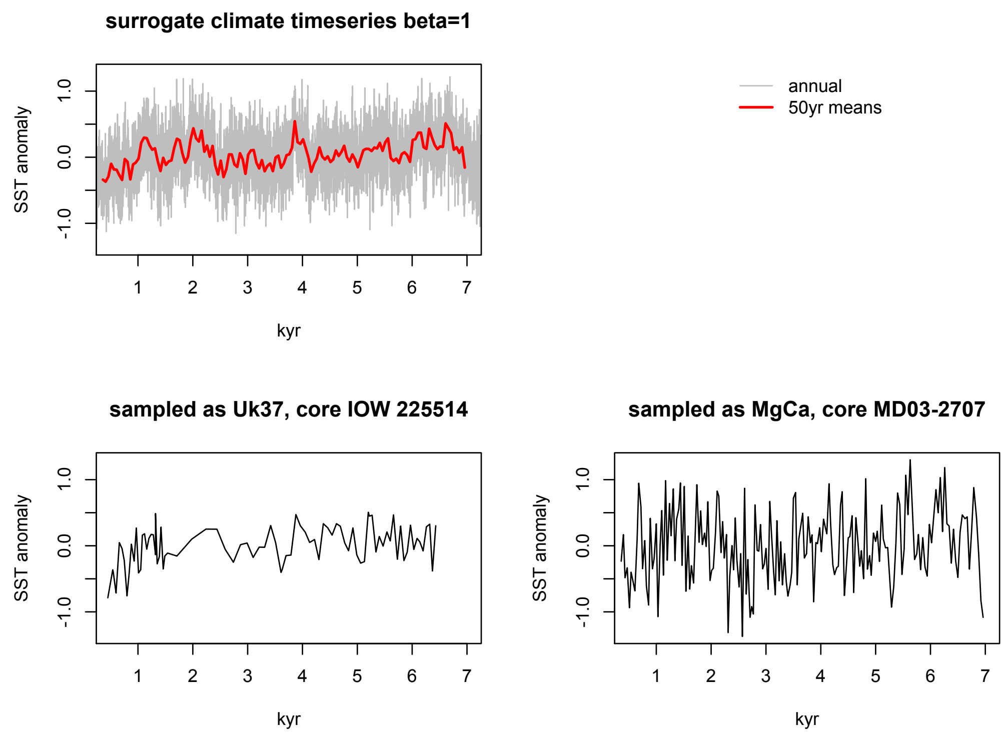

Figure 5: An example of corrupting a synthetic temperature time series according to the noise and sampling regime found in Mg/Ca and Uk37 records. Top left: A realization of a temperature time series simulated using $\beta = 1$. Lower panels: sampling according to the biophysical model for Uk37 (left) and Mg/Ca (right) gives very different proxy time series behavior. The Uk37 record has slightly suppressed variability because of the effects of bioturbation, whereas the Mg/Ca record has greater variability because of aliasing of the seasonal cycle.

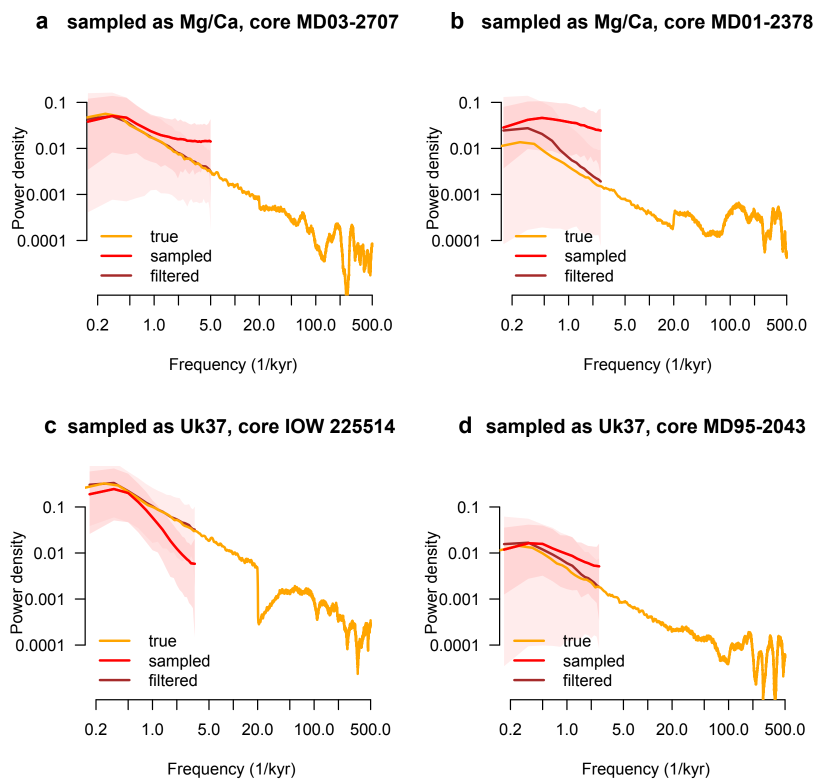

Figure 6: Demonstration of the correction process for synthetic Mg/Ca and Uk37 records. Top row: for Mg/Ca records having sampling resolutions of 36 years (**a**, MD03-2707 (Weldeab et al., 2007)) and 125 years (**b**, MD01-2378 (Xu et al., 2008)). Spectra estimated from random time series having $\beta = 1$ (orange), corrupted using a 10cm bioturbation width and a noise contribution, η of 0.45°C (red). Corruption of the spectral estimate is greatest at high frequencies because unresolved variability is preferentially aliased to these frequencies and because the relatively smaller amount of background variability is more easily disrupted in a fractional sense. After filtering, the original spectrum is recovered in expectation (red), though the 95% confidence interval is increased (shading) owing to uncertainties associated with the correction process. Note that confidence intervals are centered on their respective estimates and are darker where they overlap. Lower row: For Uk37 records having sampling resolutions of 72 years (**c**, IOW 225514 (Emeis et al., 2003)) and 130 years (**d**, MD95-2043 (Cacho et al., 2001)). Time series are simulated using an error term, η of 0.25°C. The corrupted spectra shows less variability at high frequencies because the influence of bioturbation is greater than that of measurement noise and because there is no aliasing. The low-resolution result remains less reliable than the high-resolution one, though the discrepancy is less marked than for the Mg/Ca records.

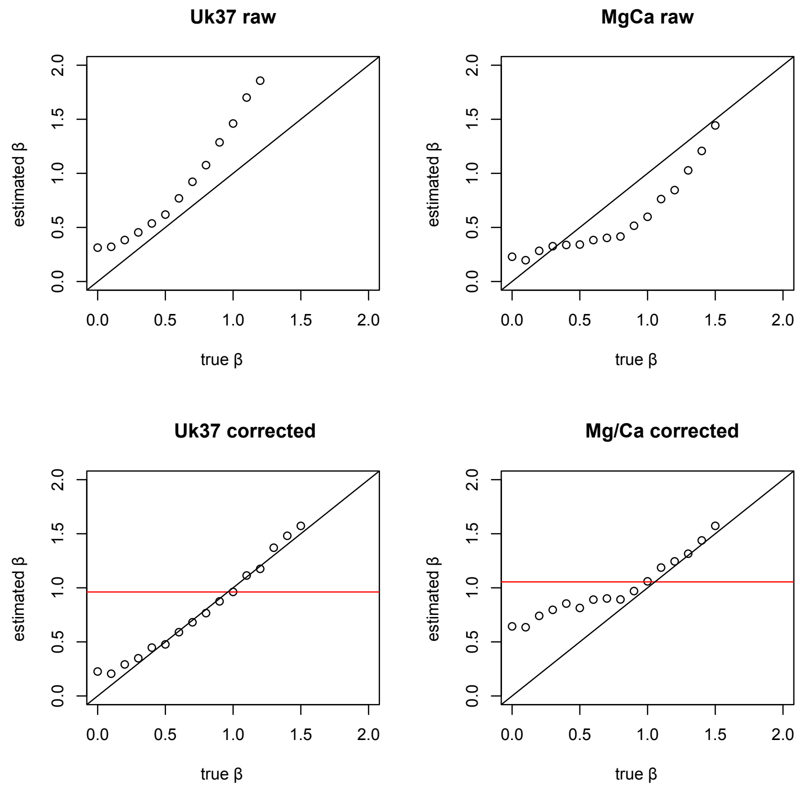

Figure 7: Test of the correction algorithm's ability to recover power-law slopes. Random time series following the spectral model of Eq. 2 with $\beta = \{0, 0.1, ...1.9, 1.5\}$ are simulated using a 10cm bioturbation width, standard deviations of η of 0.25°C and 0.45°C for Uk37 and Mg/Ca, respectively, and individual core parameters for sampling intervals. β is estimated both directly (top row) and after application of the spectral correction algorithm (lower row). For Uk37, the sampling and bioturbation generally leads to an increase in the estimated β when no correction is applied, whereas for Mg/Ca the higher noise level and aliasing leads to a smaller β. The correction filter yields good estimates when the true β is greater than 0.1 for Uk37 and 0.7 for Mg/Ca records, but for shallower power-laws there is a positive bias. The bias results from difficulties in separating signal from noise when both are close to white and because β is constrained to always be positive. Importantly, the method appears unbiased in the range of the reconstructed β from the observations (horizontal red lines).

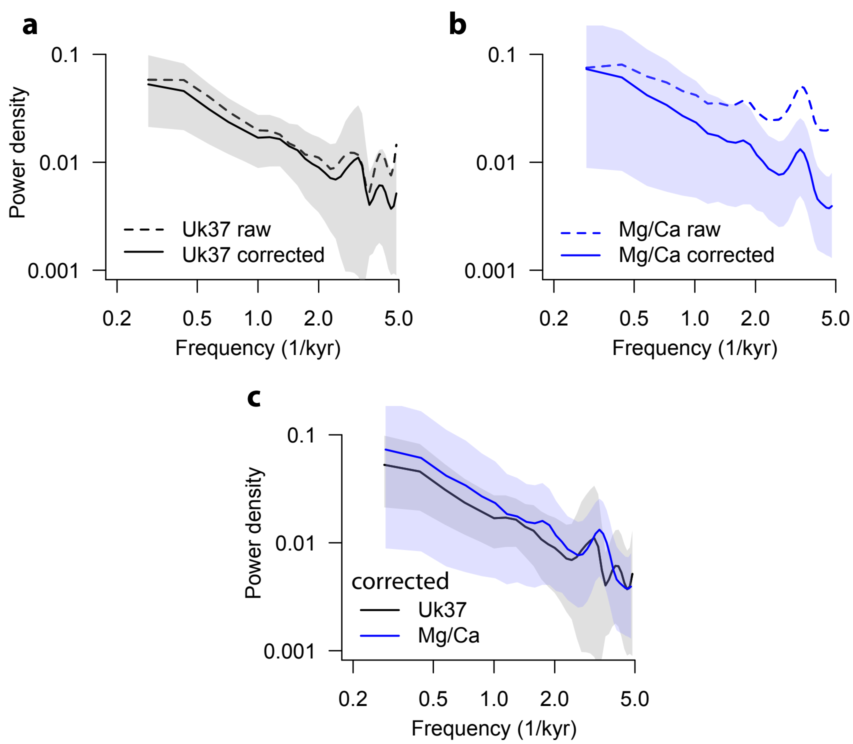

Figure 8: Spectral estimates of the proxy derived SST, raw and after correcting for sampling and noise. **(a)** for Uk37, **(b)** for Mg/Ca, **(c)** comparison of the corrected Uk37 and Mg/Ca spectra. After correction, both spectral estimates are consistent. 95% confidence intervals, indicated by shading, account for the correction process of Uk37 and Mg/Ca where appropriate.

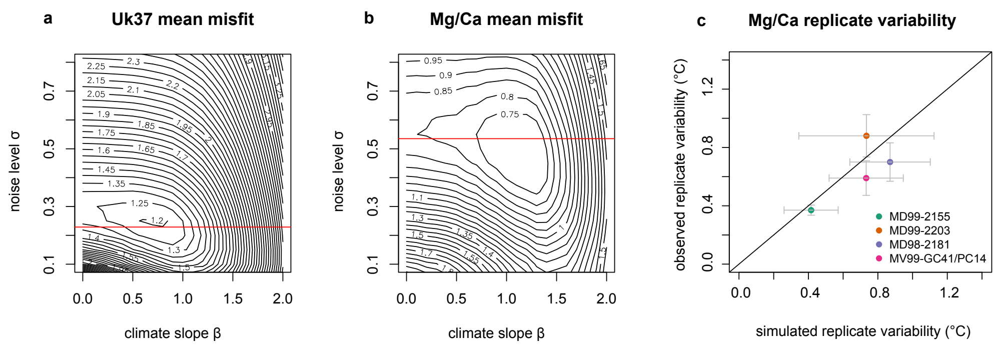

Figure 9: Misfit between observed and modeled spectra for Uk37 **(a)**) and Mg/Ca **(b)** as a function of β and the standard deviation of η. The mean misfit of all cores is shown and indicates distinct minima for both classes of records. Horizontal red lines indicate values of independently reported measurement error for Uk37 from the GHOST database (Leduc et al., 2010) and the mean observed measurement and intrasample error for Mg/Ca. The latter was estimated from the observed replicate error subtracting the contribution from seasonal aliasing indicated by the correction algorithm. **(c)** Comparison of simulated and observed replicate variability of Mg/Ca records. Error bars represent 2 standard deviations. Error bars of the simulated replicate variability are inferred from Monte Carlo experiments using synthetic records designed according to the characteristics of each record. For MD98-2181, the reported replicate standard deviation from Stott et al. (2004) is shown, whereas the other three replicater errors are calculated from data provided through personal communication.