Manuscript prepared for Earth Syst. Dynam.
with version 2015/04/24 7.83 Copernicus papers of the LaTeX class copernicus.cls.
Date: 27 October 2015

Are there multiple scaling regimes in Holocene temperature records?

Tine Nilsen[1], Kristoffer Rypdal[1], and Hege-Beate Fredriksen[1]

[1]Department of Mathematics and Statistics, University of Tromsø The Arctic University of Norway, Tromsø, Norway

Correspondence to: Tine Nilsen (tine.nilsen@uit.no)

Abstract. The concept of multiple scaling regimes in temperature time series is examined, with emphasis on the question whether or not a mono-scaling model can be rejected from observation data from the Holocene. A model for internal variability with only one regime is simpler, and allows more certain predictions on time scales of centuries when combined with existing knowledge of radiative forcing. Our analysis of spectra from stable isotope ratios from Greenland and Antarctica ice cores shows that a scale break around centennial time scales is evident for the last glacial period, but not for the Holocene. Spectra from a number of late Holocene multiproxy temperature reconstructions, and one from the entire Holocene, have also been analysed, without identifying a significant scale break. Our results indicate that a mono-scaling model cannot be rejected as a null model for the Holocene climate up to at least millennial time scales, although it can be rejected for the glacial climate state. The scale break observed from the glacial time ice core records is likely caused by the influence of Dansgaard-Oeschger events and teleconnections to the Southern hemisphere on centennial time scales. From our analysis we conclude that the two-regime model is not sufficiently justified for the Holocene to be used for temperature prediction on centennial time scales.

1 Introduction

The main focus of this paper is the scaling properties in paleotemperature records at centennial and millenial time scales. In particular we study the differences in variability between glacial/interglacial time periods, and we discuss the justification of separating temperature variability on different time scales into distinct scaling regimes. The notion of "scaling" in climatic time series is based on the observation that the natural variability of the Earth's surface temperature can be modelled as a persistent stochastic process, with superposed trends and quasi-periodic modes representing variability which is not included in the noise background. There is a considerable body of literature suggesting that long-range memory (LRM) stochastic processes are good statistical models for de-seasonalised local and global temperature records on time scales from months up to a century or more (Koscielny-Bunde et al., 1996; Rybski et al., 2006; Efstathiou et al., 2011; Rypdal et al., 2013; Østvand et al., 2014). The standard continuous-time stochastic LRM processes are the fractional Gaussian noise

(fGn) and fractional Brownian motion (fBm). The latter is the cumulative integral of the former, and both are said to be scale-invariant (or scaling). The strength of persistence, or memory, in an LRM stochastic process is described by the spectral exponent β; the power spectral density (PSD) takes a power-law form $S(f) \sim f^{-\beta}$. The fGn has $-1 < \beta < 1$ and stationary variance, while the fBm has $1 < \beta < 3$ and a non-stationary variance that grows in time like $\sigma(t) \sim t^{\beta-1}$. The fGn is persistent (exhibits long-range memory) if $\beta > 0$, and is anti-persistent if $\beta < 0$.

Because the terms "scaling regime" and "scale-break/deviation from scaling" might be ambigous, the terms are briefly explained in the following. According to the glossary of Kantelhardt (2011) a scaling regime can be identified only if a power law is valid for scales spanning at least one order of magnitude, be it frequency or time scale. "Deviation from scaling" is synonymous with violation of Kantelhardt's definition. The term "break in scaling" is used to separate scaling regimes that exists in a single time series, where each regime complies with Kantelhardt's definition and is valid for at least one order of magnitude.

Ditlevsen et al. (1996) analysed the scaling in high-resolution ice core data from Greenland. Two different overlapping time series were used to create a composite power spectrum, and from this a break in scaling was identified around centennial time scales. On time scales shorter than centennial the spectrum was flat ($|\beta| \ll 1$), while on longer time scales a non-stationary regime with $\beta \approx 1.6$ was found. One of the time series covers 0-91 kyr BP, and the other 0-3 kyr BP. This procedure of combining different time series into one power spectrum is problematic since the two time series reflect different climate states with different variability. The longer time series is dominated by the glacial state, while the short one contains only Holocene data. The different variability of the two states is seen clearly by direct inspection of the data, e.g., from comparing the Holocene part of the GRIP ice core (Fig. 9a) and the last glacial period from the same ice core (Fig. 11a). The standard deviation in the Holocene time series is less than half of the glacial one, and the latter looks more bursty. The records in Fig. 9a and Fig. 11a can be associated with different stochastic processes. The Holocene record is similar to an fGn with low persistence, while the records from the last glacial period exhibits strong intermittency and is associated with a high spectral exponent, $\beta > 1$.

A simple measure of scaling is the fluctuation analysis (FA) (for a brief review see Rypdal et al. (2013)). It defines the fluctuation function $F(\Delta t)$ as the standard deviation of the data record after it has been filtered by a simple moving average with window width Δt, and hence measures the fluctuation magnitude as a function of scale Δt. The fluctuations are scaling if $F(\Delta t) \sim \Delta t^h$, where h is the scaling exponent (Lovejoy and co-workers denote this exponent by H, and call it the *Hurst exponent*. This convention deviates from the mainstream literature on fractal processes, where the Hurst exponent is $H = h + 1$). The PSD is then of the power-law form with $\beta = 2h + 1$. An issue with FA is which mean value to relate the fluctuation deviation to; the local mean in the window, or the mean of the entire data record. The latter is problematic if the fluctuations are monotonically

growing with increasing scale. This problem can be circumvented by convolving the data record with the simple antisymmetric Haar-wavelet (Lovejoy and Schertzer., 2012) rather than performing the moving average. For fluctuations growing with scale it measures fluctuation differences versus scale, whereas for fluctuations decreasing with scale it measures fluctuation relative to the local mean. One feature that the PSD, FA and Haar fluctuation share with many other measures of scaling is that it is sensitive to trends and large-scale oscillations, i.e., it is often not able to discriminate between such variability and true scaling behaviour.

Pelletier (1998) estimated the power spectra and scaling exponents from a Deuterium record from the Vostok ice core as well as from instrumental local data, and also created composite spectra from the records. Huybers and Curry (2006) and Lovejoy and Schertzer (2012) have studied the scaling in multiple proxy data sets covering time scales from years to millions of years. Both report a break in scaling from fluctuations decreasing with scale ($h < 0,\ \beta < 1$) to fluctuations increasing with scale ($h > 0,\ \beta > 1$) on a transition time scale $\tau_c \sim 10^2$ yr. The break in scaling is seen from composite spectra of paleotemperature records based on different proxies and reconstruction techniques, where many of the records span hundreds of kyr. Since glaciation is the dominating climate state in the Quaternary, the spectra obtained in those papers are typical for glacial climate. Huybers and Curry (2006) suggest that the power-law continuum in the spectrum of surface temperature on time scales between one year and a century is a result of an inverse cascade in frequency space driven by the seasonal cycle forcing. The non-stationary scaling regime from century time scale and longer is proposed to be the result of a nonlinear response to the Milankovitch cycle forcing. From the composite spectra, they infer scaling exponents in the range $\beta = 0.37 - 0.56$ for time scales $\tau < 10^2$ yr, and $\beta = 1.29 - 1.64$ at longer time scales. Lovejoy and Schertzer (2012) introduce three different scaling regimes: the "weather" regime ($\beta \approx 2$ for time scales up to 10 days), the "macroweather" regime ($\beta \approx 0.2$ for time scales from 10 days to 10^2 yr), and the "climate" regime ($\beta \approx 1.4$ for time scales from 10^2 yr and longer). Common for the studies mentioned is that they don't make a distinction between glacial and Holocene spectra.

Other results reported in the literature support our hypothesis of different scaling in glacial and interglacial climate, with the scale break at centennial time scales absent for the Holocene. Blender et al. (2006) analysed the scaling properties of a 10 kyr long general circulation model (GCM) climate simulation, and no scale break can be detected at centennial time scales. Lovejoy et al. (2013) make similar observations and conclude that GCM's do not predict climate, only macroweather. Roe and Steig (2004) found by using a short-range memory autoregressive (AR) model that the characteristic time scales for paleotemperature ice core records were significantly shorter during the Holocene than during the last glacial period. This study is important for our reasoning, but the idea needs to be adapted to a long-memory model. We will separate the ice core records into glacial and interglacial time series and demonstrate the fundamentally different scaling properties of these climate states, and we will analyse other temperature reconstructions for the Holocene in search for a detectable

scale break.

The paper is organized as follows: in Sect. 2 we address the issues of uncertainties and limitations of proxy-based reconstructions, and the implications for the existence of separate scaling regimes are discussed. Sect. 3 describes the scaling analysis methods employed, and information about the model and the data used can be found in Sect. 4. The results from the analysis are presented in Sect. 5, and discussion and conclusion follow in Sect. 6.

2 The concept of multiple scaling regimes in the Holocene

Lovejoy and Schertzer (2012) identify two scaling regimes in a number of Holocene temperature records. In instrumental data the transition time τ_c is found to be 10-30 yr, in proxy/multiproxy reconstructions it is 40-100 yr, while for one of the ice core paleotemperature records it is approximately 2000 yr. Hence, it seems difficult to identify a universal τ_c from the data examined in that paper. For the proxy/multiproxy reconstructions that were analysed in Lovejoy and Schertzer (2012), the time period 1500-1979 was selected because it was common to all reconstructions, and the medieval warm period was avoided. However, by starting in the Little Ice Age the series are strongly influenced by steadily increasing solar, as well as anthropogenic, forcing. A pronounced linear trend has strong effect on the estimate of the scaling exponent from power spectra unless the time series is linearly detrended. This is also the case for the Haar fluctuation analysis, which was also applied in Lovejoy and Schertzer (2012).

Many of the papers cited in Sect. 1 present composite spectra based on instrumental and/or proxy data for scaling analysis. There are many problems related to this, in addition to the already mentioned aspect of combining time series from the Holocene/glacial climate state. The various data sets are representative for different degrees of spatial averaging. This will affect the shape of the spectra and the estimates of the scaling exponent, because the high-frequency variability is reduced with increasing degree of spatial averaging. For the instrumental data we can obtain global averages, while proxy/multiproxy time series represent local, regional or at best hemispheric temperature. There is an important difference between composite spectra and spectra from multiproxy reconstructions. Multiproxy reconstruction methods generally take geographical weighting into account, and the aim is to obtain realistic high/low frequency variability throughout the time period covered by the reconstruction. Most composite spectra, on the other hand, do not handle these aspects in a satisfactory manner. An example of a well designed composite spectrum can be found in Laepple and Huybers (2014), where the proxy records have been corrected for noise, and the instrumental data used were extracted from the same location where the proxy records were sampled. No attempt was made to estimate a scaling exponent from this spectrum, but it is clear that no scale-break can be observed around centennial time scales. Correcting for proxy noise is outside the scope of the present study,

and we will therefore avoid studying composite spectra based on different reconstructions spanning different time scales and climatic states, and on different proxies and reconstruction techniques.

The techniques used to estimate the scaling exponent have inherently higher uncertainties on the longest time scales, due to sparse data on these time scales. A rule of thumb is that the scaling properties for a time series of length N should be estimated only up to time scales $N/4$, since the uncertainty on time scales longer (frequencies lower) than this is too large to make meaningful estimates. Suppose, for instance, we want to establish that we have scaling in annual data on scales up to 100 yr. Then we need a series which is 400 yr long. If we want to establish the existence of a different scaling regime on time scales longer than 100 yr for a time series, we need to know with reasonable certainty the spectral estimates up to one millenium. As we will demonstrate in Sect. 5, this implies that we need record lengths spanning several millennia to bring the uncertainty of β below the limit needed to reject the mono-scaling hypothesis. More generally, we need a record length of the order of 10^2 times the smallest time scale of a scaling regime.

Instrumental temperature data are not included in our analysis of multiple scaling regimes because previous studies do not show pronounced breaks in the scaling after detrending to account for influences from anthropogenic warming (Rypdal et al., 2013). The series are too short to detect scale breaks at centennial time scales. However, we use instrumental data in Sect. 3 for illustration of interesting features of various techniques for scaling characterisation. Detection of scaling properties in regional or hemispheric proxy/multiproxy temperature time series is possible but not optimal, since these records generally cover the past 2000 years or less, and, even though they are usually given with annual resolution, some are effectively filtered to vary smoothly on annual time scale, and have effective resolution from 5 to 10 yr. For some of the available proxy/multiproxy reconstructions an enhanced power can be inferred relative to a mono-scaling spectrum on time scales longer than a century, but data are too sparse to show that this enhanced power represents a new scaling regime.

The deep ice cores sampled at Greenland and in Antartica provide the most suitable data sets for studying possible scale breaks in both the Holocene and the last glacial period, due to the high temporal resolution and long duration. It was argued by Lovejoy and Schertzer (2012) that the Holocene part of the Greenland ice core GRIP $\delta^{18}O$ is exceptionally stable compared to other proxy-based reconstructions of North-Atlantic temperature, but we observe similar stability also for the stable isotope records from Greenland ice cores GISP2 and NGRIP, and the Antarctic ice cores EPICA, Vostok and Taylor dome. When comparing the Holocene stable isotope records from deep ice core records with temperature reconstructions from other regions and based on other proxies, one needs to keep in mind that local variability and the sensitivity of the various proxies to changes in their environment will affect the resulting proxy record. A proxy record is based on an imperfect recorder of a climatic variable of interest, and some sources of noise are not related to climatic processes. Unfortunately, there is no such thing as a temperature reconstruction based on a single proxy that is representative for global temperature variability. In this manner, all temperature reconstructions

are exceptional, because the proxies record some local variability and are subject to noise from a number of sources.

3 Methods

In our data analysis we have used a number of tools, which are described in detail in the following subsections. In general, the periodogram estimator for the power spectral density (PSD) is applied for the scaling analysis of paleotemperature time series in our study, and are the basis for significance testing when studying potential scale breaks in the time series. We have chosen the PSD because most readers are familiar with it, and because it is fully adequate for the purpose. Wavelet scalograms are not used to estimate scaling properties, but to visualize particular features in some of the time series. Structure functions are discussed to point out the general importance of higher-order statistics and to justify a monoscaling model (fractional Gaussian noise) for the internal temperature variability in the Holocene. The Haar fluctuation function is discussed because it is strongly advocated as the ultimate scaling analysis technique by e.g., Lovejoy and Schertzer. (2012).

3.1 Estimation of Power Spectral Density

The periodogram is applied as an estimator for power spectral density (PSD) for evenly sampled time series of length N. The stable istope records from Greenland and Antarctic ice cores have been linearly interpolated to obtain even sampling in time. All other records were already provided with even sampling. The periodogram is defined here in terms of the discrete Fourier transform H_m as $S(f_m) = (2/N)|H_m|^2$, $m = 1, 2, \ldots, N/2$. The sampling time is the time unit, and the frequency is measured in cycles per time unit: $f_m = m/N$. $\Delta f = 1/N$ is the frequency resolution and the smallest frequency which can be represented in the spectrum, while $f_{N/2} = 1/2$ is the Nyquist frequency (the highest frequency that can be resolved). The periodogram has a poor signal to noise ratio, but since we are interested in studying the overall shape (scaling) of the spectrum, and not the power at specific spectral peaks this is not a problem here. By presenting the periodogram in a log-log plot, the scaling exponent β can be estimated by a linear fit to the power spectrum; $\log S(f) = -\beta \log f + c$. In the present study the periodogram is log-binned before fitting to ensure that all time scales are weighted equally (Østvand et al., 2014).

We have also considered other spectral estimators for the unevenly sampled stable isotope data from ice cores, such as the Lomb-Scargle periodogram (LSP), (Lomb, 1976; Scargle, 1982), or correlation slotting, Rehfeld et al. (2011). The main motivation for looking into different spectral methods is to compare biases in the spectra that could be wrongfully interpreted as breaks in the scaling. The papers by Rehfeld et al. (2011) and Broersen et al. (2000) demonstrate that irregularly sampled data cause various problems for all spectral techniques. Slotting can be problematic because the covariance estimators may not be positive semi-definite, and could hence give negative values in

the spectrum. Interpolation leads to underestimation of the spectral power at high frequencies, while the Lomb-Scargle periodogram suffers from the opposite bias: overestimation of the spectral power at high frequencies. The skill of the LS periodogram is, as demonstrated by Rehfeld et al. (2011), dependent on the skewness of the distribution of sampling intervals. The bias will therefore differ from dataset to dataset. We have tested the performance of the method on surrogate data mimicking the ice core proxy data under study. The detailed results for this test are shown in supplementary material. In general the method performs well, but not much better than interpolation plus standard periodogram. Results presented in the main paper are obtained using only interpolation and the standard periodogram. Scaling analysis of the ice core data based on the LSP is included in the supplementary material section 1.

3.2 Wavelet scalogram

The continuous wavelet transform is the convolution between a time series $x(t)$ and the rescaled mother wavelet $\Psi(t)$;

$$W(t,\tau;x(t),\Psi(t)) = \int_{-\infty}^{\infty} x(t')\frac{1}{\sqrt{\tau}}\Psi^*\left(\frac{t'-t}{\tau}\right)dt', \tag{1}$$

where the asterisk indicates complex conjungate. The wavelet scalogram (WS) is defined as $|W(t),\tau)|^2$, and is plotted versus time and time-scale. The WS is used here as a supplementary tool to the Fourier spectra. Time segments before and after the time interval where we have data were padded with zeros, as described in Torrence and Compo (1998). The region in (t,τ)-space affected by edge effects is the region above the white line in the upper part of the WS-plot shown in, e.g., Fig. 10. Due to the uneven sampling of the data in this study, linear interpolation has been performed prior to computing the WS. At each time t there is a characteristic sampling period in the orginal time series, and hence a Nyquist period. This Nyquist period is marked as the lower white curve in the WS plots. The WS below that curve does not reflect observed variability.

We have chosen two wavelet functions as the basis for our study: the Morlet wavelet which is complex valued, and the Mexican hat wavelet (second derivative of a Gaussian) which is real valued. The wavelet scalograms from these two wavelet functions provide different information. The Mexican Hat wavelet function resolves the timing of spectral peaks precisely, while the scale resolution is poor. For the Morlet wavelet function the opposite is true.

3.3 Structure functions and scaling function

A plethora of estimators have been developed for computing characteristic exponents for monoscaling long-range memory (LRM) processes. They all have strengths and weaknesses, but have in common that they give nonsense if the signal is not a monofractal process. Common for many papers by "LRM-skeptics" is the uncritical use of cookbook recipes for such estimators to data that are not

fractal, for instance climatic time series dominated by a specific trend (e.g., Mann (2011)). Hence, more important than estimating a characteristic exponent is to examine the general scaling characteristics of the data. A classical and useful method is to examine the probability density functions (PDFs) of the time series as it varies on different time scales. Rather than computing and plotting the PDFs, it is more common to compute the statistical moments of order q of the distribution, and then plot these moments as a function of time scale Δt. Given a stochastic process $x(t)$, the moments

$$S_q(\Delta t) \equiv E[|x(t+\Delta t) - x(t)|^q], \tag{2}$$

are called the structure functions (SFs) of the process. If the process is sampled at discrete times $t = 1, \ldots, N$ the empirical moments $\hat{S}_q(\Delta t) = (N - \Delta t)^{-1} \sum_{i=1}^{N-\Delta t} |x(i+\Delta t) - x(i)|^q$ constitute estimates of the structure functions. For large Δt (when Δt is no longer a small fraction of N) the number of independent terms becomes small, and the statistical uncertainty of the estimate becomes large. A useful rule is that this limits the scales we can investigate to $\Delta t < N/4$.

Let us assume that $x(t)$ is a self-similar, Gaussian, non-stationary stochastic process, i.e., a fractional Brownian motion (fBm). Then the fluctuation function $F(\Delta t)$ is monotonically increasing ($h > 0$, $\beta = 2h + 1 > 1$), and the structure functions defined on $x(t)$ take the form,

$$S_q(\Delta t) \equiv E[|x(t+\Delta t) - x(t)|^q] = \Delta t^{\zeta(q)}, \tag{3}$$

where $\zeta(q) = hq$ is the *scaling function* of the self-similar (monofractal) process. By taking the logarithm of Eq. (3) we find the linear relationship between $\log S_q(\Delta t)$ and $\log \Delta t$, where $\zeta(q)$ is the constant of proportionality, hence the SFs appear as straight lines in log-log plots with slope $\zeta(q)$. The scaling $\zeta(q)$ is a linear function of q only if the process is monofractal. If the SFs have the form Eq. (3) (i.e., if the SFs are straight lines in log-log plot) so that $\zeta(q)$ is defined, but $\zeta(q)$ is not a linear function, then the process is multifractal. If the SFs are not straight, the process is neither multi- nor monofractal, but it may still have a bursty or *intermittent* appearance.

If the process is stationary, with decreasing fluctuation function ($h < 0$, $\beta < 1$), it is denoted a fractional Gaussian noise (fGn), and the SFs are constant (flat) and contain no other information than the stationarity. However, as we shall see below, the SFs may still contain some useful information about the deviation from scaling if there are trends or oscillations in the data. In order to expose the scaling properties (if any) of the process, the trick is to form the cumulative sum $y(i) = \sum_{j=1}^{i} x(j)$ and then compute the SFs from this sum. The resulting SFs take the form

$$S_q(\Delta t) \equiv E[|y(t+\Delta t) - y(t)|^q] = \Delta t^{\zeta(q)} \tag{4}$$

where $\zeta(q) = H_u q$ is the scaling function, and H_u the scaling exponent, of the cumulative sum $y(t)$. H_u is also called the Hurst exponent of the stationary proxess $x(t)$, and is related to the spectral exponent through $H_u = (\beta+1)/2$. Since we have that $h = (\beta-1)/2$ for a nonstationary process we often see the relation $H_u = h+1$, although for SFs the two scaling exponents are defined for different

classes of processes. However, there are other estimators, like the periodogram and the wavelet variance (including the Haar wavelet fluctuation employed by Lovejoy and Schertzer (2012)), which work on both stationary and nonstationary processes. For these estimators this relation is meaningful.

The usefulness of the structure-function approach is illustrated in Figure 1. Panel (a) shows the eight structure functions corresponding to $q = 1, 2, \ldots, 8$ for the monthly global mean surface temperature (GMST) for the period 1880 – 2010 derived from the HadCRUT3 data set.

The underlying scaling of the noise is exposed by computing the SFs for the cumulative sum, as shown in Figure 1b. However, the corresponding scaling function, shown by the upper line in Figure 1d, has the slope $H_u \approx 1$, which is always the case for a signal dominated by a strong trend. The true scaling of the noise appears after a second-order polynomial fit to the record has been subtracted. The SFs for the cumulative sum of the detrended record is shown in Figure 1c, and the corresponding scaling function by the lower line in Figure 1d. This line has Hurst exponent $H_u \approx 0.85$. The straight appearance of the scaling function tells us that the GMST is monofractal, and simple tests on the PDFs at different scales show that it is Gaussian (Rypdal and Rypdal, 2010).

The scaling functions in Figure 1d have been computed from the slopes of the SFs in the regime of scales $\Delta t < 10$ yr where the SF-curves are straight. The bending of these curves for large scales are due to oscillatory modes on periods around 20 yr and 70 yr. Estimators employing only second-order statistics, like the periodogram or the Haar fluctuation employed by Lovejoy and Schertzer (2012) are not able to distinguish between scale-invariant fluctuations and variability due to trends or oscillations. Careful application of higher-order statistics like higher-order SFs has this ability. This can give us the possibility to separate distinct non-scaling dynamical features from the scaling, persistent noise background. In section 5.1 we shall demonstrate this usefulness on a multiproxy temperature reconstruction spanning two millennia.

3.4 The Haar fluctuation function

The Haar fluctuation was briefly mentioned in Sect. 1. The simple definition given in Lovejoy et al. (2013) starts by defining the fluctuation on scale Δt as,

$$\Delta x_t(\Delta t) = \left| \frac{2}{\Delta t} \sum_{i=t}^{t+\Delta t/2} x_i - \frac{2}{\Delta t} \sum_{i=t+\Delta t/2}^{t+\Delta t} x_i \right|, \tag{5}$$

and then the Haar structure function is given by

$$S_q^{\text{Haar}}(\Delta t) = \frac{1}{N - \Delta t} \sum_{t=1}^{N-\Delta t} \Delta x_t(\Delta t)^q. \tag{6}$$

The Haar fluctuation function used extensively by Lovejoy and Schertzer in arguing for the existence of transitions between scaling regimes is defined as

$$F^{\text{Haar}}(\Delta t) = \sqrt{S_2^{\text{Haar}}(\Delta t)}. \tag{7}$$

In Figure 2 we illustrate some features of the Haar fluctuation applied to the instrumental GMST. Figure 2a shows the GMST and a trend computed by fitting a second order polynomial. The black thick curve in Figure 2b is the Haar fluctuation function computed from the record in a log-log plot (the upper curve is shifted by a factor 10).

There are (at least) two different ways to model this record as a simple stochastic process. One is to assume that it is a linear combination of an fGn ($\beta < 1$) and an fBm ($\beta > 1$). The former will dominate the fluctuation function on the small scales, and the latter on the long scales. Hence this is a model that exhibits a scale break and two scaling regimes. We have estimated the slopes of the fluctuation function in these two regimes, and found $H \approx -0.1$ ($\beta = 2h + 1 \approx 0.8$) on the short scales, and $h = 0.3$ ($\beta \approx 1.6$) on the long scales. We then computed the weights of each process from using the estimated variance of the GMST on the shortest and longest scales, respectively, and computed an ensemble of realisations of their linear combination. The red curves in Figure 2b constitute 20 realisations in such an ensemble (multiplied by a factor 10). These curves demonstrate that the observed Haar-fluctuation is consistent with this model, but also that the uncertainty in the model prediction on scales longer than a decade is so large that that the observed fluctuation function here could also be consistent with a model where $h < 0$ ($\beta < 1$). Another obvious way to model the record is as a linear combination of the fGn and the quadratic trend. The resulting ensemble is shown as the blue curves in Figure 2b. The observed record is consistent with this model too, but the big difference is that in this case the model prediction on large scales is much more certain, and hence constitutes a "better" statistical model. An extra bonus is that the quadratic trend is physically well understood, since it corresponds closely to present knowledge about greenhouse forcing.

Thus, we have the choice between explaining the observation with a poor statistical model (many parameters, large prediction uncertainty, and no physics explaining the scale break) and a much better model (fewer parameters, lower uncertainty, and a clear physical explanation).

A common estimator for scaling exponents is the wavelet variance, i.e., to plot the variance of the wavelet coefficients versus scale in a log-log plot, and it is common to normalise the wavelet such that the slope for an fGn will be the spectral exponent β. In practice this can be obtained by squaring the Haar fluctuation $S_2^{1/2}(\Delta t)$ and multiplying by the scale Δt, i.e., we compute $\Delta t S_2(\Delta t)$. In Figure 3 we generate an ensemble of 10 fGns of 1000 data points with $H = -0.1$ ($\beta = 0.8$) and plot $S_2^{1/2}(\Delta t)$ for all realisations in the ensemble in Figure 3a , and $\Delta t S_2(\Delta t)$ for the same ensemble in Figure 3c. If we use the entire ensemble to estimate the slope we will get a quite accurate result (getting better the larger the ensemble), but if we estimate the slope from one realisation we make greater errors if we include the longer time scales in the fit. This is why some authors recommend not to include longer scales than $1/4$ of the record length and others recommend no more than $1/10$. If we fit a straight line we have no reason to believe that the Haar fluctuation gives less accurate estimates than the Haar wavelet variance. However, by inspecting the two for one particular realisation in the ensemble, as is done in Fig. 3b and d, we observe that the Haar

fluctuation can present a break in the curve that visually is much more pronounced than in the corresponding Haar wavelet variance. This is nothing but a visual illusion (the two plots contain the same information); a curve that changes its slope from negative to positive is more easily perceived to represent different qualities than the curve that only changes its positive slope somewhat. Since all the curves are produced from realisations of the fGn, all scale breaks are spurious, and caused by the diverging statistical uncertainty in the high-scale end. Thus, if the underlying scaling is close to $1/f$ noise, then weak trends or oscillations tend to appear as breaks in the Haar fluctuation curve, but are much less visible in the wavelet variance curve. A sound approach to graphical tools like this should avoid visualisations like the Haar fluctuation, which exaggerates such spurious breaks.

4 Data

The scaling is analysed in seven proxy/multiproxy temperature reconstructions representing late Holocene temperature, one temperature reconstruction representing the entire Holocene time period, in addition to six reconstructions of stable isotope ratios from the deep ice cores GRIP, GISP2 and NGRIP from Greenland, and EPICA, Taylor dome and Vostok from Antarctica. Information and analysis results from GISP2, NGRIP, Taylor and Vostok is provided in supplementary material. From the available ice core time series we extract sub-series covering only the Holocene and only the last glacial period, respectively. For the GRIP ice core we also extract a time series covering 0 – 85 kyr BP. Since the exact timing of the transition between the Holocene and the last glacial period is slightly different for Greenland and Antarctica, we have chosen the start and end of the time series carefully for each series, such that the transition is not contained in any of the "Holocene only" or the "glacial only" time series.

4.1 Proxy/multiproxy late Holocene temperature reconstructions

We have chosen seven proxy- or multiproxy based temperature reconstructions for our study, and in order to avoid the trend effect from anthropogenic warming we have discarded data after 1850 AD (see Table 1). All time series are given with annual resolution. A few of the reconstructions are based partly on the same raw proxy records, but we include all since the reconstruction methods are different. The Jones et al. (1998) multiproxy reconstruction represents northern hemisphere temperature. The Briffa et al. (2001) reconstruction represents the continental region 20° N – 90° N and is constructed from tree rings. The Esper et al. (2002) reconstruction is also based on tree rings and represent the continental region 30° N – 80° N. The Huang (2004) reconstruction is based on borehule temperatures, integrated with instrumental temperatures and the multiproxy reconstruction by Mann et al. (1999). The Moberg et al. (2005) multiproxy reconstruction represent northern hemisphere temperature, and is smoothed on the shortest time scales, so estimates of the scaling exponents are restricted to time scales from 4 years and longer. The Mann et al. (2009) multiproxy reconstruction

represents global temperature, and is smoothed up to decadal time scales. The Neukom et al. (2014) multiproxy reconstruction represents southern hemisphere temperature.

4.2 Multiproxy full Holocene temperature reconstruction

The temperature reconstruction described in Marcott et al. (2013) is included in our study because it covers the entire Holocene time period. The reconstruction is based on 73 proxy records with temporal resolution varying from 20 to 500 years. The spatial distribution of proxy data is near global, and there is a high percentage of data sets from marine sites. The proxy records were interpolated to 20-year resolution before constructing the temperature reconstruction, and the final record is presented with a 20-year resolution. This data set is unique in our study because the reconstructed temperature gets gradually smoother as one goes back in time. This is observed from the time series itself in Figure 8a. From the supplementary material of the Marcott et al. (2013) paper it is clear that the proxy records covering the most recent time also in general exhibit the best temporal resolution. The reconstructed temperature data for the past 1500 years therefore represent high-frequency variability in a more realistic way than the remaining part of the reconstruction.

4.3 The GRIP ice core

The European multinational research project "Greenland Ice Core Project" (GRIP) completed drilling a 3028 m deep ice core from central Greenland in 1992 (Dansgaard et al., 1993). Two GRIP data sets are used in this study, one with high temporal resolution covering 0 – 91 kyr BP (Ditlevsen et al., 1996), and one with lower temporal resolution covering 0 – 250 kyr BP (Greenland Ice-Core Project (GRIP) Members, 1993; Johnsen et al., 1997). The high-resolution data set was provided by Peter Ditlevsen at the Centre for Ice and Climate, Niels Bohr Institute, University of Copenhagen, personal communication. Both data sets are used to estimate the scaling exponents, but the results shown in Sect. 5 are for the high resolution time series. Both temperature reconstructions are based on $\delta^{18}\mathrm{O}$.

4.4 The EPICA ice core

The European Project for Ice Coring in Antarctica (EPICA) drilled two deep ice cores in Antarctica between 1996 and 2006. Here we focus on the core from dome C at the East Antarctic Plateau, covering the past 740 000 years (EPICA community members, 2004; Jouzel et al., 2007). The temperature reconstruction is based on $\delta\mathrm{D}$.

5 Results

5.1 Results for late Holocene multiproxy reconstructions

Three approaches are used to detect a scale break from the spectra of the seven multiproxy temperature reconstructions. The first is to assume a scale break at exactly 100 years, and then estimate β for long and short time scales, and determine the uncertainties for each estimate. By this approach we demonstrate that scale breaks may occur by chance from a mono-scaling model, without being statistically significant. The second approach is to use a procedure for automatic detection of a scale break from a two-scaling regimes hypothesis, and show that a wide range of time scales τ_c for the break, and a wide range of scaling exponents β_1, β_2, arise by applying the procedure to a Monte Carlo ensemble of monoscaling time series. We also employ the structure function approach, and show that the enhanced power at large scales is associated with an oscillation with characteristic scale around 500 yr.

Fig. 4 illustrates the procedure and results for the Moberg temperature reconstruction, using the first approach. The scaling exponent β is estimated from the standard periodogram of the reconstructed data, for time scales shorter than 10^2 yr ($\beta_{1,\mathrm{data}}$) and for time scales longer than 10^2 yr ($\beta_{2,\mathrm{data}}$), as shown in Fig. 4b. A Monte Carlo (MC) ensemble of synthetic fGn's with 2000 members is then constructed with $\beta_{1,\mathrm{data}}$, and from the spectra (Fig. 4c), the same estimation technique is used to estimate $\beta_{1,\mathrm{MC}}$ and $\beta_{2,\mathrm{MC}}$ for each realization. From the distribution of the estimated $\beta_{1,\mathrm{MC}}$ and $\beta_{2,\mathrm{MC}}$, the 95% confidence ranges are computed. Fig. 4d shows the mean and 95% confidence range for $\beta_{2,\mathrm{MC}}$. Since the blue line ($\beta_{2,\mathrm{data}}$) is within the confidence range for a MC ensemble of fGn's with β=0.8, the single-scaling regime hypothesis cannot be rejected. Results for all seven reconstructions are shown in Table 1.

For the Esper at al. (2002) reconstruction the estimate of $\beta_{1,\mathrm{data}}$ is slightly outside the confidence range, but this is due a bias of the synthetic fBm for β slightly higher than unity (see supplementary material section 1.1 for further details). This deviation should therefore be ignored.

From the second approach we obtain for each reconstruction two values of β and a time for the scale break. The procedure is to fit two line segments with slopes β_1 and β_2 to the log-log spectrum, such that they join at $f = f_c = 1/\tau_c$. The two slopes and the transition frequency f_c are the parameters to be fitted by an ordinary least-square procedure. Results for the seven temperature reconstructions are provided in Table 2, where also the differences in β-values are included. The scale-break hypothesis of Lovejoy and Schertzer (2012) states that the the difference $\beta_2 - \beta_1$ should be around unity. This procedure has also been tested on a Monte Carlo ensemble of mono-scaling fGn's. Fig. 5 shows a histogram of the differences in estimated β_2 and β_1. The histogram shows that the scale breaks detected by this procedure in the multiproxy records are not unlikely to be detected in records with a single scaling regime, i.e., their detection does not reject the single-scaling regime hypothesis. A histogram of τ_c also shows a broad distribution, (figure not shown).

In Figure 6a we plot SFs for the cumulative sum of the Moberg multiproxy reconstruction. The SFs are straight in the log-log plot up to around 500 yr, but then there is a broad bump. By examining the record it becomes apparent that this bump is associated with an oscillation with period of order of a millennium that involves the Medieval Warm Anomaly (MWA) high and the Little Ice Age (LIA) low. The fact that this oscillation shows up in the high-order SFs indicates that its amplitude is larger than consistent with the underlying persistent noise. If we fit the SFs by straight lines up to 500 yr we obtain the scaling function in Figure 6b with slope $H_u \approx 0.87$, in very good agreement with what was found from the instrumental data. If this oscillation were a manifestation of a new scaling regime with $\beta > 1.4$ ($h > 0.2$, $H_u > 1.2$), we should expect the SFs in Figure 6a to be straight with slope q on scales $> \tau_c$, and the scaling function obtained by fitting lines to the SFs on these scales to have slope $H_u = 1$ (this is easily demonstrated by Monte Carlo simulations). What we observe, however, is a downward bend caused by the oscillation discussed above. It can be correctly argued that this bend is not statistically significant, since we only have one sample of it, but it demonstrates very clearly that there are cases where fluctuation measures like the periodogram or the Haar fluctuation function will suggest a new scaling regime with higher scaling exponents (see Figure 7b), while the SF method will suggest oscillations.

The issue of modelling the fluctuations on multi-century time scale as a second scaling regime or an oscillation is illustrated in Figure 7. The idea is the same as in Figure 2, but now the "trend" is modelled as a growing oscillation on the form $At\sin[\omega(t-\varphi)]$. The Moberg record and the fitted trend is shown in Figure 7a, and Figure 7b shows the Haar fluctuation of the record along with realisations of a model comprised of a linear combination of an fGn and an fBm (red curves), and realizations of a model comprised of a linear combination of an fGn and the oscillatory trend. Again, the observation is consistent with both models, but the latter exhibits smaller uncertainties at longer scales, and hence is a better statistical model. Rypdal and Rypdal (2014) attribute the oscillatory trend to a combination of volcanic and solar forcing, and show that the residual after subtracting the response to this forcing is well modelled as an fGn with h compatible with what was used for the short scales in Figure 7b. Thus, compared to a two-scaling regime model, the simpler and more accurate statistical model of the Moberg record is to model the internal variability as a persistent fGn for all scales up to the length of the record, superposed on a forced oscillatory trend.

An important feature shown by the scaling functions in Figures 1 and 6 is that the background noise in GMST is monofractal. It is also Gaussian. This means that a fractional Gaussian noise is a good model for these fluctuations, and hence that all essential information is contained in the scaling exponent and the variance. This is also true for ice core data in the Holocene, while during the last glaciation, ice core data are neither monofractal nor multifractal. In an accompanying paper from our research group (Rypdal and Rypdal, 2015) it is demonstrated that if the transitions between stadial and interstadials associated with DO events and glacial/interglacial transitions are removed from the ice-core records, the remaining fluctuations scales roughly as a $1/f$ noise ($\beta \approx 1$) on time scale

longer than a century. In other words, over the length of the Antarctic ice core record (800 kyr) the temperature variability can be described as a series of glacial/interglacial transitions, and within the glacial periods; a series of stadial/interstadial transitions, superposed on a background $1/f$ noise. This suggests that Holocene variability should also exhibit this scaling of the climate noise, and the analysis of the Moberg record we have made here does not reject that hypothesis.

5.2 Results for the full Holocene multiproxy reconstruction

The reconstruction by Marcott et al. (2013) has been analysed with the periodogram in a particular way to take into account the increasing smoothness of the record as one goes backward in time. If we compute the standard periodogram for the full time series, the resulting spectral exponent is $\beta = 2.9$. The power is artificially low at high frequencies, and this is corrected by dividing the time series into segments S_n of lengths $2^n \times 400$ yr, with $n = 0, 1, 2 \ldots, 5$, and starting with the most recent period. Hence, $S_1 = 50 - 450$ yr BP, $S_2 = 50 - 850$ yr BP, $S_3 = 50 - 1650$ yr BP, $S_4 = 50 - 3250$ yr BP, $S_5 = 50 - 6450$ yr BP, and $S_6 = 50 - 11290$ yr BP (longest possible record, shorter than $2^5 \times 400$). The periodogram was estimated for each segment, and then a new power spectrum was created using only parts of each segment assumed to be trustworthy with regard to preserved variability. All of S_1 was included, while for $S_2, \ldots, S_6$ only the low-frequency parts were included (none overlapping). By this composition, the resulting power spectrum represents the variability on all time scales more correctly. Figure 8c shows the spectra of all six segments, in addition to the corrected spectrum (blue dots, black line). The value of β=1.3 is estimated from this line. The corrected spectrum still does not represent the true scaling of the global temperature, but it is a better representation than the periodogram of the original record.

5.3 Results for ice core time series

For the time series plots, time on the horizontal axis is given in years BP (before present), where "present" is defined as 1950 AD. The spectral analysis is presented in a double-logarithmic plot. The raw periodogram is plotted in gray, while the log-binned version is marked by black points. The spectral index β is estimated from the log-binned periodogram in the region shown by the blue line. Finally, the blue, shaded area indicate the 95% confidence range estimated from an ensemble of synthetic fGn's/fBm's with β and variance estimated from the log-binned periodogram. The plot of the wavelet scalogram is included in this section only for the GRIP Holocene/past85 kyr record.

For the last glacial period, we present time series and periodograms for a time interval of ≈ 80 kyr. Spectral analysis results are also included for a combined Holocene/last glacial period time series from the GRIP ice core to illustrate that analysis of such records will be dominated by the glacial climate and suppress the characteristics of Holocene climate.

5.3.1 Results from the GRIP ice core

Fig. 9a shows the δ^{18}O time series of the Holocene part of the high-resolution GRIP ice core, and Fig. 9b the periodogram from the same time series. Fig. 9c displays the same time series as shown in (a), but with the earliest 2500 yr removed. Fig 9d shows the periodogram for the time series in (c). The rationale for removing the earliest part of the Holocene record can be seen from Fig. 9a, where one observes a decrease in δ^{18}O around 8 kyr BP. This particular decrease is often observed in paleotemperature records from the northern hemisphere, and especially in records from the North-Atlantic region. The feature is known as the 8.2 kyr event, and the temperature change was probably caused by a large pulse of freshwater into the North-Atlantic Ocean associated with the collapse of the Laurentide ice sheet (Alley and Agustsdottir, 2005). In Fig. 9b, β is estimated to be ≈ 0.3 for time scales up to 10^3 yr. No scale break is detected on centennial time scales. The low value of β is typical for local temperature data from continental sites (Blender and Fraedrich, 2003; Fraedrich and Blender, 2003). On time scales longer than a millennium we can infer a higher β, but still $\beta < 1$. Since the 8.2 kyr event might affect the scaling we also analysed the shorter record (Fig. 9c). The periodogram for this time series is essentially flat. Fig. 10 shows the Mexican hat and Morlet wavelet scalograms for the full Holocene section of the GRIP ice core. The 8.2 kyr event clearly increases the power at millennial time scales, and this event is the source of the increased power observed at that time scale in Fig. 9b. From the periodogram of the Holocene part of the low-resolution GRIP time series we estimate $\beta \approx 0.1$ (not shown in figure).

Fig. 11a displays the δ^{18}O time series for the GRIP ice core from the last glacial period, and Fig. 11b the periodogram for the same time period. In Fig. 11a the Dansgaard-Oeschger (DO) events are observable as rapid warming over decadal time scales, followed by more gradual cooling (Bond and Lotti, 1995). In Fig. 11b we find $\beta \approx 1.8$ for time scales longer than 10^2 yr and shorter than 10^4 yr. On centennial time scales and shorter, the spectrum is flatter. This means that a hypothesis of a scale break at centennial time scales is plausible under glacial climate conditions, even though such a scale break could not be identified from the Holocene time series. From the low-resolution GRIP data set we estimate $\beta \approx 1.3$ for time scales longer than centennial, and a scale break is seen at this scale (figure not shown).

Fig. 12a shows the past 85 kyr time series of the high-resolution GRIP ice core, and Fig. 12b the periodogram for the same time series. In Fig. 12b, $\beta \approx 1.6$ for time scales longer than centennial, and a scale break is visible at this scale. The periodogram in Fig. 12 is very similar to that in Fig. 11, indicating that the information from the Holocene is suppressed to a high degree in the periodogram of this time series.

5.3.2 Results from the EPICA ice core

Fig. 13a shows the Holocene time series of the EPICA ice core, and Fig. 13b the periodogram for the same time series. In Fig. 13a the Antarctic equivalent to the Northern Hemisphere Holocene climate optimum (HCO) occurred between 11500 and 9000 yr BP (Masson et al., 2000). In Fig. 13b, $\beta \approx 0$ for time scales shorter than 10^3 yr.

Fig. 14a shows the time series of the EPICA ice core from the last glacial period, and Fig. 14b the periodogram for the same time period. We observe from the time series that the fluctuations do not coincide with the DO events in the GRIP ice core with respect to timing and amplitude. Like the glacial part of the GRIP ice core the EPICA glacial time series Fig. 14a has higher fluctuation levels than the Holocene counterpart. In Fig. 14b we estimate $\beta \approx 1.6$ for 10^3 yr $< \tau < 10^4$ yr. The scale break in this figures appears at 10^3 yr.

6 Discussion and conclusions

In this paper we have examined a number of paleoclimatic temperature records to assess the feasibility of detection with confidence multiple scaling regimes in Holocene climate, and in particular a break in scaling around centennial time scales. Seven proxy/multiproxy reconstructions from the late Holocene, and one for the entire Holocene, have been selected for analysis due to high temporal resolution and coverage in time, and six reconstructions from deep ice cores sampled at Greenland and Antarctica also meet our requirements for temporal coverage and resolution.

For the seven proxy-based temperature reconstructions, our first approach was to assume a break at exactly 100 years. Obviously there are few data points available for estimation on the longer scales using this procedure and the estimated values of β_2 are within the uncertainties of a mono-scaling model for all seven reconstructions. The scale break is therefore not statistically significant. For the second approach, our systematic procedure detects a break in scaling for all reconstructions. The time scale for the break varies significantly between reconstructions and is in most cases not even located near centennial time scales. The differences $\beta_2 - \beta_1$ varies over a great range and takes on both positive and negative values. This procedure has also also been tested on a Monte Carlo ensemble of fGn's and demonstrates that we will find such apparent breaks even in data that should not have breaks.

The discussion and correction of the temperature reconstruction by Marcott et al. (2013) illustrates the potential pitfall of uncritically selecting paleoclimatic time series for scaling analysis. The time series may be an excellent temperature reconstruction for many purposes, but for scaling analysis one needs to correct for the the fact that the data are increasingly low pass filtered as one goes backward in time.

In ice-core data, a scale break at centennial time scales can only be seen in records from the last glacial period. The time series for the Holocene from the GRIP and EPICA ice cores both exhibit

weak persistent scaling (Figs. 9 and 13). The scaling exponent is estimated to $\beta \approx 0.3$ and $\beta \approx 0$ for the two ice cores respectively up to millenium time scale. No break in scaling can be observed at centennial time scale. The low value of β obtained is consistent with the scaling exponents observed over land from the paleoclimate model run presented in Blender et al. (2006). On time scales longer than millenial we do not have enough data points to make confident estimates of β. From the wavelet scalograms we argue that the increase in power seen at the longest time scales in the GRIP periodogram can be attributed to the 8.2 kyr event.

The scaling properties of the GRIP and EPICA last glacial period are significantly different from the Holocene. A scale break at centennial time scales is identified with confidence from Figs. 11 and 14. We interpret this scale break as being associated with the variability of Dansgaard-Oeschger events and teleconnections to the Southern Hemisphere, (WAIS Divide Project Members, 2015). A number of theories and models exist for the mechanism of these events, see e.g., Dokken et al. (2013). This and other studies indicate that this variability is internal and not a direct response to external forcing.

From the GRIP time series including both the Holocene period and the last glacial period we obtain a power spectrum very similar to that of the glacial climate, Fig. 12. Since the glacial climate state is the dominating state in the Quaternary, the Holocene temperature variability is strongly suppressed when time series covering 100 kyr or longer are used to estimate scaling exponents. Our analyses of Holocene records, on the other hand, show that a scale break on centennial time scales is not a universal feature, and in those cases it appears to be present, it cannot be detected with sufficient certainty.

Faced with the results we have presented here one may ask what the practical implications are. Is scaling in climatic time series a useful concept? Our perception is that a scaling law may be useful as a statistical (stochastic) model when a causal description turns out to be very complex, i.e., when the viable alternative is something like a general circulation model. Such a statistical model does not have to exhibit long-memory scaling (a more standard model is a short-memory autoregressive process), but there is strong evidence, for internal variability of surface temperature data, that an fGn is a much better model than an AR(1) process for time scales at least up to centuries (Rypdal and Rypdal, 2014). Thus, for prediction on time scales up to decades, a mono-scaling model with $\beta < 1$ is what should be used, (Lovejoy et al., 2015). More interesting, however, is whether long-memory scaling is important for prediction on century time scales and beyond, and here the issue of non-stationary scaling ($\beta > 1$) for such time scales becomes crucial. What is the proper value of β to use in such prediction efforts in a warming Holocene climate? The conclusion we draw from our results is that, unless we ignore the knowledge that the present climate state of the Earth is an interglacial, we should still use $\beta < 1$. One argument that can be raised against that conclusion is that, even though we cannot reject the hypothesis of single-regime scaling based on available Holocene data, we cannot exclude that two scaling regimes is true either. Moreover, the latter is supported

from records spanning hundreds of kyr which encompass both glacial and interglacial climate. This stalemate reflects that we are faced with a model selection problem where the outcome depends on which available knowledge we prefer to emphasise.

The multiple regime model stresses the information we have on scaling in the second-order statistics such as power spectra ($\beta > 1$) on time scales up to hundreds of kyr (Quaternary scaling), and infers that this scaling should be a guideline for prediction independent of whether the initial state is glacial or interglacial. It essentially ignores the fact that Quaternary climate is characterised by several intermittencies, (Lovejoy and Schertzer, 2012). The dominating Quaternary climate state is the glacial, and temperature proxies from the last glacials take the form of a non-Gaussian intermittent stochastic process, displayed in its full glory by the Dansgaard-Oeschger events. One type of intermittency implies that the probability density functions (PDFs) are heavy-tailed on short time scales and approach Gaussian on longer time scales. This typically happens if the signal is bursty, but without long-range correlation between bursts. In this case the high-order structure functions are not straight lines in a log-log plot. Another, and more restricted, class of intermittent processes are those that are multifractal. Here we have correlated bursts, straight structure functions, but curved scaling function. Hence structure functions of higher order than two are needed to characterise the process. For a Gaussian process the power spectral density can be inferred from the second-order structure function and hence does not convey information beyond second-order statistics. Moreover, Quaternary climate is characterised by the glacial-interglacial transitions, which adds more intermittency, and all this intermittency makes prediction based only on Quaternary scaling very difficult.

The single-regime model, on the other hand, ignores the information available on time scales beyond the Holocene, but makes use of the fact that our present climate state is an interglacial, and that second-order statistics is sufficient to describe the scaling on the time-scales that is available to us in the Holocene. As discussed above, and by Rypdal and Rypdal (2015), single-regime scaling can be rejected by data that goes way beyond the Holocene if this scaling is supposed to account for DO-events and glacial/interglacial transistions, but there is no statistically significant empirical evidence that the scaling inferred by glacial-state data is present in the interglacial climate state (Rypdal and Rypdal, 2015).

The issue discussed in this paper is an example of a more general problem concerning scaling analysis that needs to be addressed in a systematic manner. On geological time scales the Earth is an evolving system. There are cycles, but the Earth rarely repeats itself. The Eemian was similar to the Holocene, but also very different, the most striking difference being the evolution of human civilisations. Thus, the dynamics of the Earth is non-stationary in a very fundamental sense. This makes scaling analysis, and modelling of Earth processes based on such analysis, a quite problematic issue. It has little meaning to talk about a universal scaling in Earths climate since the scaling characteristic on a given range of scales up to a chosen maximal scale $\tau_{\max}$ will depend on the eon, era, period, epoch, or age the analysis is done. In other words, the result will depend on the time t around which

the time range $\tau_{\max}$ is centred. A scaling analysis of a given Earth-system variable must therefore be conditioned by two essential parameters; the range $\tau_{\max}$ of scales considered, and the positioning t of this range in time. The time series and the wavelet scalogram of the GRIP temperature series for the past $t_{\max} = 90$ kyr illustrates the issue, as shown in Figure 15. The central time parameter t is along the horisontal axis and the scale τ along the vertical. We have no data for the future, which means that the transform cannot be computed correctly above the upper white line in the figure. Likewise, the area below the lower white curve is influenced by the interpolation made due to uneven sampling of the time series. It is apparent that the scalogram is different in the first 11.5 kyr (the Holocene) from the remaining 80 kyr (the last glacial). There is generally lower power on all scales in the Holocene, and the increase in power with increasing scale as t is kept constant is lower. We can also observe from this scalogram that the longest interstadials (warm stages) associated with DO-events exhibit variability very similar to the Holocene.

The literature reveals that there is no consensus of how this issue of non-stationarity could be handled. Ice cores restrict the information we can obtain to somewhat less than $\tau_{\max} = 1$ Myr BP. This is the range of time scales considered in the work of Lovejoy and Schertzer (2012) and Huybers and Curry (2006), and the period is the Quaternary in which the Earth's climate has been in a bistable state shifting between glacials and interglacials. The methodology and interpretations are based on this choice of the parameters $(\tau_{\max}, t)$. We don't see anything wrong with that, as long as one is mindful on that this is a choice, and recognises that there are other, equally valid, choices. If the issue is understanding of the present and future climate in our present interglacial state, we don't believe this choice is useful, simply because it ignores the knowledge that the Earth at present resides in an interglacial state and probably will continue to do so as long as there is human civilisation and anthropogenic forcing on this planet. Nevertheless, the recent results by (Rypdal and Rypdal, 2015) suggests that something useful can be learnt about Holocene climate from data spanning the last glacials and interglacials, namely that periods between the abrupt climate transitions are characterised by monoscaling of $1/f$-type, very similar to the scaling of the temperature variability in our present climate.

Acknowledgements. This paper was supported by the the Norwegian Research Council (KLIMAFORSK programme) under grant no. 229754. We are grateful for discussions with Martin Rypdal and Dmitry Divine, technical assistance from Ola Løvsletten and Peter Ditlevsen for providing the high-resolution GRIP data set.

References

Alley, R. B. and Agustsdottir, A. M.: The 8k event: cause and consequences of a major Holocene abrupt climate change, Quaternary Science Reviews, 24, 1123–1149, doi:10.1016/j.quascirev.2004.12.004, 2005.

Blender, R. and Fraedrich, K.: Long time memory in global warming simulations, Geophys. Res. Lett., 30, doi:10.1029/2003GL017666, 2003.

Blender, R., Fraedrich, K., and Hunt, B.: Millennial climate variability: GCM-simulation and Greenland ice cores, Geophys. Res. Lett., 33, doi:10.1029/2005GL024919, 2006.

Bond, G. C. and Lotti, R.: Iceberg Discharges into the North Atlantic on Millennial Time Scales During the Last Glaciation, Science, 267, 1005–1010, doi:10.1126/science.267.5200.1005, 1995.

Briffa, K. R., Osborn, T. J., Schweingruber, F. H., Harris, I. C., Jones, P. D., Shiyatov, S. G., and Vaganov, E. A.: Low-frequency temperature variations from a northern tree ring density network, J. Geophys. Res.-Atmos, 106, 2929–2941, doi:10.1029/2000JD900617, 2001.

Broersen, P. M. T., de Waele, S., and Bos, R.: The accuracy of time series analysis for laser-doppler velocimetry, Selected Papers from the 10th International Symposium Lisbon, Portugal July 10?13, 2000, in: Laser Techniques for Fluid Mechanics, 2000.

Dansgaard, W., Johnsen, S. J., Clausen, H. B., Dahl-Jensen, D., Gundestrup, N. S., Hammer, C. U., Hvidberg, C. S., Steffensen, J. P., Sveinbjörnsdottir, A. E., Jouzel, J., and Bond, G.: Evidence for general instability of past climate from a 250-kyr ice-core record, Nature, 364, 218–220, doi:10.1038/364218a0, 1993.

Ditlevsen, P. D., Svensmark, H., and Johnsen, S.: Contrasting atmospheric and climate dynamics of the last-glacial and Holocene periods, Nature, 379, 810–812, doi:10.1038/379810a0, 1996.

Dokken, T. M., Nisancioglu, K. H., Li, C., Battisti, D. S., and Kissel, C.: Dansgaard-Oeschger cycles: Interactions between ocean and sea ice intrinsic to the Nordic seas, Paleoceanography, 28, 491–502, doi:10.1002/palo.20042, 2013.

Efstathiou, M. N., Tzanis, C., Cracknell, A. P., and Varotsos, C. A.: New features of land and sea surface temperature anomalies, Int. J. Remote Sensing, 32, 3231–3238, doi:10.1080/01431161.2010.541504, 2011.

EPICA community members: Eight glacial cycles from an Antarctic ice core, Nature, 429, 623–628, doi:10.1038/nature02599, 2004.

Esper, J., Cook, E. R., and Schweingruber, F. H.: Low-Frequency Signals in Long Tree-Ring Chronologies for Reconstructing Past Temperature Variability, Science, 295, 2250–2253, doi:10.1126/science.1066208, 2002.

Fraedrich, K. and Blender, R.: Scaling of Atmosphere and Ocean Temperature Correlations in Observations and Climate Models, Phys. Rev. Lett., 90, 108 501, doi:10.1103/PhysRevLett.90.108501, 2003.

Greenland Ice-Core Project (GRIP) Members: Climate instability during the last interglacial period recorded in the GRIP ice core, Nature, 364, 203–207, doi:10.1038/364203a0, 1993.

Huang, S.: Merging information from different resources for new insights into climate change in the past and future, Geophys. Res. Lett., 31, doi:10.1029/2004GL019781, 2004.

Huybers, P. and Curry, W.: Links between annual, Milankovitch and continuum temperature variability, Nature, 441, doi:10.1038/nature04745, 2006.

Johnsen, S. J., Clausen, H. B., Dansgaard, W., Gundestrup, N. S., Hammer, C. U., Andersden, U., Andersen, K. K., Hvidberg, C. S., Dahl-Jensen, D., Steffensen, J. P., Shoji, H., Sveinbjörnsdottir, A. E., White, J., Jouzel, J., and Fisher, D.: The 18-O record along the Greenland Ice Core Project deep ice core

and the problem of possible Eemian climatic instability, J. Geophys. Res.-Oceans, 102, 26, 397–26, 410, doi:10.1029/97JC00167, 1997.

Jones, P. D., Briffa, K. R., Barnett, T. P., and Tett, S. F. B.: High-resolution palaeoclimatic records for the last millennium: interpretation, integration and comparison with General Circulation Model control-run temperatures, The Holocene, 8, 455–471, doi:10.1191/095968398667194956, 1998.

Jouzel, J., Masson-Delmotte, V., Cattani, O., Dreyfus, G., Falourd, S., Hoffmann, G., Minster, B., Nouet, J., Barnola, J. M., Chappellaz, J., Fischer, H., Gallet, J. C., Johnsen, S., Leuenberger, M., Loulergue, L., Luethi, D., Oerter, H., Parrenin, F., Raisbeck, G., Raynaud, D., Schilt, A., Schwander, J., Selmo, E., Souchez, R., Spahni, R., Stauffer, B., Steffensen, J. P., Stenni, B., Stocker, T. F., Tison, J. L., Werner, M., and Wolff, E. W.: Orbital and Millennial Antarctic Climate Variability over the Past 800,000 Years, Science, 317, 793–796, doi:10.1126/science.1141038, 2007.

Kantelhardt, J. W.: Fractal and Multifractal Time Series, in: Mathematics of Complexity and Dynamical Systems, edited by Meyers, R. A., pp. 463–487, Springer New York, doi:10.1007/978-1-4614-1806-1_30, 2011.

Koscielny-Bunde, A. B., Havlin, S., and Goldreich, Y.: Analysis of daily temperature fluctuations, Physica A, 231, 393–396, doi:10.1016/0378-4371(96)00187-2, 1996.

Laepple, T. and Huybers, P.: Ocean surface temperature variability: Large model-data differences at decadal and longer periods, P. Natl. A. Sci., 111, 16 682–16 687, doi:10.1073/pnas.1412077111, 2014.

Lomb, N. R.: Least-squares frequency analysis of unequally spaced data, Astrophys. Space Sci., 39, 447–462, 1976.

Lovejoy, S. and Schertzer., D.: Haar wavelets, fluctuations and structure functions: convenient choices for geophysics, Nonlinear Proc. Geoph, 19, 513–527, doi:10.5194/npg-19-513-2012, 2012.

Lovejoy, S. and Schertzer, D.: Low Frequency Weather and the Emergence of the Climate, pp. 231–254, 196, American Geophysical Union, doi:10.1029/2011GM001087, 2012.

Lovejoy, S., Schertzer, D., and Varon, D.: Do GCMs predict the climate ... or macroweather?, Earth Syst. Dynam., 4, 439–454, 2013.

Lovejoy, S., del Rio Amador, L., and Hébert, R.: The Scaling LInear Macroweather model (SLIM): using scaling to forecast global scale macroweather from months to decades, Earth System Dynamics Discussions, 6, 489–545, 2015.

Mann, M.: On long range dependence in global surface temperature series, Climatic Change, 107, 267–276, doi:10.1007/s10584-010-9998-z, 2011.

Mann, M. E., Bradley, R. S., and Hughes, M. K.: Northern Hemisphere temperatures during the past millennium: Inferences, uncertainties, and limitations, Geophys. Res. Lett., 26, 759–762, doi:10.1029/1999GL900070, 1999.

Mann, M. E., Zhang, Z., Rutherford, S., Bradley, R. S., Hughes, M. K., Shindell, D., Ammann, C., Faluvegi, G., and Ni, F.: Global Signatures and Dynamical Origins of the Little Ice Age and Medieval Climate Anomaly, Science, 326, 1256–1260, doi:10.1126/science.1177303, 2009.

Marcott, S. A. a. D. S., Clark, P. U., and Mix, A. C.: A Reconstruction of Regional and Global Temperature for the Past 11,300 Years, Science, 339, 1198–1201, doi:10.1126/science.1228026, 2013.

Masson, V., Vimeux, F., Jouzel, J., Morgan, V., Delmotte, M., Ciais, P., Hammer, C., Johnsen, S., Lipenkov, V. Y., Mosley-Thompson, E., Petit, J.-R., Steig, E. J., Stievenard, M., and Vaikmae, R.: Holocene Cli-

765 mate Variability in Antarctica Based on 11 Ice-Core Isotopic Records, Quaternary Res., 54, 348–358, doi:10.1006/qres.2000.2172, 2000.

Moberg, A., Sonechkin, D. M., Holmgren, K., Datsenko, N. M., and Karlén, W.: Highly variable Northern Hemisphere temperatures reconstructed from low-and high-resolution proxy data, Nature, 433, 613–617, doi:10.1038/nature03265, 2005.

770 Neukom, R., Gergis, J., Karoly, D. J., Wanner, H., Curran, M., Elbert, J., Gonzalez-Rouco, F., Linsley, B. K., Moy, A. D., Mundo, I., Raible, C. C., Steig, E. J., van Ommen, T., Vance, T., Villalba, R., Zinke, J., and Frank, D.: Inter-hemispheric temperature variability over the past millennium, Nature Climate Change, 4, 362–367, doi:10.1038/nclimate2174, 2014.

Østvand, L., Nilsen, T., Rypdal, K., Divine, D., and Rypdal, M.: Long-range memory in internal and forced
775 dynamics of millenium-long climate model simulations, Earth Sys. Dyn., 5, 295–308, doi:10.5194/esd-5-295-2014, 2014.

Pelletier, J. D.: The power spectral density of atmosphereic temperature from time scales of 10^{-2} to 10^{6} yr., Earth and Planetary Science Letters, 158, 157–164, doi:10.1016/S0012-821X(98)00051-X, 1998.

Rehfeld, K., Marwan, N., Heitzig, J., and Kurths, J.: Comparison of correlation analysis techniques for irregu-
780 larly sampled time series, Nonlinear Processes in Geophysics, 18, 389–404, doi:10.5194/npg-18-389-2011, 2011.

Roe, G. H. and Steig, E. J.: Characterization of Millennial-Scale Climate Variability, J. Climate, 17, 1929–1944, doi:10.1175/1520-0442(2004)017<1929:COMCV>2.0.CO;2, 2004.

Rybski, D., Bunde, A., Havlin, S., and von Storch, H.: Long-term persistence in climate and the detection
785 problem, Geophys. Res. Lett., 33, doi:10.1029/2005GL025591, 2006.

Rypdal, K., Østvand, L., and Rypdal, M.: Long-range memory in Earth's surface temperature on time scales from months to centuries, J. Geophys. Res., 118, 7046–7062, doi:10.1002/jgrd.50399, 2013.

Rypdal, M. and Rypdal, K.: Testing Hypotheses about Sun-Climate Complexity Linking, Phys. Rev. Lett., 104, 128 501–4, doi:10.1103/PhysRevLett.104.128501, 2010.

790 Rypdal, M. and Rypdal, K.: Long-memory effects in linear-response models of Earth's temperature and implications for future global warming, J. Climate, 27, doi:10.1175/JCLI-D-13-00296.1, 2014.

Rypdal, M. and Rypdal, K.: Late Quaternary temperature variability described as abrupt transitions on a $1/f$ noise background, submitted to Earth Syst.Dyn. Discuss., 2015.

Scargle, J. D.: Studies in astronomical time series analysis. ll. Statistical aspects of spectral analysis of unevenly
795 sampled data, Astrophys. J., 263, 835–853, doi:10.1086/160554, 1982.

Torrence, C. and Compo, G. P.: A Practical Guide to Wavelet Analysis, Bull. Amer. Met. Soc., 79, 61–78, doi:10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2, 1998.

WAIS Divide Project Members: Precise interpolar phasing of abrupt climate change during the last ice age, Nature, 520, 661–665, doi:10.1038/nature14401, 2015.

Table 1. Results using approach 1 for multiproxy temperature reconstructions

Reconstruction	Time period	$\beta_{1,\mathrm{data}}$	conf. range for $\beta_{1,\mathrm{MC}}$	$\beta_{2,\mathrm{data}}$	conf. range for $\beta_{2,\mathrm{MC}}$
Jones et al. 1998	1000 - 1850	0.5	(0.4, 0.7)	1.2	(-0.8, 1.7)
Briffa et al. 2001	1402 - 1850	0.6	(0.4, 0.8)	2.9	(-2.0, 3.0)
Esper et al. 2002	831 - 1850	1.3	(0.8, 1.2)	1.2	(0.2, 3.3)
Huang 2004	1500 - 1850	0.7	(0.6, 1.0)	2.3	(-4.4, 6.0)
Moberg et al. 2005	0 - 1850	0.8	(0.6, 1.0)	1.2	(0.0, 1.5)
Mann et al. 2008	500 - 1850	2.5	(1.9, 2.6)	1.6	(1.5, 3.1)
Neukom et al. 2014	1000 - 1850	0.6	(0.4, 0.8)	1.3	(-0.8, 1.9)

Table 2. Results using approach 2 for multiproxy temperature reconstructions

Data set	$\beta_{1,\mathrm{data}}$	$\beta_{2,\mathrm{data}}$	τ_c (yr)	$\beta_2 - \beta_1$
Jones et al. 1998	0.5	0.9	38	0.4
Briffa et al. 2001	0.9	0.2	22	-0.7
Esper et al. 2002	1.4	1.0	38	-0.4
Huang 2004	0.8	2.2	94	1.4
Moberg et al. 2005	0.7	2.6	353	1.9
Mann et al. 2008	3.1	0.9	47	-2.2
Neukom et al. 2014	0.5	0.8	9	0.3

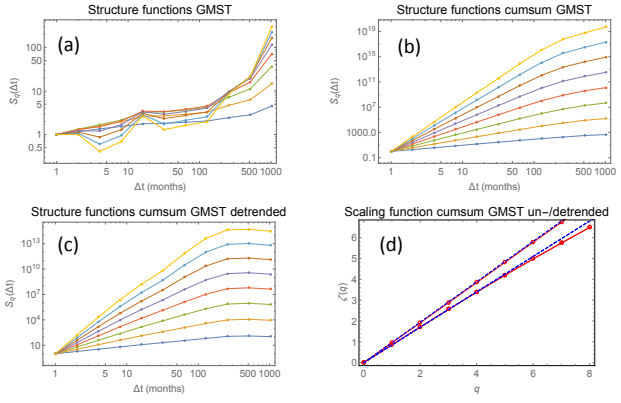

Figure 1. (a): Structure function estimates (empirical moments) $S_q(\Delta t) = (N-\Delta t)^{-1} \sum_{i=1}^{N-\Delta t} |T(t_i+\Delta t) - T(t_i)|^q$ for the GMST (HadCRUT3) monthly record 1880-2010; $T(t_i);\ i = 1, \ldots, N$. (b): Structure function for the cumulative sum $y_{t_i} = \sum_{j=1}^{i} T(t_j)$. (c): Structure function for the cumulative sum of the quadratically detrended GMST. (d): Scaling functions for the undetrended cumulative sum (upper line) and the detrended cumulative sum (lower line).

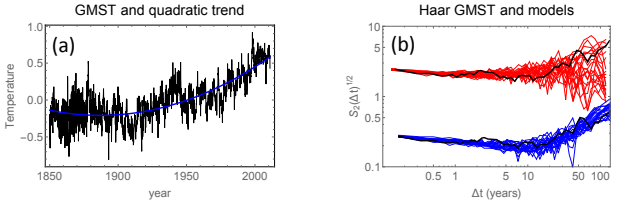

Figure 2. (a): The instrumental global mean surface temperature (GMST) 1850 – 2010 (black). A second order polynomial least-square fit to the GMST record (blue). (b): Black curves are the Haar fluctuation function of the GMST, the upper is multiplied by 10. The red curves are Haar fluctuation functions of 20 realisations of a model comprised of a linear combination of an fGn with $h = -0.1$ and an fBm with $h = 0.3$. The blue curves are the same of a model comprised of a linear combination of an fGn with $h = -0.1$ and the second-order polynomial trend.

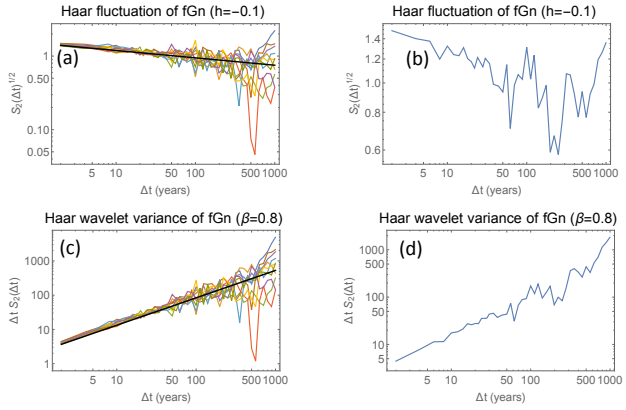

Figure 3. (a) The Haar fluctuation of 20 realisations of an fGn with h=-0.1, corresponding to $\beta = 0.8$. (b) The Haar-fluctuation of one realisation in the ensemble. (c) The Haar wavelet variance of the same 20 realisations as in panel (a). (d) The Haar wavelet variance of the same realisation as in panel (b).

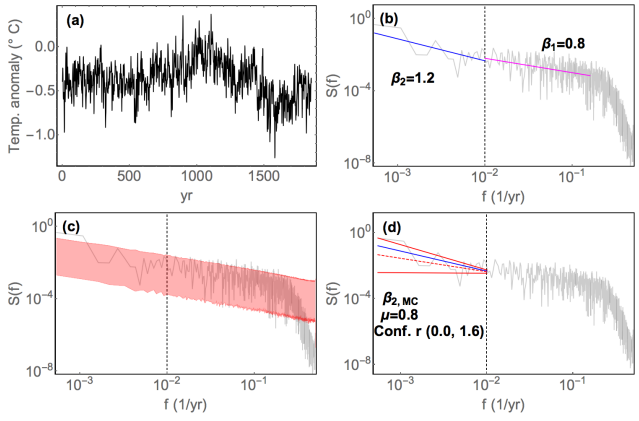

Figure 4. a) The Moberg et al. (2005) reconstructed temperature for the Northern hemisphere. (b) Estimated values of β_1 and β_2. (c) 95% confidence range for periodograms in Monte Carlo study. (d) 95% confidence range for estimates of β_2.

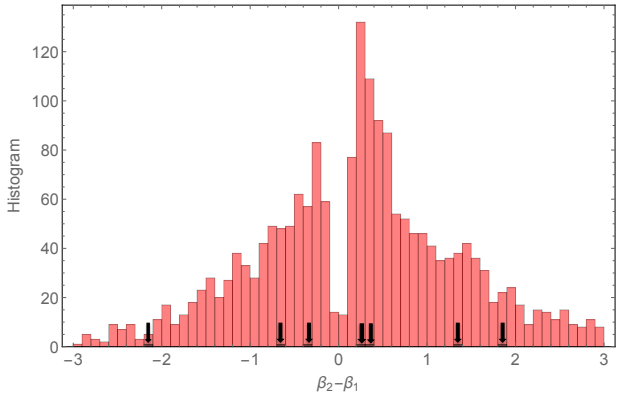

Figure 5. Differences in β_2 and β_1 for a Monte Carlo ensemble with 2000 members of synthetic LRM processes with β=0.7. The black arrows indicate the differences from the multiproxy reconstructions.

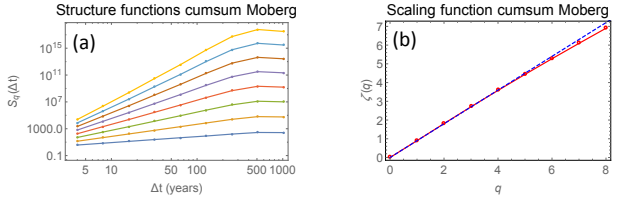

Figure 6. (a): Structure functions for the cumulative sum of the Moberg NH reconstruction year 0–1979. (b): Scaling functions for the cumulative sum computed from straight line fits to the SFs in the scale range 1 – 500 yr.

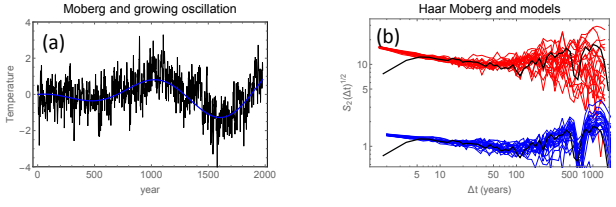

Figure 7. (a): The Moberg reconstruction (black). A fit of the function $At\sin[\omega(t-\varphi)]$ (blue). (b): Black curves are the Haar fluctuation function of the Moberg record, the upper is multiplied by 10. The red curves are Haar fluctuation functions of 20 realisations of a model comprised of a linear combination of an fGn with $h=-0.2$ and an fBm with $h=0.3$. The blue curves are the same of a model comprised of a linear combination of an fGn with $h=-0.2$ and the trend $At\sin[\omega(t-\varphi)]$.

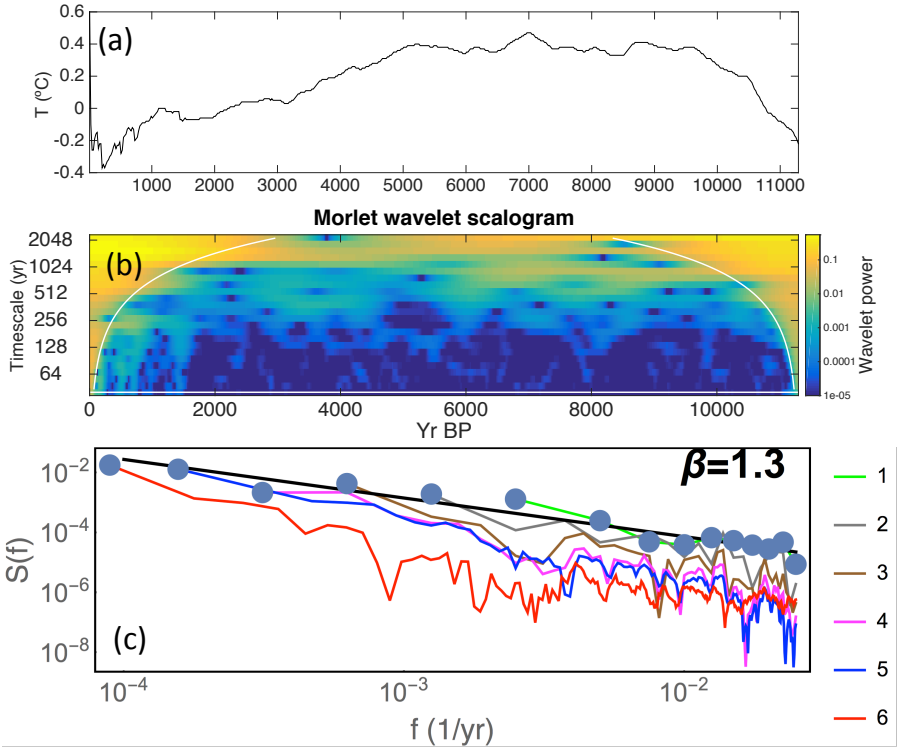

Figure 8. (a) The reconstructed time series covering the entire Holocene time period. (b) Wavelet scalogram for the same time period. (c) Section 1-6 of the reconstruction described in section 5.2, and composite spectrum (black line, blue dots). The estimated β is estimated for this line.

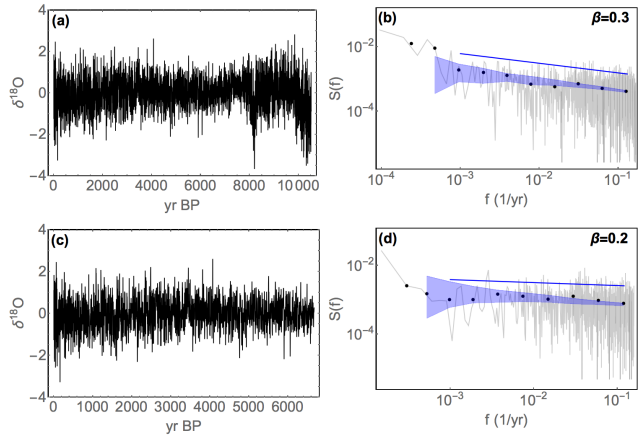

Figure 9. (a): δ^{18}O anomalies from the Holocene part of the high-resolution GRIP ice core. (b): Periodogram. The raw periodogram is shown in gray, the log-binned version by black dots. β is estimated from the log-binned periodogram in the region marked by the blue line. The confidence range is shown by the blue, shaded area, estimated from a Monte Carlo ensemble of synthetic fGns with the estimated value of β and variance from the log-binned periodogram. (c) Same figure as in (a) except the oldest section has been removed. (d) Periodogram for the time series in (c).

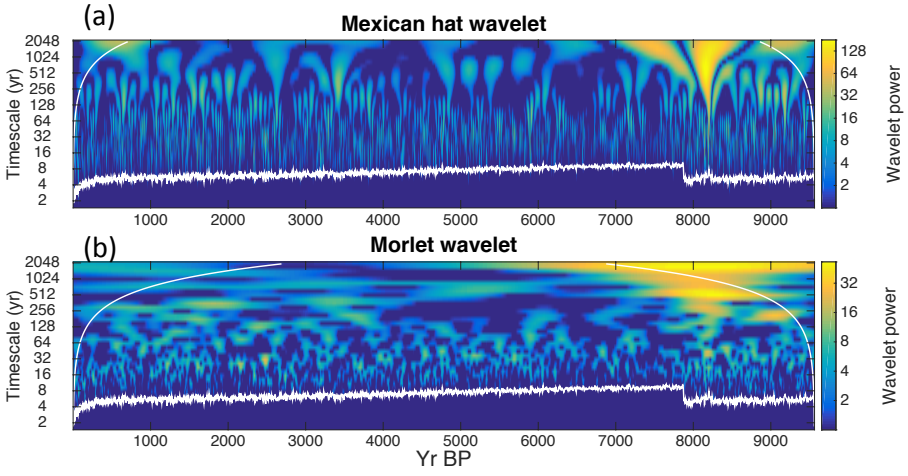

Figure 10. Top: the Mexican hat wavelet scalogram for the Holocene part of the GRIP ice core, and bottom: the Morlet wavelet scalogram for the same time series. The lower white curve in each plot denotes the varying Nyquist frequency, and the upper white curve the area affected by edge effects. Studies are restricted to the area between the two curves. The color bar to the right of the figure is used to indicate the the magnitude of the wavelet power.

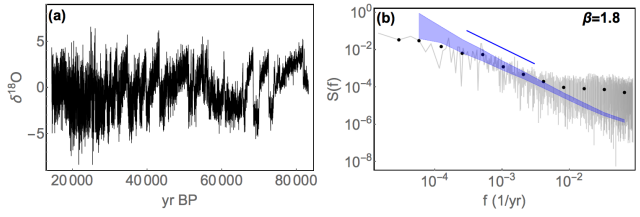

Figure 11. (a): δ^{18}O anomaly time series from the last glacial period of the high-resolution GRIP ice core. (b): Periodogram for the time series in (a).

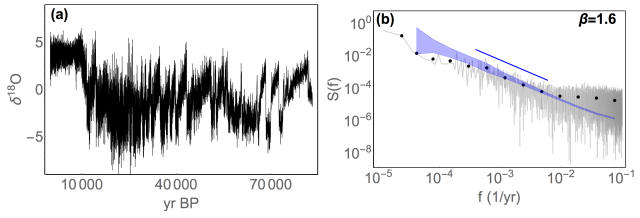

Figure 12. (a): δ^{18}O anomalies from the past 85 kyr of the high-resolution GRIP ice core. (b): Periodogram for the same time series.

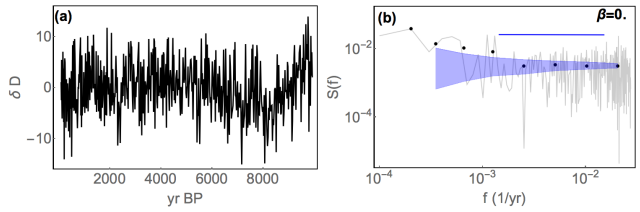

Figure 13. (a): δD anomalies from the Holocene part of the EPICA ice core. (b): Periodogram.

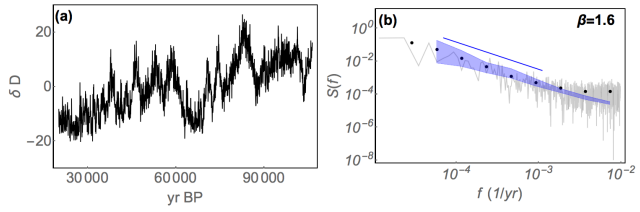

Figure 14. (a): δD anomalies from the last glacial period part of the EPICA ice core. (b): Periodogram for the time series in (a).

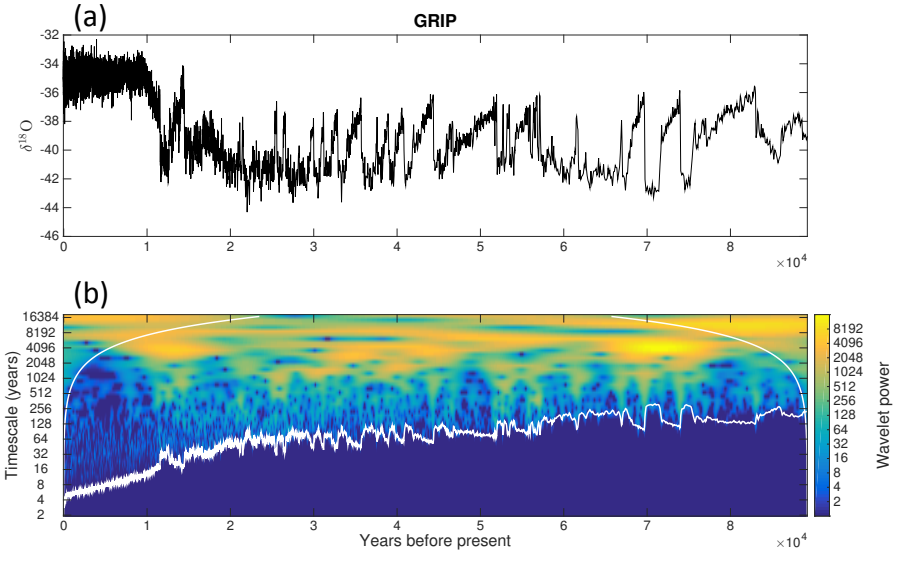

Figure 15. Upper panel: The δ^{18}O proxy time series for Greenland temperature from the GRIP ice core for the period 0 – 90 kyr BP. Lower panel: The Morlet wavelet scalogram for the signal in the upper panel.