Scaling regimes and linear / nonlinear responses of last millennium climate to volcanic and solar forcings

S. Lovejoy[1], and C. Varotsos[2]

[1]Physics, McGill University, 3600 University St., Montreal, Que., Canada

[2]Climate Research Group, Division of Environmental Physics and Meteorology, Faculty of Physics, University of Athens, University Campus Bldg. Phys. V, Athens 15784, Greece

Correspondence to: S. Lovejoy (lovejoy@physics.mcgill.ca) and C. Varotsos (covar@phys.uoa.gr)

Abstract. At scales much longer than the deterministic predictability limits (about 10 days), the statistics of the atmosphere undergoes a drastic transition, the high frequency weather acts as a random forcing on the lower frequency macroweather. In addition, up to decadal and centennial scales the equivalent radiative forcings of solar, volcanic and anthropogenic perturbations are small compared to the mean incoming solar flux. This justifies the common practice of reducing forcings to radiative equivalents (which are assumed to combine linearly), as well as the development of linear stochastic models, including for forecasting at monthly to decadal scales.

In order to clarify the validity of the linearity assumption and determine its scale range, we use last Millennium simulations, both with the simplified Zebiac- Cane (ZC) model and the NASA GISS E2-R fully coupled GCM. We systematically compare the statistical properties of solar only, volcanic only and combined solar and volcanic forcings over the range of time scales from one to 1000 years. We also compare the statistics to multiproxy temperature reconstructions. The main findings are: a) that the variability of the ZC and GCM models are too weak at centennial and longer scales, b) for longer than ≈50 years, the solar and volcanic forcings combine subadditively (nonlinearly) compounding the weakness of the response, c) the models display another nonlinear effect at shorter time scales: their sensitivities are much higher for weak forcing than for strong forcing (their intermittencies are different) and we quantify this with statistical scaling exponents.

1. Introduction

1.1 Linearity versus nonlinearity

The GCM approach to climate modeling is based on the idea that whereas weather is an initial value problem, the climate is a boundary value problem (Bryson, 1997; Pielke, 1998). This means that although the weather's sensitive dependence on initial conditions (chaos, the "butterfly effect") leads to a loss of predictability at time scales of about 10 days, nevertheless averaging over enough "weather" leads to a convergence to the model's "climate". This climate is thus the state to which averages of model outputs converge for fixed atmospheric compositions and boundary conditions (i.e. control runs).

The question then arises as to the response of the system to small changes in the boundary conditions: for example anthropogenic forcings are less than 2 W/m^2, and at least over scales of several years, solar and volcanic forcings are of similar magnitude or smaller (see e.g. Fig. 1a and the quantification in Fig. 2). These numbers are of the order of 1% of the mean solar radiative flux so that we may anticipate that the atmosphere responds fairly linearly. This is indeed that usual assumption and it justifies the reduction of potentially complex forcings to overall radiative forcings (see Meehl et al., 2004 for GCM investigations at annual scales and Hansen et al., 2005 for greenhouse gases). However, at long enough scales, linearity clearly breaks down, indeed starting with the celebrated "Daisy world" model (Watson and Lovelock, 1983), there is a whole literature that uses energy balance models to study the strongly nonlinear interactions/feedbacks between global temperatures and albedoes. There is no debate that temperature-albedo feedbacks are important at the multimillenial scales of the glacial- interglacial transitions. While some authors (e.g. Roques et al., 2014) use time scales as short as 200 years for the critical ice-albedo feedbacks, others have assumed that the temperature response to solar and volcanic forcings over the last millennium are reasonably linear (e.g. Østvand et al., 2014; Rypdal and Rypdal, 2014), while Pelletier (1998) and Fraedrich et al., (2009) assume linearity to even longer scales.

It is therefore important to establish the times scales over which linear responses are a reasonable assumption. However, clearly even over scales where typical responses to small forcings are relatively linear, the response may be nonlinear if the forcing is – volcanic or volcanic- like, i.e. if it is sufficiently "spikey" or intermittent.

1.2 Atmospheric variability: scaling regimes

Before turning our attention to models, what can we learn empirically? Certainly, at high enough frequencies (the weather regime), the atmosphere is highly nonlinear. However, at about ten days, the atmosphere undergoes a drastic transition to a lower frequency regime, and this "macroweather" regime is potentially quasi- linear in its responses. Indeed, the basic atmospheric scaling regimes were identified some time ago - primarily using spectral analysis (Lovejoy and Schertzer, 1986; Pelletier, 1998; Shackleton and Imbrie, 1990; Huybers and Curry, 2006). However, the use of real space fluctuations provided a clearer picture and a simpler interpretation. It also showed that the usual view of atmospheric variability, as a sequence of narrow scale range processes (e.g. nonlinear oscillators), has seriously neglected the main source of variability, namely the scaling "background spectrum" (Lovejoy, 2014). What was found is that for virtually all atmospheric fields, there was a transition from the behavior of the mean temperature fluctuations scaling $\langle \Delta T(\Delta t) \rangle \approx \Delta t^{H}$ with $H>0$ to a lower frequency scaling regime with $H<0$ at scales $\Delta t >\approx 10$ days; the macroweather regime. The trasntion scale of around 10 days, can be theoretically predicted on the basis of the scaling of the turbulent wind due to solar forcing (via the imposed energy rate density; see (Lovejoy and Schertzer, 2010; Lovejoy and Schertzer, 2013; Lovejoy et al., 2014). Whereas the weather is naturally identified with the high frequency $H>0$ regime and with temperature values "wandering" up and down like a drunkard's walk, the lower frequency $H<0$ regime is characterized by fluctuations tending to cancel out – effectively starting to converge. This converging regime is a low frequency type of weather, described as "macroweather" (Lovejoy, 2013; Lovejoy et al., 2014). For the GCM control runs, macroweather effectively continues to asymptotically long times; in the real world, it continues to time scales of 10-30 years (industrial) and 50-100 years (pre-industrial) after which a new $H>0$ regime is observed; it is natural to associate this new regime with the climate (see Fig. 5 of Lovejoy et al., 2013;, see also Franzke et al., 2013). Other papers analyzing macroweather scaling include Koscielny-Bunde et al., (1998); Eichner et al., (2003); Kantelhardt et al., (2006); Rybski et al., (2006); Bunde et al., (2005); Østvand et al., (2014); Rypdal and Rypdal, (2014); Fredriksen and Rypdal, (2015).

The explanation for the "macroweather" to climate transition (at scale τ_c) appears to be that over the "macroweather" time scales - where the fluctuations are "cancelling" - other, slow processes which presumably include both external climate forcings and other slow (internal) land-ice or biogeochemical processes – become stronger and stronger. At some point (τ_c) their variability dominates. A significant point where opinions diverge is the value of the global transition scale τ_c during the preindustrial Holocene; and the possibility that there are large regional variations in τ_c during the Holocene so that Greenland ice core data may not be globally representative, see Lovejoy (2015a) for a discussion.

1.3 Scaling in the numerical models

There have been several studies of the low frequency control run responses of GCMs (Vyushin et al., 2004; Zhu et al., 2006; Fraedrich et al., 2009; Lovejoy et al., 2013; Fredriksen and Rypdal, 2015) finding that they are scaling down to their lowest frequencies. This scaling is a consequence of the absence of a characteristic time scale for the long-time model convergence; it turns out that the relevant scaling exponents are very small: empirically the GCM convergence is “ultra slow” (Lovejoy et al., 2013) (section 3.4). Most earlier studies focused on the implications of the long – range statistical dependencies implicit in the scaling statistics. Unfortunately, due to this rather technical focus, the broader implications of the scaling have not been widely appreciated.

More recently, using scaling fluctuation analysis, behavior has been put into the general theoretical framework of GCM climate modeling (Lovejoy et al., 2013). From the scaling point of view, it appears that the climate arises as a consequence of slow internal climate processes combined with external forcings (especially volcanic and solar - and in the recent period - anthropogenic forcings). From the point of view of the GCMs, the low frequency (multicentennial) variability arises exclusively as a response to external forcings, although potentially - with the addition of (known or currently unknown) slow processes such as land-ice or biogeochemical processes - new internal sources of low frequency variability could be included. Ignoring the recent (industrial) period, and confining ourselves to the last millennium, the key question for GCM models is whether or not they can reproduce the climate regime where the decline of the “macroweather” fluctuations ($H<0$) is arrested and the increasing $H>0$ climate regime fluctuations begin. In a recent publication (Lovejoy et al., 2013), four GCMs simulating the last millennium were statistically analyzed and it was found that their low frequency variability (especially below (100 yrs)$^{-1}$) was somewhat weak, and this was linked to both the weakness of the solar forcings (when using sunspot-based solar reconstructions with $H>0$), and – for strong volcanic forcings - with the statistical type of the forcing ($H<0$, Lovejoy and Schertzer, 2012a; Bothe et al., 2013a,b; Zanchettin et al., 2013; see also Zanchettin et al., 2010 for the dynamics on centennial time scales).

1.4 This paper

The weakness of the responses to solar and volcanic forcings at multicentennial scales raises question a linearity question: is the response of the combined (solar plus volcanic) forcing roughly the sum of the individual responses? Additivity is often implicitly assumed when climate forcings are reduced to their equivalent radiative forcings and Mann et al., (2005) already pointed out that – at least - in the Zebiac-Cane (ZC) model discussed below that they are not additive. Here we more precisely analyze this question and quantify the degree of sub-additivity as a function of temporal scale (section 3.4). A related linear/nonlinear issue pointed out by Clement et al., (1996), is that due to the nonlinear model response, there is a high sensitivity to a small forcing and a low sensitivity to a large forcing. Systems in which strong and weak events have different statistical behaviors display stronger or weaker “clustering” and are often termed “intermittent” (from turbulence). When they are also scaling, the weak and strong events are characterized by different scaling exponents that quantify how the respective clustering changes with scale. In section 4, we investigate this quantitatively and confirm that it is particularly strong for volcanic forcing, and that for the ZC model the response (including that of a GCM), is much less intermittent, implying that the model strongly (and nonlinearly) smooths the forcing.

In this paper, we establish analysis methodologies that can address these issues and apply them to model outputs that cover the the required range of time scales: Last Millenium model outputs. Unfortunately - although we consider the NASA GISS E2-R Last Millenium simulations, there seem to be no full Last Millenium GCM simulations that have the entire suite of volcanic only, solar only and solar plus volcanic forcings and responses, therefore we have use the simplified Zebiak-Cane model outputs published by Mann et al., (2005) (and even this lacked control runs to directly quantify the internal variability).

Although the Zebiak –Cane model lacks several important mechanisms- notably for our purposes deep ocean dynamics - there are clearly sources of low frequency variability present in the model. For example, Goswami and Shukla, (1991) using 360 year control runs found multidecadal and multicentennial nonlinear variability due to the feedbacks between SST anomalies, low level convergence and atmospheric heating. In addition, in justifying his Millenium ZC simulations, (Mann et al., 2005) specifically cited model centennial scale variability as a factor motivating their study.

2. Data and analysis

2.1 Discussion

During the pre-industrial part of the last millennium, the atmospheric composition was roughly constant, and the earth’s orbital parameters varied by only a small amount. The main forcings used in GCM climate models over this period are thus solar and volcanic (in the GISS-E2-R simulations discussed below, reconstructed land use changes are also simulated but the corresponding forcings are comparatively weak and will not be discussed further). In particular, the importance of volcanic forcings was demonstrated by Minnis et al., (1993) who investigated the volcanic radiative forcing caused by the 1991 eruption of Mount Pinatubo, and found that volcanic aerosols produced a strong cooling effect. Later, Shindell et al., (2003) used a stratosphere-resolving general circulation model to examine the effect of the volcanic aerosols and solar irradiance variability on pre-industrial climate change. They found that the best agreement with historical and proxy data was obtained using both forcings. However, solar and volcanic forcings induce different responses because the stratospheric and surface influences in the solar case reinforce one another but in the volcanic case they are opposed. In addition, there are important differences in solar and volcanic temporal variabilities (including seasonality) that statistically link volcanic eruptions with the onset of ENSO events (Mann et al., 2005). Decreased solar irradiance cools the surface and stratosphere (Cracknell and Varotsos 2007, 2011; Kondratyev and Varotsos, 1995a,b). In contrast, volcanic eruptions cool the surface, but aerosol heating warms the sunlit lower stratosphere (Shindell et al., 2003; Miller et al., 2012). This leads to an increased meridional gradient in the lower stratosphere, but a reduced gradient in the tropopause region (Chandra et al., 1996; Varotsos et al., 1994, 2009).

Vyushin et al., (2004) suggested that volcanic forcings improve the low frequency variability scaling performance of atmosphere-ocean models compared to all other forcings (see however the comment by Blender and Fraedrich, (2004), which also discusses earlier papers on the field e.g. Fraedrich and Blender, (2003); Blender and Fraedrich, (2004). Weber, (2005) used a set of simulations with a climate model, driven by reconstructed forcings in order to study the Northern Hemisphere temperature response to volcanic and solar forcing, during 1000-1850. It was concluded that the response to solar forcing equilibrates at interdecadal timescales, while the response to volcanic forcing never equilibrates due to the fact that the time interval between volcanic eruptions is typically shorter than the dissipation time scale of the climate system (in fact they are scaling so that eruptions occur over all observed time scales, see below).

At the same time, Mann et al. (2005) investigated the response of El Niño to natural radiative forcing changes during 1000-1999, by employing the Zebiak–Cane model for the coupled ocean–atmosphere system in the tropical Pacific. They found that the composite feedback of the volcanic and solar radiative forcing to past changes, reproduces the fluctuations in the variability of the historic El Niño records (e.g., Efstathiou et al., 2011; Varotsos 2013).

Finally, as discussed below Lovejoy and Schertzer, (2012a) analysed the time scale dependence of several solar reconstructions Lean, (2000); Wang et al., (2005); Krivova et al., (2007); Steinhilber et al., (2009); Shapiro et al., (2011) and the two main volcanic reconstructions Crowley, (2000) and Gao et al., (2008), (referred to as “Crowley” and “Gao” in the following). The solar forcings were found to be qualitatively quite different depending on whether the reconstructions were based on sunspots or ^{10}Be isotopes from ice cores with the former increasing with time scale and the latter decreasing with time scale.

This quantitative and qualitative difference brings into question the reliability of the solar reconstructions. By comparison, the two volcanic reconstructions were both statistically similar in type; they were very strong at annual and sometimes multiannual scales but they quickly decrease with time scale $(H<0)$ explaining why they are weak at centennial and millennial scales. We re-examine these findings below.

2.2 The climate simulation of Mann et al. (2005) using the Zebiak-Cane model

Mann et al., (2005) used the Zebiak–Cane model of the tropical Pacific coupled ocean – atmosphere system (Zebiak and Cane, 1987) to produce a 100-realization ensemble for solar forcing only, volcanic forcing only and combined forcings over the last millennium. Figure 1a shows the forcings and mean responses of the model which were obtained from: ftp://ftp.ncdc.noaa.gov/pub/data/paleo/climate_forcing/mann2005/mann2005.txt. No anthropogenic effects were included. Mann et al., (2005) modeled the region between ± 30^{o} of latitude - by scaling the Crowley volcanic forcing reconstruction with a geometric factor 1.57 to take the limited range of latitudes into account. Figure 1b shows the corresponding GISS-E2-R simulation responses for three different forcings as discussed in Schmidt et al., (2013) and Lovejoy et al., (2013). Although these were averaged over the northern hemisphere land only (a somewhat different geography than the ZC simulations), one can see that the low frequencies seem similar even if the high frequencies are somewhat different. We quantify this below.

3. Methods

3.1 Comparing simulations with observations as functions of scale

The ultimate goal of weather and climate modelling (including forecasting) is to make simulations $T_{sim}(t)$ as close as possible to observations $T_{obs}(t)$. Ignoring measurement errors and simplifying the discussion by only considering a single spatial location (i.e. a single time series), the goal is to achieve simulations with $T_{sim}(t)=T_{obs}(t)$. However, this is not only very ambitious for the simulations, even when considering the observations, $T_{obs}(t)$ is often difficult to evaluate if only because data are often sparse or inadequate in various ways. However, a necessary condition for $T_{sim}(t)=T_{obs}(t)$ is the weaker statistical equality: $T_{sim}(t)\overset{d}{=}T_{obs}(t)$ where " $\overset{d}{=}$ " means equal in probability distributions (we can say that $a\overset{d}{=}b$ if $\Pr(a>s)=\Pr(b>s)$ where "Pr" indicates "probability"). Although $T_{sim}(t)\overset{d}{=}T_{obs}(t)$ is only a necessary (but not sufficient) condition for $T_{sim}(t)=T_{obs}(t)$, it is much easier to empirically verify.

Starting in the 1990s, with the advent of ensemble forecasting systems, the Rank Histogram (RH) method was proposed (Anderson, 1996) as a simple nonparametric test of $T_{sim}(t)\overset{d}{=}T_{obs}(t)$, and this has led to a large literature, including recently Bothe et al., (2013a, b). From our perspective there are two limitations of the RH method. First, it is non-parametric so that its statistical power is low. More importantly, it essentially tests the equation $T_{sim}(t)\overset{d}{=}T_{obs}(t)$ at a single unique time scale/resolution. This is troublesome since the statistics of both $T_{sim}(t)$ and $T_{obs}(t)$ series will depend on their space-time resolutions; recall that averaging in space alters the temporal statistics, e.g. $5^{o}\times5^{o}$ data are not only spatially, but also are effectively temporally smoothed with respect to $1^{o}\times1^{o}$ data. This means that even if $T_{sim}(t)$ and $T_{obs}(t)$ have nominally the same temporal resolutions they may easily have different high frequency variability. Possibly more importantly - as claimed in Lovejoy et al., (2013) and below - the main difference between $T_{sim}(t)$ and $T_{obs}(t)$ may be that the latter has more low frequency variability than the former, and this will not be captured by the RH technique which operates only at the highest frequency available. This problem is indirectly acknowledged, see for example the discussion of correlations in Marzban et al., (2011). The potential significance of the low frequencies becomes obvious when $H>0$ for the low frequency range. In this case – since the series tends to "wander", small differences in the low frequencies may translate into very large differences in RH, and this even if the high frequencies are relatively accurate.

A straightforward solution is to use the same basic idea – i.e. to change the sense of equality from deterministic to probabilistic (" = " to " $\overset{d}{=}$ ") – but to compare the statistics systematically over a range of time scales. The simplest way is to check the equality $\Delta T_{sim}(\Delta t)\overset{d}{=}\Delta T_{obs}(\Delta t)$ where ΔT is the fluctuation of the temperature over a time period Δt (see the discussion in Lovejoy and Schertzer, (2013) box 11.1). In general, knowledge of the probabilities is equivalent to knowledge of (all) the statistical moments (including the non-integer ones), and for technical reasons it turns out to be easier to check $\Delta T_{sim}(\Delta t)\overset{d}{=}\Delta T_{obs}(\Delta t)$ by considering the statistical moments.

3.2 Scaling Fluctuation Analysis

In order to isolate the variability as a function of time scale Δt, we estimated the fluctuations $\Delta F(\Delta t)$ (forcings, W/m^2), $\Delta T(\Delta t)$ (responses, *K*). Although it is traditional (and often adequate) to define fluctuations by absolute differences $\Delta T(\Delta t) = |T(t+\Delta t) - T(t)|$, for our purposes this is not sufficient. Instead we should use the absolute difference of the means from *t* to $t+\Delta t/2$ and from $t+\Delta t/2$ to $t+\Delta t$. Technically, the latter corresponds to defining fluctuations using Haar wavelets rather than "poor man's" wavelets (differences). In a scaling regime, the fluctuations vary with the time lag in a power law manner:

$$\Delta T = \varphi \Delta t^H \tag{1}$$

where φ is a controlling dynamical variable (e.g. a dynamical flux) whose mean $\langle\varphi\rangle$ is independent of the lag Δt (i.e. independent of the time scale). This means that the behaviour of the mean fluctuation is $<\Delta T> \approx \Delta t^H$ so that when $H>0$, on average fluctuations tend to grow with scale whereas when $H<0$, they tend to decrease. Note that the symbol "*H*" is in honour of Harold Edwin Hurst (Hurst, 1951). Although in the case of quasi-Gaussian statistics, it is equal to his eponymous exponent, the *H* used here is valid in the more general multifractal case and is generally different.

Fluctuations defined as differences are adequate for fluctuations increasing with scale $(H>0)$. When $H>0$, the rate at which average differences increase with time lag Δt directly reflects the increasing importance of low frequencies with respect to high frequencies. However, in physical systems the differences tend to increase even when $H<0$. This is because correlations $\langle T(t+\Delta t)T(t)\rangle$ tend to decrease with the time lag Δt and this directly implies that the mean square differences $\left(\left\langle \Delta T(\Delta t)^2\right\rangle\right)$ increase (mathematically, for a stationary process: $\left\langle \Delta T(\Delta t)^2\right\rangle = \left\langle \left(T(t+\Delta t)-T(t)\right)^2\right\rangle = 2\left(\left\langle T^2\right\rangle - \left\langle T(t+\Delta t)T(t)\right\rangle\right)$. This means that when $H<0$, differences cannot correctly characterize the fluctuations. For $H<0$ the high-frequency details dominate the differences and prevent these differences to decrease with increasing scale Δt .

The Haar fluctuation which is useful for -1<*H*<1 is particularly easy to understand since with proper "calibration" in regions where $H>0$, its value can be made to be very close to the difference fluctuation, while in regions where $H<0$, it can be made close to another simple to interpret "anomaly fluctuation". The latter is simply the temporal average of the series over a duration Δt of the series with its overall mean removed (in Lovejoy and Schertzer, 2012b this was termed a "tendency" fluctuation which is a less intuitive term). In this case, the decrease of the Haar fluctuations for increasing lag Δt characterizes how effectively averaging a (mean zero) process (the anomaly) over longer time scales reduces its variability.

Here, the calibration is affected by multiplying the raw Haar fluctuation by a factor of 2 which brings the values of the Haar fluctuations very close to both the corresponding difference and anomaly fluctuations (over time scales with $H>0$, $H<0$ respectively). This means that in regions where $H>0$, to good accuracy, the Haar fluctuations can be treated as differences whereas in regions where $H<0$ they can be treated as anomalies. While other techniques such as Detrended Fluctuation Analysis (Peng et al., 1994) perform just as well for determining exponents, they have the disadvantage that their fluctuations are not at all easy to interpret (they are the standard deviations of the residues of polynomial regressions on the running sum of the original series). Indeed, the DFA fluctuation function is typically presented without any units.

Once estimated, the variation of the fluctuations with time scale can be quantified by using their statistics; the q^{th} order structure function $S_q(\Delta t)$ is particularly convenient:

$$S_q(\Delta t) = \left\langle \Delta T(\Delta t)^q \right\rangle \tag{2}$$

where "$\langle\ \rangle$" indicates ensemble averaging (here, we average over all disjoint intervals of length Δt). Note that although q can in principle be any value, here we restrict to q>0 since divergences may occur – indeed for multifractals, are expected - for q<0). In a scaling regime, $S_q(\Delta t)$ is a power law:

$$S_q(\Delta t) = \left\langle \Delta T(\Delta t)^q \right\rangle \propto \Delta t^{\xi(q)};\ \xi(q) = qH - K(q) \tag{3}$$

where the exponent $\xi(q)$ has a linear part qH and a generally nonlinear and convex part $K(q)$ with $K(1)=0$. $K(q)$ characterizes the strong non Gaussian, multifractal variability; the "intermittency". Gaussian processes have $K(q)=0$. The root-mean-square (RMS) variation $S_2(\Delta t)^{1/2}$ (denoted simply $S(\Delta t)$ below) has the exponent $\xi(2)/2 = H - K(2)/2$. It is only when the intermittency is small ($K(q)\approx 0$) that we have $\xi(2)/2 \approx H = \xi(1)$. Note that since the spectrum is a second order statistic, we have the useful relationship for the exponent β of the power law spectra: $\beta = 1+\xi(2) = 1+2H-K(2)$ (this is a corollary of the Wiener-Khintchin theorem). Again, only when $K(2)$ is small do we have the commonly used relation $\beta \approx 1+2H$; in this case, $H>0$, $H<0$ corresponds to $\beta>1$, $\beta<1$, respectively. To get an idea of the implications of the nonlinear $K(q)$, note that a high q value characterizes the scaling of the strong events whereas a low q characterizes the scaling of the weak events (q is not restricted to integer. The scalings are different whenever the strong and weak events cluster to different degrees, the clustering in turn is precisely determined by another exponent - the codimension - which is itself is uniquely determined by $K(q)$. We return to the phenomenon of "intermittency", in section 4, it is particularly pronounced in the case of volcanic forcings.

Figure 2a shows the result of estimating the Haar fluctuations for the solar and volcanic forcings. The solar reconstruction that was used is a hybrid obtained by "splicing" the annual resolution sunspot based reconstruction (Fig. 2b, top; back to 1610, although only the more recent part was used by Mann et al. (2005) with a ^{10}Be based reconstruction (Fig. 2b, bottom) at much lower resolution (≈ 40-50 yrs). In Fig. 2a, the two rightmost curves are for two different ^{10}Be reconstructions; at any given time scale, their amplitudes differ by nearly a factor of 10 yet they both have Haar fluctuations that diminish with scale ($H \approx -0.3$). Figure 2b (top) clearly shows the qualitative difference with "wandering" ($H > 0$, sunspot based) and Fig. 2b (bottom), the cancelling ($H < 0$, ^{10}Be based) solar reconstructions (Lovejoy and Schertzer, 2012a). In the "spliced" reconstruction used here, the early ^{10}Be part (1000-1610) at low resolution was interpolated to annual resolution; the interpolation was close to linear so that we find $H \approx 1$ over the scale range 1-50 yrs, with the $H < 0$ part barely visible over the range 100-600 years (roughly the length of the ^{10}Be part of the reconstruction).

The reference lines in Fig. 2a have slopes -0.4, -0.3, 0.4 showing that both solar and volcanic forcings are fairly accurately scaling (although because of the "splicing" for the solar, only up until ≈200-300 yrs) but with exactly opposite behaviours: whereas the solar fluctuations increase with time scale, the volcanic fluctuations decrease with scale. For time scales beyond 200-300 yrs, the solar forcing is stronger than the volcanic forcing (they "cross" at roughly 0.3 W/m^2).

3.3 Linearity and nonlinearity

There is no question that - at least in the usual deterministic sense - the atmosphere is turbulent and nonlinear. Indeed, the ratio of the nonlinear to the linear terms in the dynamical equations – the Reynolds number - is typically about 10^{12}. Due to the smaller range of scales, in the numerical models it is much lower, but it is still ≈ 10^3 to 10^4. Indeed it turns out that the variability builds up scale by scale from large to small scales so that - since the dissipation scale is about 10^{-3} m - the resulting (millimetre scale) variability can be enormous; the statistics of this buildup are quite accurately modelled by multifractal cascades (see the review Lovejoy and Schertzer, 2013, especially ch. 4 for cascade analyses of data and model outputs). The cascade based Fractionally Integrated Flux model (FIF, Schertzer and Lovejoy, 1987) is a nonlinear stochastic model of the weather scale dynamics, and can be extended to provide nonlinear stochastic models of the macroweather and climate regimes (Lovejoy and Schertzer, 2013, ch. 10).

However, ever since Hasselmann, (1976), it has been proposed that sufficiently space-time averaged variables may respond linearly to sufficiently space-time averaged forcings. In the resulting (low frequency) phenomenological models, the nonlinear deterministic (high frequency) dynamics act as a source of random perturbations; the resulting stochastic model is usually taken as being linear. Such models are only justified if there is a physical scale separation between the high frequency and low frequency processes. The existence of a relevant break (at 2- 10 day scales) has been known since Panofsky and Van der Hoven, (1955) and was variously theorized as the "scale of migratory pressure systems of synoptic weather map scale" (Van der Hoven, 1957) and later as the "synoptic maximum" (Kolesnikov and Monin, 1965). From the point of view of Hasselman-type linear stochastic modelling (now often referred to as "Linear Inverse Modelling (LIM)", e.g., Penland and Sardeshmuhk, (1995); Newman et al., (2003); Sardeshmukh and Sura, (2009)), the system is regarded as a multivariate Ornstein-Uhlenbeck (OU) process. At high frequencies, an OU process is essentially the integral of a white noise (with spectrum $\omega^{-\beta_h}$ with $\beta_h = 2$), whereas at low frequencies it is a white noise, (i.e. $\omega^{-\beta_l}$ with $\beta_l = 0$). In the LIM models, these regimes correspond to the weather and macroweather, respectively. Recently Newman, (2013) has shown predictive skill for global temperature hindcasts is somewhat superior to GCM's for 1-2 year horizons.

In the more general scaling picture going back to Lovejoy and Schertzer, (1986), the transition corresponds to the lifetime of planetary structures. This interpretation was quantitatively justified in (Lovejoy and Schertzer, 2010) by using the turbulent energy rate density. The low and high frequency regimes were scaling and had spectra significantly different than those of OU processes (notably with $0.2 < \beta_l < 0.8$) with the two regimes now being referred to as "weather" and "macroweather" (Lovejoy and Schertzer, 2013). Indeed, the main difference with respect to the classical LIM is at low frequencies. Although the difference in β_l may not seem so important, the LIM value $\beta_l = 0$, (white noise) has no low frequency predictability whereas the actual values $0.2 < \beta_l < 0.8$ (depending mostly on the land or ocean location) corresponds to potentially huge predictability (the latter can diverge as β_l approaches 1). A new "ScaLIng Macroweather Model" (SLIMM) has been proposed as a set of fractional order (but still linear) stochastic differential equations with predictive skill for global mean temperatures out to at least 10 years (Lovejoy et al., 2015; Lovejoy, 2015b). However, irrespective of the exact statistical nature of the weather and macroweather regimes, a linear stochastic model may still be a valid approximation over significant ranges.

These linear stochastic models (whether LIM or SLIMM) explicitly exploit the weather/macroweather transition and may have some skill up to macroweather scales perhaps as large as decades. However, at long enough time scales, another class of phenomenological model is often used, wherein the dynamics are determined by radiative energy balances. Energy balance models focus on slower (true) climate scale processes such as sea ice – albedo feedbacks and are generally quite nonlinear, being associated with nonlinear features such as tipping points and bifurcations (Budyko, 1969). These models are typically zero or one dimensional in space (i.e. they are averaged over the whole earth or over latitude bands) and may be deterministic or stochastic (see Nicolis, 1988 for an early comparison of the two approaches). See Dijkstra, (2013) for a survey of the classical deterministic dynamical systems approach as well as the more recent stochastic "random dynamical systems" approach, (see also Ragone, et al., 2014). Although energy balance models are almost always nonlinear, there have been several suggestions that linear energy balance models are in fact valid up to millennial and even multimillennial scales.

Finally, we could mention the existence of empirical evidence of stochastic linearity between forcings and responses in the macroweather regime. Such evidence comes for example, from the apparent ability of linear regressions to "remove" the effects of volcanic, solar and anthropogenic forcings (Lean and Rind, 2008). This has perhaps been quantitatively demonstrated in the case of anthropogenic forcing where use is made of the globally, annually averaged CO_2 radiative forcings (as a linear surrogate for all anthropogenic forcings). When this radiative forcing was regressed against similarly averaged temperatures, it gave residues with amplitudes ±0.109K (Lovejoy, 2014a) which is almost exactly the same as GCM estimates of the natural variability (e.g., Laepple et al., (2008)). Notice that in this case the identification of the global temperature T_{globe} as the sum of a regression determined anthropogenic component (T_{anth}) with residues as natural variability (T_{nat}) is in fact only a confirmation of *stochastic* linearity (i.e. $T_{globe} \overset{d}{=} T_{anth} + T_{nat}$). Since presumably the actual residues would have been different if there had been no anthropogenic forcing. Indeed, when the residues were analysed using fluctuation analysis, it was only their statistics that were close to the pre-industrial multiproxy statistics.

3.4 Testing linearity: the additivity of the responses

We can now test the linearity of the model responses to solar and volcanic forcings. First consider the model responses (Fig. 3a). Compare the response to the volcanic only forcing (green) curve; with the response from the solar only forcing (black). As expected from Fig. 2a, the former is stronger than the latter up (until centennial scales) reflecting the stronger volcanic forcing. At scales $\Delta t \approx> 100$ yrs however, we see that the solar only has a stronger response, also as expected from Fig. 2a. Now consider the response to the combined volcanic and solar forcing (brown). Unsurprisingly, it is very close to the volcanic only until

$\Delta t \approx 100$ yrs; however at longer time scales, the combined response seems to decrease following the volcanic forcing curve; it seems that at these longer time scales the volcanic and solar forcings have negative feedbacks so that the combined response to solar plus volcanic forcing is actually less than for pure solar forcing, they are "subadditive".

In order to quantify this we can easily determine the expected solar and volcanic response if the two were combined additively (linearly). In the latter case, the solar and volcanic fluctuations would not interfere with each other, and since these forcings are statistically independent, the responses would also be statistically independent, the response variances would add.

A linear response means that temperature fluctuations due to only solar forcing $(\Delta T_s(\Delta t))$ and only volcanic forcing $(\Delta T_v(\Delta t))$ would be related to the temperature fluctuations of the response to the combined solar plus volcanic forcings $(\Delta T_{s,v}(\Delta t))$ as:

$$\Delta T_{s,v}(\Delta t) = \Delta T_s(\Delta t) + \Delta T_v(\Delta t) \tag{4}$$

This is true regardless of the exact definition of the fluctuation: as long as the fluctuation is defined by a linear operation on the temperature series any wavelet will do. Therefore, squaring both sides and averaging ("$\langle\ \rangle$") and assuming that the fluctuations in the solar and volcanic forcings are statistically independent of each other (i.e., $\langle \Delta T_s(\Delta t)\Delta T_v(\Delta t)\rangle = 0$), we obtain:

$$\left\langle \Delta T_{s,v}(\Delta t)^2 \right\rangle = \left\langle \Delta T_s(\Delta t)^2 \right\rangle + \left\langle \Delta T_v(\Delta t)^2 \right\rangle \tag{5}$$

The implied additive response structure function $S(\Delta t) = \left(\left\langle \Delta T_s(\Delta t)^2 \right\rangle + \left\langle \Delta T_v(\Delta t)^2 \right\rangle\right)^{1/2}$ is shown in Fig. 3b along with the ratio of the latter to the actual (nonlinear) solar plus volcanic response (top: $\left(\left\langle \Delta T_s(\Delta t)^2 \right\rangle + \left\langle \Delta T_v(\Delta t)^2 \right\rangle\right)^{1/2} \Big/ \left\langle \Delta T_{s,v}(\Delta t)^2 \right\rangle^{1/2}$). It can be seen that the ratio is fairly close to unity for time scales below about 50 yrs. However beyond 50 yrs there is indeed a strong negative feedback between the solar and volcanic forcings. This is seen more clearly in Fig. 3c which shows that at $\Delta t \approx 400$ years, that the negative feedback is strong enough to reduce the theoretical additive fluctuation amplitudes by a factor of ≈ 2 (the fall-off at the largest Δt is probably an artefact of the poor statistics at these scales). It should be noted that in addition to linearity, the latter holds assuming statistical independence (top curve in Fig. 3c) of the solar and volcanic forcing. For comparison, the bottom curve in Fig. 3c illustrates the results obtained when analyzing the series constructed by directly summing the two response series (instead of assuming statistical independence). It is clearly seen that the basic result still holds but it is a little less strong (a factor of ≈ 1.5). The reason for the difference is that the cancellation of the cross terms assumed by statistical independence is only approximately valid on single realizations, especially at the lower frequencies where the statistics are worse (even on a single realization, at any given scale - except the very longest - there are several fluctuations so that there is still some averaging).

The calculations above ignored the model's internal variability, this was considered small due to the averaging over 100 realizations of the ZC model with the same forcings: the internal is expected to largely cancel out. While it is true that a definitive answer to this requires running the model in "control mode" so as to capture only the internal variability (as was done in for the GISS model, see Fig. 4), there are nevertheless several reasons why the internal variability is almost certainly smaller than the response due to the forcings:

i) We can get a typical order of magnitude of the internal variability from the GISS model, Fig. 4; we see that for a single realization - without averaging over 100 realizations as in Fig. 3a – that the typical centennial variability is $\approx \pm 0.05$K and decreasing with a power law with exponent $\approx \xi(2) / 2 \approx -0.2$. After averaging for 100 realizations, we expect this to decrease by $(100)^{0.5} = 10$, i.e. to $\pm$ 0.005K. This is much smaller than the centennial scale variability of the ZC responses in Fig. 3a (from the graph, these are about $\approx \pm (10^{-1.2}) / 2 \approx \pm$ 0.03K.

ii) We can use the fact that a) the observed responses are upper bounds on the internal variability and b) that the internal variability must decrease with scale (otherwise the model's climate diverges rather than converges for long times. Exponents near the GISS vale $\xi(2) / 2 \approx -0.2$ are common, see e.g. Lovejoy et al., (2013). From Fig. 2, we see that the ZC solar response at $\approx$ 20 years is $\pm$ 0.03K, so this is an upper bound for the internal variability at all scales longer than $\approx$ 20 years. However, over the range $\approx$ 50-500 years (relevant for the subadditivity conclusion), the solar response variability is considerably larger than this noise value: from the graph, $\approx \pm (10^{-0.8}) / 2 \approx \pm$ 0.08K.

We conclude that it is unlikely that the internal variability is strong enough to account for the results.

In the ZC model, all forcings are input at the surface so that here the subadditivity is due to the differing seasonality, fluctuation intensities and spatial distributions of the solar and volcanic forcings. In the GISS-E2-R GCM simulations, the response to the solar forcing is too small to allow us to determine if it involves a similar solar-volcanic negative feedback (Fig. 4). In GCMs with their vertically stratified atmospheres or the real atmosphere, non additivity is perhaps not surprising given the difference between the solar and volcanic vertical heating profiles. If such negative feedbacks are substantiated in further simulations, it would enhance the credibility of the idea that current GCMs are missing critical slow (multi centennial, multi millennial) climate processes. No matter what the exact explanation, non additivity underlines the limitations of the convenient reduction of climate forcings to radiative forcing equivalents. It also indicates that at scales longer than about 50 yrs energy budget models must nonlinearly account for albedo-temperature interactions (i.e. that linear energy budget models are inadequate at these time scales, and that albedo-temperature interactions must at least be correctly parametrized).

Also shown for reference in Fig. 3a are the fluctuations for three multiproxy estimates of annual northern hemisphere temperatures (1500-1900; pre-industrial, Moberg et al., 2005; Huang, 2004; Ljungqvist, 2010, the analysis was taken from Lovejoy and Schertzer, 2012c). Although it should be borne in mind that the ZC model region (the Pacific) does not coincide with the proxy region (the northern hemisphere), the latter is the best model validation available. In addition, since we compare model and proxy fluctuation statistics as functions of time scale, the fact that the spatial regions are somewhat different is less important than if we had attempted a direct year by year comparison of model outputs with the multiproxy reconstructions.

In Fig. 3a, we see that the responses of the volcanic only and the combined volcanic and solar forcings fairly well reproduce the RMS multiproxy statistics until ≈ 50 yrs; however at longer time scales, the model fluctuations are substantially too weak – roughly 0.1 K (corresponding to ±0.05 K) and constant or falling, whereas at 400 yr scales, the RMS multiproxy temperature fluctuations are ≈ 0.25 K (±0.125) and rising. Indeed, in order to account for the ice ages, they must continue to rise until ≈ 5 K (±2.5 K) at glacial-interglacial scales of 50 – 100 kyrs, (the “glacial-interglacial window”: according to paleodata, this rise continues in a smooth, power law manner with $H>0$ until roughly 100 kyrs, see Lovejoy and Schertzer, 1986, Shackleton and Imbrie, 1990 Pelletier, 1998, Schmitt et al., 1995, Ashkenazy et al., 2003, Huybers and Curry, 2006, and Lovejoy et al., 2013).

In Fig. 4, we compare the RMS Haar fluctuations from the ZC model combined (volcanic and solar forcing) response with those from simulations from the GISS-E2-R GCM with solar only forcing and a control run (no forcings, black; see Lovejoy et al., (2013) for details; the GISS-E2-R solar forcing was the same as the spliced series used in the ZC simulations). We see that the three are remarkably close over the entire range; for the GISS model, this indicates that the solar only forcing is so small that the response is nearly the same as for the unforced (control) run. The ZC combined solar and volcanic forcing is clearly much weaker than the pre-industrial multiproxies (dashed blue, same as in Fig. 3a). The reference line with slope -0.2 shows the convergence of the control to the model climate; the shallowness of the slope (-0.2) implies that the convergence is ultra slow. For example, fluctuations from a 10 yr run control run are only reduced by a factor of $(10/3000)^{-0.2} \approx 3$ if the run is extended to 3 kyrs.

Finally, in Fig. 5, we compare the responses to the volcanic forcings for the Zebiak-Cane model and for the GISS-E2-R GCM for two different volcanic reconstructions (Gao et al., 2008), and Crowley, 2000) (the reconstruction used in the ZC simulation). For reference, we again show the combined ZC response and the preindustrial multiproxies. We see that the GISS GCM is much more sensitive to the volcanic forcing than the Zebiak-Cane model; indeed, it is too sensitive at scales $\Delta t <\approx 100$, but nevertheless becomes too weak at scales $\Delta t \approx> 200$ years. Indeed, since the volcanic forcings continue to decrease with scale, we expect the responses to keep diminishing with scale at larger Δt.

Note that the spatial regions covered by the ZC simulation, the GISS outputs and the multiproxy reconstructions are not the same. For the latter, the reason is that there is no perfectly appropriate (regionally defined) multiproxy series whereas for the GISS outputs, we reproduced the structure function analysis from a published source. Yet, the differences in the regions may not be so important since we are only making statistical comparisons. This is especially true since all the series are for planetary scale temperatures (even if they are not identical global sized regions) and in addition, we are mostly interested in the fifty year (and longer) statistics which may be quite similar.

4. Intermittency: a multifractal trace moment analysis

4.1 The Trace moment analysis technique

In the previous sections we considered the implications of linearity when climate models were forced separately with two different forcings compared with the response to the combined forcing; we showed that the ZC model was subadditive. However, linearity also constrains the relation between the fluctuations in the forcings and the responses. For example at least since the work of Clement et al., (1996), in the context of volcanic eruptions, it has been recognized that the models are typically sensitive to weak forcing events but insensitive to strong ones, i.e. they are nonlinear, and Mann et al., (2005) noticed this in their ZC simulations.

In a scaling regime, both forcings and responses will be characterized by a hierarchy of exponents (i.e. the function $\xi(q)$ in Eq. 3 or equivalently by the exponent H and the function $K(q)$), the differences in the statistics of weak and strong events are reflected in these different exponents; high order moments (large q) are dominated by large fluctuations and conversely for low order moments. The degree of convexity of $K(q)$ quantifies the degree of these nonlinear effects (indeed, how they vary over time scales Δt). Such "intermittent" behaviour was first studied in the context of turbulence (Kolmogorov, 1962; Mandelbrot, 1974).

In order to quantify this, recall that if the system is linear, the response is a convolution of the system Green's function with the forcing, in spectral terms it acts as a filter. If it is also scaling, then the filter is a power law: ω^{-H} where ω is the frequency, (mathematically, if $\tilde{T}(\omega)$ and $\tilde{F}(\omega)$ are the Fourier transforms of the response and forcing, for a scaling linear system, we have: $\tilde{T}(\omega) \propto \omega^{-H} \tilde{F}(\omega)$ such a filter corresponds to a fractional integration of order H). In terms of fluctuations this implies: $\Delta T(\Delta t) = \Delta t^{H} \Delta F(\Delta t)$ (assuming that the fluctuations are appropriately defined). Therefore, by taking q^{th} powers of both sides and ensemble averaging, we see that in linear scaling systems we have: $\xi_T(q) = qH + \xi_F(q)$ (c.f. eq. (3) with $\xi_T(q)$ and $\xi_F(q)$ the structure function exponents for the response and the forcing respectively). If $\xi_T(q)$ and $\xi_F(q)$ only differ by a term linear in q, then $K_T(q) = K_F(q)$, so that if over some regime, we find empirically $K_T(q) \neq K_F(q)$ (i.e. the intermittencies are different), then we may conclude that that the system is nonlinear (note that this result is independent of whether the linearity is deterministic or only statistical in nature).

Let us investigate the nonlinearity of the exponents by returning to Eq. (1), (2) and (3) in more detail. Up until now we have studied the statistical properties of the forcings and responses using the RMS fluctuations e.g. we have used the following equation but only for the value $q = 2$:

$$\left\langle \Delta T(\Delta t)^{q} \right\rangle \propto \left\langle \varphi_{\lambda'}^{q} \right\rangle \Delta t^{qH} = \Delta t^{\xi(q)};\ \xi(q) = qH - K(q) \qquad (6)$$

(see Eq. (1)) the exponent $K(q)$ (implicitly defined in (3)) is given explicitly by:

$$\left\langle \varphi_{\lambda'}^{q} \right\rangle = \Delta t^{K(q)};\ \ \frac{\tau_{eff}}{\Delta t} \qquad (7)$$

where τ_{eff} is the effective outer scale of the multifractal cascade process, φ gives rise to the strong variability and λ' is the cascade ratio from this outer scale to the scale of interest Δt .

If the driving flux φ was quasi-Gaussian, then $K(q)=0$, $\xi(q) = qH$ and the exponent $\xi(2) = 2H = \beta - 1$ would be sufficient for a complete characterization of the statistics. However geophysical series are often far from Gaussian, even without statistical analysis, a visual inspection (the sharp spike" of varying amplitudes, see Fig. 1a) of the volcanic series makes it obvious that it is particularly extreme in this regard. We expect - at least in this case - that the $K(q)$ term will readily be quite large (although note the constraint $K(1)=0$ and the mean of φ (the $q=1$ statistic) is independent of scale). To characterize this, note that since $K(1)=0$, we have $\xi(1)=H$ and then use the first two derivatives of $\xi(q)$ at $q=1$ to estimate the tangent (linear approximation) to $K(q)$ near the mean (C_1) and the curvature of $K(q)$ near the mean characterized by α. This gives

$$\left.\begin{aligned}(C_1)&=K'(1)=H-\xi'(1)\\ \alpha&=K''(1)/K'(1)=\xi''(1)/\left(\xi'(1)-H\right)\end{aligned}\right\} \tag{8}$$

The parameters C_1, α are particularly convenient since – thanks to a kind of multiplicative central limit theorem - there exist multifractal universality classes (Schertzer and Lovejoy, 1987). For such universal multifractal processes, the exponent function $K(q)$ can be entirely (i.e. not only near $q=1$) characterized by the same two parameters:

$$K(q)=\frac{C_1}{\alpha-1}\left(q^{\alpha}-q\right);\ 0\le\alpha\le 2 \tag{9}$$

In the universality case (9), it can be checked that the estimate in (8) (near the mean) is satisfied so that C_1, α characterize all the statistical moments (actually, (6), (7) are only valid for $q<q_c$; for $q>q_c$, the above will break down due to multifractal phase transitions; the critical q_c is typically >2, so that here we confine our analyses to $q\le 2$ and do not discuss the corresponding extreme - large q - behaviour).

A drawback of the above fluctuation method for using $\xi(q)$ to estimate $K(q)$ (6) is that if C_1 is not too big, then for the low order moments q, the exponent $\xi(q)$ may be dominated by the linear (qH) term, so that the multifractal part $(K(q))$ of the scaling is not too apparent. A simple way of directly studying $K(q)$ is to transform the original series so as to estimate the flux φ at a small scale, essentially removing the (qH) part of the exponent. It can then be degraded by temporal averaging and the scaling of the various statistical moments - the exponents $K(q)$ - can be estimated directly. To do this, we divide (1) by its ensemble average so as to estimate the normalized flux at the highest resolution by:

$$\varphi'=\frac{\varphi}{\langle\varphi\rangle}=\frac{\Delta T}{\langle\Delta T\rangle} \tag{10}$$

where the ensemble average ("$\langle\ \rangle$") is estimated by averaging over the available data (here a single series), and the fluctuations Δt are estimated at the finest resolution (here 1 yr).

4.2 Trace moment analysis of forcings, responses and multiproxies

We now test (7); for convenience, we use the symbol λ as the ratio of a convenient reference scale – here the length of the series, $\tau_{ref} = 1000$ yrs to the resolution scale Δt (for some analyses, 400 yrs was used instead, see the captions in Fig. 6). In an empirical study, the outer scale τ_{eff} is not known a priori, it must be empirically estimated; denote the scale at which the cascade starts by λ'

Starting with (7), the basic prediction of multiplicative cascades is that the normalized moments φ' (10) obey the generic multiscaling relation:

$$M(q) = \left\langle \varphi_\lambda'^{q} \right\rangle = \lambda'^{K(q)} = \left(\frac{\tau_{eff}}{\Delta t}\right)^{K(q)} = \left(\frac{\lambda}{\lambda_{eff}}\right)^{K(q)};\ \lambda' = \frac{\tau_{eff}}{\Delta t} = \frac{\lambda}{\lambda_{eff}};\ \lambda_{eff} = \frac{\tau_{ref}}{\tau_{eff}} \qquad (11)$$

We can see that τ_{eff} can readily be empirically estimated since a plot of $\mathrm{Log}_{10}M$ versus $\mathrm{Log}_{10}\lambda$ will have lines (one for each q, slope $K(q)$) converging at the outer scale $\lambda = \lambda_{eff}$ (although for a single realisation such as here, the outer scale will be poorly estimated since clearly for a single sample (series) there is no variability at the longest time scales, there is a single long-term value that generally poorly represents the ensemble mean). Figure 6a shows the results when ΔT is estimated by the absolute second difference at the finest resolution. The solar forcing (upper right) was only shown for the recent period (1600-2000) over which the higher resolution sunspot based reconstruction was used, the earlier 1000-1600 part was based on a (too) low resolution ^{10}Be "splice" as discussed above, see Fig. 2b. In the solar plot (upper left), but especially in the volcanic forcing plot (upper right), we see that the scaling is excellent over nearly the entire range (the points are nearly linear) and in addition, the lines plausibly "point" (i.e. cross) at a unique outer scale $\lambda = \lambda_{eff}$ which is not far from the length of the series, see Table 1 for estimates of the corresponding time scales. From these plots we see that the responses to the volcanic forcing "spikiness" (intermittency) are much stronger than to the corresponding responses to the weaker solar "spikiness". The model atmosphere therefore considerably dampens the intermittency, but in addition this effect is highly nonlinear so that the intermittency of the combined volcanic and solar forcing (bottom left) is actually a little less than the volcanic only intermittency (bottom right). Table 1 gives a quantitative characterization of the intermittency strength near the mean, using the C_1 parameter.

It is interesting at this stage to compare the intermittency of the ZC outputs with those of the GISS-E2-R GCM (Fig. 6b) and with multiproxy temperature reconstructions (Fig. 6c). In Fig. 6b, we see that the GISS-E2-R trace moments rapidly die off at large scales (small λ) so that the intermittency is limited to small scales to the right of the convergence point. In this Figure, we see that the lines converge at $\mathrm{Log}_{10}\lambda \approx 1.1-1.5$ corresponding to τ_{eff} in the range roughly 10–30 yrs. Since the intermittency builds up scale by scale from large scales modulating smaller scales in a hierarchical manner, and since this range of scales is small, the intermittency will be small. The partial exception is for the upper right plot which is for the GISS-E2-R response to the large Gao volcanic forcing (recall that the ZC model uses the weaker, Crowley volcanic reconstruction whose response is strongly intermittent, see Fig. 6b, the upper left plot). This result shows that contrary to the ZC model whose response is strongly intermittent (highly non Gaussian) over most of the range of time scales, the GISS-E2-R response is nearly Gaussian implying that the (highly non Gausssian) forcings are quite heavily (nonlinearly) damped.

This difference in the model responses to the forcing intermittency is already interesting, but it does not settle the question as to which model is more realistic. To attempt to answer this question, we turn to Fig. 6c which shows the trace moment analysis for six multiproxy temperature reconstructions over the same (pre-industrial) period as the GISS-E2-R model (1500-1900; unlike the ZC model, the GISS-E2-R included anthropogenic forcings so that the period since 1900 was not used in the GISS-E2-R analysis). Statistical comparisons of nine multiproxies were made in ch. 11 of Lovejoy and Schertzer, (2013), (for reasons of space, only six of these are shown in Fig. 6c) where it was found that the pre 2003 multiproxies had significantly smaller multicentennial and lower frequency variability than the more recent multiproxies used as reference in Fig. 4 and 5. However, Fig. 6c shows that the intermittencies are all quite low (with the partial exception of the Mann series, see the upper right plot). This conclusion is supported by the comparison with the red curves. These indicate the generic envelope of trace moments of quasi-Gaussian processes for $q \leq 2$ it shows how the latter converge (at large scales, small λ, to the left) to the flat $(K(q)=0)$ Gaussian limit. We see that the actual lines are only slightly outside this envelope showing that they are only marginally more variable than quasi-Gaussian processes.

The comparison of the GISS-E2-R outputs (Fig. 6b) with the multiproxies (Fig. 6c) indicates that they are both of low intermittency and are more similar to each other than to the ZC multiproxy statistics. One is therefore tempted to conclude that the GISS-E2-R model is more realistic than the ZC model with its much stronger intermittency. However this conclusion may be premature since the low multiproxy and GISS intermittencies may be due to limitations of both the multiproxies and the GISS-E2-R model.

Multicentennial and multimillenial scale ice core analyses displays significant paleotemperature intermittency ($C_1 \approx 0.05-0.1$, Schmitt et al., 1995 see the discussion in ch. 11 of Lovejoy and Schertzer, 2013) so that the multiproxies may be insufficiently intermittent.

5. Conclusions

From the point of view of GCM's, climate change is a consequence of changing boundary conditions (including composition), the latter are the climate forcings. Since forcings of interest (such as anthropogenic forcings) are typically of the order of 1% of the mean solar input the responses are plausibly linear. This justifies the reduction of the forcings to a convenient common denominator: the "equivalent radiative forcing", a concept which is useful only if different forcings add linearly, if they are "additive". An additional consequence of linearity is that the climate sensitivities are independent of whether the fluctuations in the forcings are weak or strong. Both consequences of linearity clearly have their limits. For example, at millennial and longer scales, energy balance models commonly discard linearity altogether and assume that nonlinear albedo responses to orbital changes are dominant. Similarly, at monthly and annual scales, the linearity of the climate sensitivity has been questioned in the context of sharp, strong volcanic forcings.

In view of the widespread use of the linearity assumption, it is important to quantitatively establish its limits and this can best be done using numerical climate models. A particularly convenient context is provided by the Last Millennium simulations, which (in the preindustrial epoch) are primarily driven by the physically distinct solar and volcanic forcings (forcings due to land use changes are very weak). The ideal would be to have a suite of the responses of fully coupled GCM's which include solar only, volcanic only and combined solar and volcanic forcings and control runs (for the internal variability) so that the responses could be evaluated both individually and when combined. Unfortunately, the optimal set of GCM products are the GISS E2-R millennium simulations with solar only and solar plus volcanic forcing and a control run (this suite is missing the volcanic only responses). We therefore also considered the outputs of a simplified climate model, the Zebiac-Cane (ZC) model (Mann et al., 2005) for which the full suite of external forcing response was available.

Following a previous study, we first quantified the variability of the forcings as a function of time scale by considering fluctuations. These were estimated by using the difference between the averages of the first and second halves of intervals Δt ("Haar" fluctuations). This definition was necessary in order to capture the two qualitatively different regimes, namely those in which the average fluctuations increase with time scale $(H > 0)$ and those in which they decrease with scale $(H < 0)$. Whereas the solar forcing was small at annual scales, it generally increased with scale. In comparison, the volcanic forcing was very strong at annual scales but rapidly decreased, the two becoming roughly equal at about 200 yrs. By considering the response to the combined forcing we were then able to examine and quantify their non-additivity (nonlinearity). By direct analysis (Fig. 3b, c), it was found that in the ZC model, additivity of the radiative forcings only works up until roughly 50 yr scales; at 400 yr scales, there are negative feedback interactions between the solar and volcanic forcings that reduce the combined effect by a factor of ≈ 1.5 - 2. This "subadditivivity" makes their combined effects particularly weak at these scales. Although this result seems statistically robust for the ZC Millenium simulations, until the source of the nonlinearity is pin-pointed and the results reproduced with full-blown coupled GCM's, they must be considered tentative (the conclusions would also be strengthened if ZC control runs output were available to estimate the internal variability), many more simulations with diverse forcings are needed to completely settle the issue..

In order to investigate possible nonlinear responses to sharp, strong events (such as volcanic eruptions), we used the fact that if the system is linear and scaling, then the difference between the structure function exponents $(\xi(q))$ for the forcings and responses is itself a linear function of the order of moment q (moments with large q are mostly sensitive to the rare large values, small q moments are dominated by the frequent low values). By using the trace moment analysis technique, we isolated the nonlinear part of $\xi(q)$ (i.e. the function $K(q)$) which quantifies the intermittent (multifractal, highly non-Gaussian) part of the variability (associated with the "spikiness" of the signal). Unsurprisingly we showed that the volcanic intermittency was much stronger than the solar intermittency, but that in both cases, the model responses were highly smoothed, they were practically nonintermittent (close to Gaussian) hence that the model responses to sharp, strong events were not characterized by the same sensitivity as to the more common weaker forcing events.

By examining model outputs, we have found evidence that the response of the climate system is reasonably linear with respect to the forcing up to time scales of 50 yrs at least for weak (i.e. not sharp, intermittent) events. But the sharp, intermittent events such as volcanic eruptions that occasionally disrupt the linearity at shorter time scales, become rapidly weaker at longer and longer time scales (with scaling exponent $H \approx -0.3$). In practice, linear stochastic models may therefore be valid from over most of the macroweather range, from ≈ 10 days to over 50 years. However, given their potential importance, it would be worth designing specific coupled climate model experiments in order to investigate this further.

6. Acknowledgements:

The ZC simulation outputs and corresponding solar and volcanic forcings were taken from ftp://ftp.ncdc.noaa.gov/pub/data/paleo/climate_forcing/mann2005/mann2005.txt. We thank J. Lean (solar data Fig. 2b (top), Judith.Lean@nrl.navy.mil), A. Shapiro (solar data, Fig. 2b (bottom) Alexander Shapiro, alexander.shapiro@pmodwrc.ch) and G. Schmidt (the GISS-E2-R simulation outputs, gavin.a.schmidt@nasa.gov) for graciously providing data and model outputs. The ECHAM5 based Millenium simulations analyzed in table 1 were available from: https://www.dkrz.de/Klimaforschung-en/konsortial-en/millennium-experiments-1?set_language=en. Mathematica and MatLab codes for performing the Haar fluctuation analyses are available from: http://www.physics.mcgill.ca/~gang/software/index.html. This work was unfunded, there were no conflicts of interest.

References

Anderson, J. L.: A method for producing and evaluating probabilistic forecasts from ensemble model integrations, J. Climate, 9, 1518–1530,1996.

Ashkenazy, Y., D. Baker, H. Gildor, and Havlin, S.: Nonlinearity and multifractality of climate change in the past 420,000 years, Geophys. Res. Lett., 30, 2146 doi: 10.1029/2003GL018099, 2003

Blender, R., and Fraedrich, K.: Comment on "Volcanic forcing improves atmosphere–ocean coupled general circulation model scaling performance" by D. Vyushin, I. Zhidkov, S. Havlin, A. Bunde, and S. Brenner, Geophys. Res. Lett., 31, L22213, doi: 10.1029/2004GL020797, 2004.

Bothe, O., Jungclaus, J. H., and Zanchettin, D.: Consistency of the multi-model CMIP5/PMIP3-past1000 ensemble, Climate of the Past, 9 (6), 2471-2487, 2013a.

Bothe, O., Jungclaus, J. H., Zanchettin, D., and Zorita, E.: Climate of the last millennium: Ensemble consistency of simulations and reconstructions, Climate of the Past, 9 (3), 1089-1110, 2013b.

Bryson, R. A.: The Paradigm of Climatology: An Essay, Bull. Amer. Meteor. Soc., 78, 450-456, 1997.

Budyko, M. I.: The effect of solar radiation varaitions on the climate of the earth, Tellus, 21, 611-619, 1969.

Bunde, A., Eichner, J. F., .Kantelhardt, J. W, and Havlin, S.: Long-term memory: a natural mechanism for the clustering of extreme events and anomalous residual times in climate records, Phys. Rev. Lett. , 94, 1-4 doi: 10.1103/PhysRevLett.94.048701, 2005.

Chandra, S., Varotsos, C., and .Flynn, L. E. The mid-latitude total ozone trends in the northern hemisphere, Geophys Res Lett., 23(5), 555-558, 1996.

Clement, A. C., Seager, R., Cane, M. A., and Zebiak, S. E.: An ocean dynamical thermostat, 2190–2196, 1996.

Cracknell, A. P., and Varotsos, C. A.: New aspects of global climate-dynamics research and remote sensing. Int. J. Remote Sens., *32*(3), 579-600, 2011.

Cracknell, A. P., & Varotsos, C. A.: The Antarctic 2006 ozone hole. Int. J. Remote Sens.,*28*(1), 1-2, 2007.

Crowley, T. J. :Causes of Climate Change Over the Past 1000 Years, Science, 289, 270 doi: 10.1126/science.289.5477.270, 2000.

Dijkstra, H.: *Nonlinear Climate Dynamics*, 357 pp., Cambridge University Press, Cambridge, 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 Sens., 32(11), 3231-3238, 2011.

Eichner, J. F., Koscielny-Bunde, E., Bunde, A., Havlin, S., and Schellnhuber, H.-J.: Power-law persistance and trends in the atmosphere: A detailed studey of long temperature records, Phys. Rev. E, 68, 046133-046131-046135 doi: 10.1103/PhysRevE.68.046133, 2003.

Fraedrich, K., Blender, R., and Zhu, X.: Continuum Climate Variability: Long-Term Memory, Scaling, and 1/f-Noise, International Journal of Modern Physics B, 23, 5403-5416, 2009.,

Franzke, J., Frank, D., Raible, C. C., Esper, J., and Brönnimann, S.: Spectral biases in tree-ring climate proxies Nature Clim. Change, 3, 360-364 doi: doi: 10.1038/Nclimate1816, 2013.

Fredriksen, H.-B., and Rypdal, K.: Scaling of Atmosphere and Ocean Temperature Correlations in Observations and Climate Models, J. Climate doi: doi.org/10.1175/JCLI-D-15-0457.1, 2015.

Gao, C. G., Robock, A., and Ammann, C.:, Volcanic forcing of climate over the past 1500 years: and improved ice core-based index for climate models, J. Geophys. Res., 113, D23111 doi: 10.1029/2008JD010239, 2008.

Goswami, B. N., and Shukla, J.: Aperiodic Variability in the Cane—Zebiak Model, J. of Climate, 6, 628-638, 1991.

Hansen, J., Sato, M. K. I., Ruedy, R., Nazarenko, L., Lacis, A., Schmidt, G. A., and Bell, N.:, Efficacy of climate forcings, J. Geophys. Res., 110, D18104 doi:10.1029/2005JD005776, 2005.

Hasselmann, K.: Stochastic Climate models, part I: Theory, Tellus, 28, 473-485, 1976

Huang, S.: Merging Information from Different Resources for New Insights into Climate Change in the Past and Future, Geophys.Res, Lett., 31, L13205 doi:10.1029/2004 GL019781, 2004.

Hurst, H. E.: Long-term storage capacity of reservoirs, Trans. Amer. Soc. Civil Eng., 116, 770-808, 1951.

Huybers, P., and Curry, W.: Links between annual, Milankovitch and continuum temperature variability, Nature, 441, 329-332 doi:10.1038/nature04745, 2006.

Kantelhardt, J. W., Koscielny-Bunde, E., Rybski, D., Braun, P., Bunde, A., and Havlin, S.: Long-term persistence and multifractality of precipitation and river runoff record, J. Geophys. Res., 111 doi: doi:10.1029/2005JD005881, 2006.

Kolesnikov, V. N., and Monin, A. S.: Spectra of meteorological field fluctuations, Izvestiya, Atmospheric and Oceanic Physics, 1, 653-669, 1965.,

Kolmogorov, A. N.: A refinement of previous hypotheses concerning the local structure of turbulence in viscous incompressible fluid at high Raynolds number, Journal of Fluid Mechanics, 83, 349, 1962,

Kondratyev, K. Y., and Varotsos, C. A.: Volcanic eruptions and global ozone dynamics, Int. J. Remote Sens., 16 (10), 1887-1895, 1995a.

Kondratyev, K. Y., and Varotsos, C. A.: Atmospheric greenhouse - effect in the context of global climate-change NUOVO CIMENTO DELLA SOCIETA ITALIANA DI FISICA C-GEOPHYSICS AND SPACE PHYSICS 18, 123-151, 1995b

Koscielny-Bunde, E., Bunde, A., Havlin, S., Roman, H. E., Goldreich, Y., and Schellnhuber, H. J.: Indication of a universal persistence law governing atmospheric variability, Phys. Rev. Lett. , 81, 729-732, 1998.

Krivova, N. A., Balmaceda, L., and Solanski, S. K.: Reconstruction of solar total irradiance since 1700 from the surface magnetic field flux, Astron. and Astrophys, 467, 335-346 doi: 10.1051/0004-6361:20066725, 2007.

Laepple, T., Jewson, S., and Coughlin, K.:, Interannual temperature predictions using the CMIP3 multi-model ensemble mean, Geophys. Res. Lett., 35, L10701, doi:10.1029/2008GL033576, 2008.

Lean, J. L.: Evolution of the Sun's Spectral Irradiance Since the Maunder Minimum, Geophys. Res Lett., 27, 2425-2428, 2000.

Lean, J. L., and Rind, D. H.: How natural and anthropogenic influences alter global and regional surface temperatures: 1889 to 2006, Geophys. Res. Lett., 35, L18701 doi: 10.1029/2008GL034864, 2008.

Ljungqvist, F. C.: A new reconstruction of temperature variability in the extra - tropical Northern Hemisphere during the last two millennia, Geografiska Annaler: Physical Geography, *92 A*(3), 339 - 351 doi:10.1111/j .1468 - 0459.2010 .00399.x, 2010.

Lovejoy, S.: What is climate?, EOS, 94, (1), 1 January, p1-2, 2013.

Lovejoy, S.: Scaling fluctuation analysis and statistical hypothesis testing of anthropogenic warming, Climate Dyn., 42, 2339-2351 doi:10.1007/s00382-014-2128-2, 2014a.

Lovejoy, S.: A voyage through scales, a missing quadrillion and why the climate is not what ou expect, Climate Dyn., 44, 3187-3210, doi: 10.1007/s00382-014-2324-0, 2014b.,

Lovejoy, S.:, The macroweather to climate transition in the Holocene: regional and epoch to epoch variability (comments on “Are there multiple scaling regimes in Holocene temperature records?”by T. Nilsen, K. Rypdal, and H.-B. Fredriksen), Earth Syst. Dynam. Discus., 6, C1–C10, 2015a.

Lovejoy, S.: Using scaling for macroweather forecasting including the pause, Geophys. Res. Lett., 42, 7148–7155 doi:10.1002/2015GL065665, 2015b.

Lovejoy, S., and Schertzer, D.: Scale invariance in climatological temperatures and the local spectral plateau, Annales Geophysicae, 4B, 401-410, 1986.,

Lovejoy, S., and Schertzer,D.: Towards a new synthesis for atmospheric dynamics: space-time cascades, Atmos. Res., 96, 1-52 doi:10.1016/j.atmosres.2010.01.004, 2010.

Lovejoy, S., and Schertzer, D.: Stochastic and scaling climate sensitivities: solar, volcanic and orbital forcings, Geophys. Res. Lett., 39, L11702, doi:10.1029/2012GL051871, 2012a.,

Lovejoy, S., and Schertzer, D.: Low frequency weather and the emergence of the Climate, in Extreme Events and Natural Hazards: The Complexity Perspective, edited by A. S. Sharma, A. Bunde, D. N. Baker and V. P. Dimri, pp. 231-254, AGU monographs, Washington D.C., 2012b,

Lovejoy, S., and Schertzer, D.: ,Haar wavelets, fluctuations and structure functions: convenient choices for geophysics, Nonlinear Proc. Geophys., 19, 1-14, doi:10.5194/npg-19-1-2012, 2012c.

Lovejoy, S., and Schertzer, D.: The Weather and Climate: Emergent Laws and Multifractal Cascades, 496 pp., Cambridge University Press, Cambridge, 2013.

Lovejoy, S., Schertzer, D., and Varon, D.: Do GCM’s predict the climate…. or macroweather?, Earth Syst. Dynam. , 4, 1–16 doi:10.5194/esd-4-1-2013, 2013.,

Lovejoy, S., Muller, J. P., and Boisvert, J. P.: On Mars too, expect macroweather, Geophys. Res. Lett., 41, 7694-7700, doi:10.1002/2014GL061861, 2014.

Lovejoy, S., del Rio Amador, L., and Hébert, R.: The ScaLIng Macroweather Model (SLIMM): using scaling to forecast global-scale macroweather from months to Decades, Earth Syst. Dynam., 6, 1–22, http://www.earth-syst-dynam.net/6/1/2015/, doi:10.5194/esd-6-1-2015, 2015.

Mandelbrot, B. B.: Intermittent turbulence in self-similar cascades: divergence of high moments and dimension of the carrier, Journal of Fluid Mechanics, 62, 331-350, 1974.

Mann, M. E., Cane, M. A., Zebiak, S. E., and Clement, A.: Volcanic and solar forcing of the tropical pacific over the past 1000 years, J. Clim., 18, 447-456, 2005.

Marzban, C., Wang, R., Kong, F., and Leyton, S.: On the effect of correlations on rank histograms: reliability of temperature and wind speed forecasts from fine scale ensemble reforecasts, Mon. Weather Rev., 139, 295–310, doi:doi:10.1175/2010MWR3129.1, 2011.

Meehl, G. A., Washington, W. M., .Ammann, C. M.,Arblaster, J. M., Wigley, T. M. L., and Tebaldi, C.: Combinations Of Natural and Anthropogenic Forcings In Twentieth-Century Climate, J . of Clim. , 17, 3721-3727, 2004.

Miller, G. H., Geirsdóttir, Á., Zhong, Y., Larsen, D. J., Otto Bliesner, B. L., Holland, M. M., and Anderson, C.: Abrupt onset of the Little Ice Age triggered by volcanism and sustained by sea-ice/ocean feedbacks, Geophys. Res. Lett., 39, L02708 doi:10.1029/2011GL050168, 2012.

Minnis, P., Harrison, E. F., Stowe, L. L., Gibson, G. G.,Denn, F. M., Doelling, D. R., and Smith Jr, W. L.: Radiative Climate Forcing by the Mount Pinatubo Eruption, Science, 259 (5100), 1411-1415, 1993.

Moberg, A., Sonnechkin, 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(7026), 613-617, 2005.

Newman, M.: An Empirical Benchmark for Decadal Forecasts of Global Surface Temperature Anomalies, J. of Clim., 26, 5260-5269, doi:10.1175/JCLI-D-12-00590.1, 2013.

Newman, M. P., Sardeshmukh, P. D., and. Whitaker, J. S.:, A study of subseasonal predictability, Mon. Wea. Rev., 131, 1715-1732, 2003.

Nicolis, C.: Transient climatic response to increasing CO2 concentration: some dynamical scenarios, Tellus A, 40A, 50-60, doi:10.1111/j.1600-0870.1988.tb00330.x, 1988.

Østvand, L., Nilsen, T., Rypdal, K., Divine, D., and Rypdal, M.:, Long-range memory in millennium-long ESM and AOGCM experiments, Earth System Dynamics, 5, ISSN 2190-4979.s 2295 - 2308.s, doi:10.5194/esd-5-295-2014, 2014.

Panofsky, H. A., and Van der Hoven, I.: Spectra and cross-spectra of velocity components in the mesometeorlogical range, Quarterly J. of the Royal Meteorol. Soc., 81, 603-606, 1955.

Pelletier, J., D.: The power spectral density of atmospheric temperature from scales of 10^{-2} to 10^{6} yr, EPSL, 158, 157-164, 1998.

Peng, C.-K., Buldyrev, S. V., Havlin, S., Simons, M., Stanley, H. E.,and Goldberger, A. L.: Mosaic organisation of DNA nucleotides, Phys. Rev. E, 49, 1685-1689, 1994.,

Penland, C., and Sardeshmuhk, P. D.: The optimal growth of tropical sea surface temperature anomalies, J. Climate, 8, 1999-2024, 1995.

Pielke, R.: Climate prediction as an initial value problem, Bull. of the Amer. Meteor. Soc., 79, 2743-2746, 1998.

Ragone, F., Lucarini, V., and Lunkeit, F.: A new framework for climate sensitivity and prediction: a modelling perspective, Climate Dynamics, 1-13, 2014.

Roques, L., Chekroun, M. D., Cristofol, M., Soubeyrand, S., and Ghi, M.: Parameter estimation for energy balance models with memory, Proc. Roy. Soc. A, 470 20140349 doi: DOI: 10.1098/rspa.2014.0349, 2014.

Rybski, D., Bunde, A. Havlin, S., and von Storch, H.: Long-term persistance in climate and the detection problem, Geophys. Resear. Lett., 33, L06718-06711-06714, doi:10.1029/2005GL025591, 2006.

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 (14), 5240 - 5258, doi:10.1175/JCLI-D-13-00296.1, 2014.

Sardeshmukh, P. D., and Sura, P.:, Reconciling non-gaussian climate statistics with linear dynamics, J. of Climate, 22, 1193-1207, 2009.

Schertzer, D., and Lovejoy, S.: Physical modeling and Analysis of Rain and Clouds by Anisotropic Scaling of Multiplicative Processes, J Geophys Res, 92, 9693-9714, 1987.

Schmidt, G. A., et al.: Using paleo-climate model/data comparisons to constrain future projections in CMIP5, Clim. Past Discuss., *9*, 775-835, doi:10.5194/cpd-9-775-2013, 2013.

Schmitt, F., Lovejoy, S., and Schertzer, D.: Multifractal analysis of the Greenland Ice-core project climate data., Geophys. Res. Lett, 22, 1689-1692, 1995.

Shackleton, N. J., and Imbrie, J.: The δ18O spectrum of oceanic deep water over a five-decade band, Climatic Change, 16, 217-230, 1990.

Shapiro, A. I., Schmutz, W., Rozanov, E., Schoell, M., .Haberreiter, M., Shapiro, A. V., and Nyeki, S.: A new approach to long-term reconstruction of the solar irradiance leads to large historical solar forcing, Astronomy & Astrophysics, 529, A67, doi: doi.org/10.1051/0004-6361/201016173, 2011.

Shindell, D. T., Schmidt, G. A., Miller, R. I., and Mann, M. E.:Volcanic and Solar Forcing of Climate Change during the Preindustrial Era, J. Clim., 16, 4094-4107, 2003.

Steinhilber, F., Beer, J., and Frohlich, C.: Total solar irradiance during the Holocene, Geophys. Res. Lett., 36, L19704, doi:10.1029/2009GL040142, 2009.

Van der Hoven, I.:, Power spectrum of horizontal wind speed in the frequency range from 0.0007 to 900 cycles per hour, Journal of Meteorology, 14, 160-164, 1957.

Varotsos, C. A.: The global signature of the ENSO and SST-like fields. Theor. Applied Clim., 113(1-2), 197-204, 2013.

Varotsos, C., Efstathiou, M., and Tzanis, C.: Scaling behaviour of the global tropopause. Atmos Chem Phys, 9(2), 677-683, 2009.

Varotsos, C., Kalabokas, P., and Chronopoulos, G.: Association of the laminated vertical ozone structure with the lower-stratospheric circulation. J. Applied Meteorol, 33(4), 473-476, 1994.

Vyushin, D., Zhidkov, I., Havlin, S., . Bunde, A., and Brenner, S.: Volcanic forcing improves atmosphere-ocean coupled, general circulation model scaling performance. Geophy. Res. Lett., 31, L10206, doi:10.1029/2004GL019499, 2004.

Wang, Y.-M., Lean, J. L., and Sheeley, N. R. J.: Modeling the Sun's magnetic field and irradiance since 1713, Astrophys J. , *625*, 522–538, 2005.

Watson, A. J. and Lovelock, J. E.: Biological homeostasis of the global environment: the parable of Daisyworld, Tellus, 35B, 284-289, 1983.

Weber, S. L.: A timescale analysis of the Northern Hemisphere temperature response to volcanic and solar forcing, Climate of the Past, 1, 9–17, 2005.

Zanchettin, D., Rubino, A., and Jungclaus, J. H.: Intermittent multidecadal-to-centennial fluctuations dominate global temperature evolution over the last millennium, Geophys. Res. Lett., 37 (14), L14702, 2010.

Zanchettin, D., Rubino, A., Matei, D., Bothe, O., and Jungclaus, J. H.:, Multidecadal-to-centennial SST variability in the MPI-ESM simulation ensemble for the last millennium, Climate Dynamics, 40 (5-6), 1301-1318, 2013.

Zebiak, S. E., and Cane, M. A.: A Model El Niño – Southern Oscillation, Mon. Wea. Rev., 115, 2262–2278, 1987. .

Zhu, X., Fraederich, L., and Blender, R.:Variability regimes of simulated Atlantic MOC, Geophys. Res. Lett., 33, L21603, doi:10.1029/2006GL027291, 2006.

Tables:

Table 1. The scaling exponent estimates for the forcings and ZC model responses.

	Forcings		**Responses**			**Control Runs**	
	Solar	**Volcanic**	**Solar**	**Volcanic**	**Combined**	**GISS**	**ECHAM5**
H	0.40	-0.21	0.031	-0.17	-0.15	-0.26	-0.4
C_1	0.095	0.48	0.022	0.054	0.038	<0.01	<0.01
α	1.04	0.31	1.82	2.0	2.0	–	–
$\xi(2)/2$	0.33	-0.47	-0.01	-0.28	-0.23	<0.01	<0.01
β	1.66	0.06	0.98	0.44	0.54	0.47	0.2
τ_{eff}	630 yrs	300yrs	100yrs	100 yrs	250 yrs	–	–

Table 1 shows the scaling exponent estimates for the forcings and ZC model responses. For the solar (forcing and response), only the recent 400 yrs (sunspot based) series were used, for the others, the entire 1000 yrs range was used, see figure 6a. The RMS exponent was estimated from Eq. (6), (9): H was estimated from the Haar fluctuations, α , C_1 were estimated from the trace moments (Fig. 6a). Note that the external cascade scales are unreliable since they were estimated from a single realization. The control runs at the right are for the GISS-E2-R model discussed in the text and (ECHAM5) from the fully coupled COSMOS-ASOB Millenium long term simulations based on the Hamburg ECHAM5 model for 800–4000AD.

Figures and Captions:

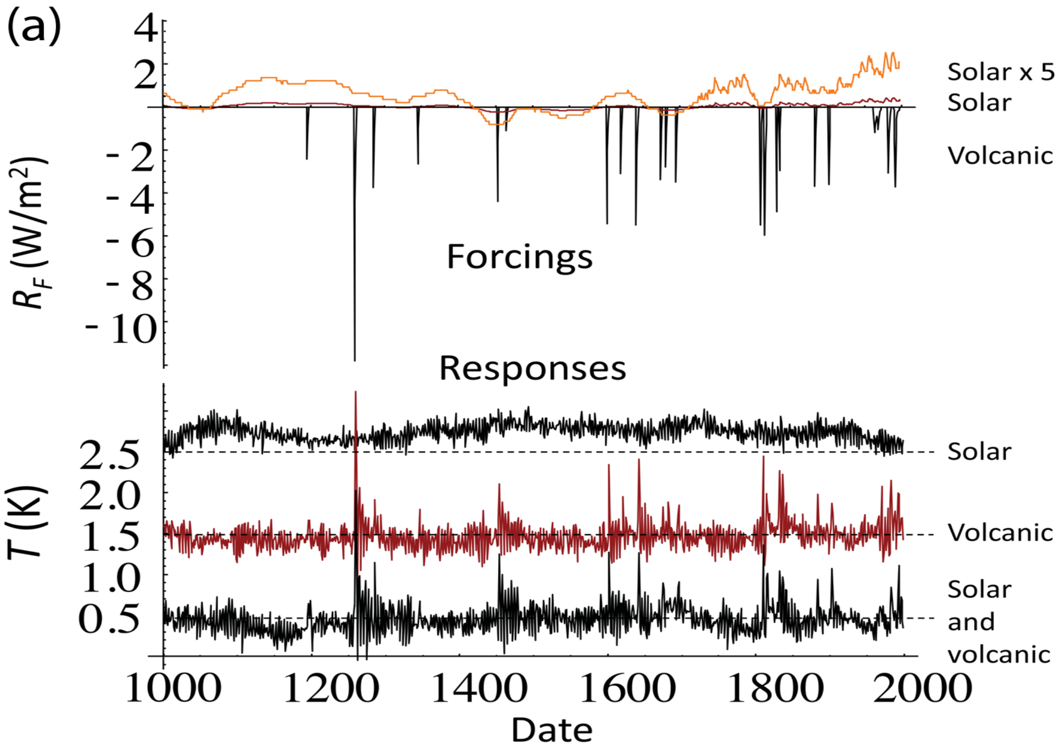

Figure 1a. *Top graph:* The radiative forcings R_F (top, W/m^2) and responses T(K) from 1000-2000 AD for the Zebiak–Cane model, from Mann et al., (2005), integrated over the entire simulation region. The forcings are reconstructed solar (brown), solar blown up by a factor 5 (orange) and volcanic (red). For the solar forcing (top series), note the higher resolution and wandering character for the recent centuries – this part is based on sunspots, not ^{10}Be.

Bottom graph: The responses are for the solar forcing only (top), volcanic forcing only (middle) and both (bottom); they have been offset in the vertical for clarity by 2.5, 1.5, 0.5K, respectively.

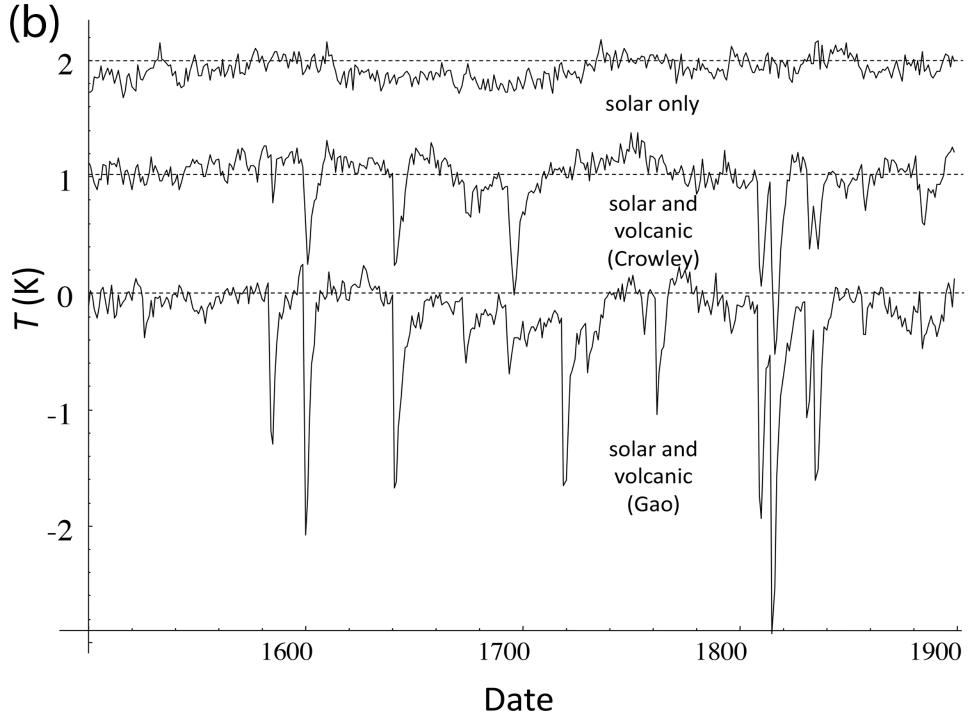

Figure 1b. GISS-ER-2 responses averaged over land, the northern hemisphere at annual resolution. The industrial part since 1900 was excluded due to the dominance of the anthropogenic forcings. The solar forcing is the same as for the ZC model, it is mostly sunspot based (since 1610). The top row is for the solar forcing only, the middle series is the response to the solar and Crowley reconstructed volcanic forcing series (i.e. the same as used in the ZC model); the bottom series uses the solar and reconstructed volcanic forcing series from Gao et al., (2008). Each series has been offset in the vertical by 1K for clarity (these are anomalies so that the absolute temperature values are unimportant).

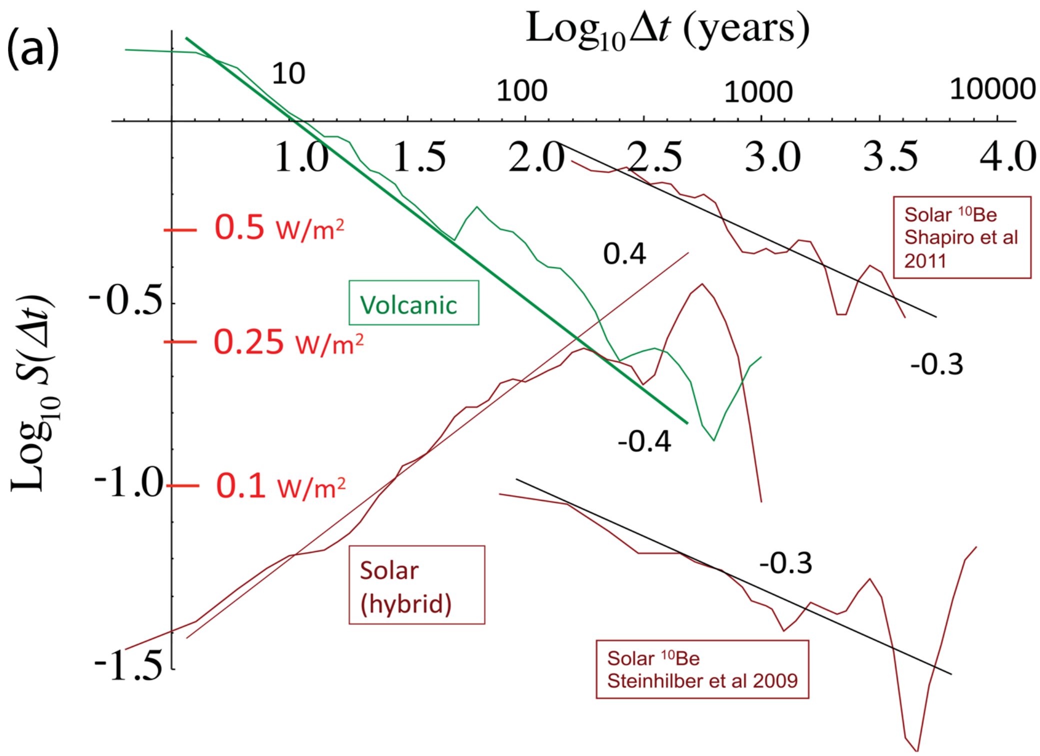

Figure 2a. The RMS Haar fluctuation $S(\Delta t)$ for the solar and volcanic reconstructions used in the ZC simulation for lags Δt from 2 to 1000 years (left). The solar is a "hybrid" obtained by "splicing" the sunspot-based reconstruction (Fig. 2b, top) with a ^{10}Be based reconstruction (Fig. 2b, bottom). The two rightmost curves are for two different ^{10}Be reconstructions (Shapiro et al., 2011; Steinhilber et al., 2009). Although at any given scale, their different assumptions lead to amplitudes differing by nearly a factor of 10, their exponents are virtually identical and the amplitudes diminish rapidly with scale.

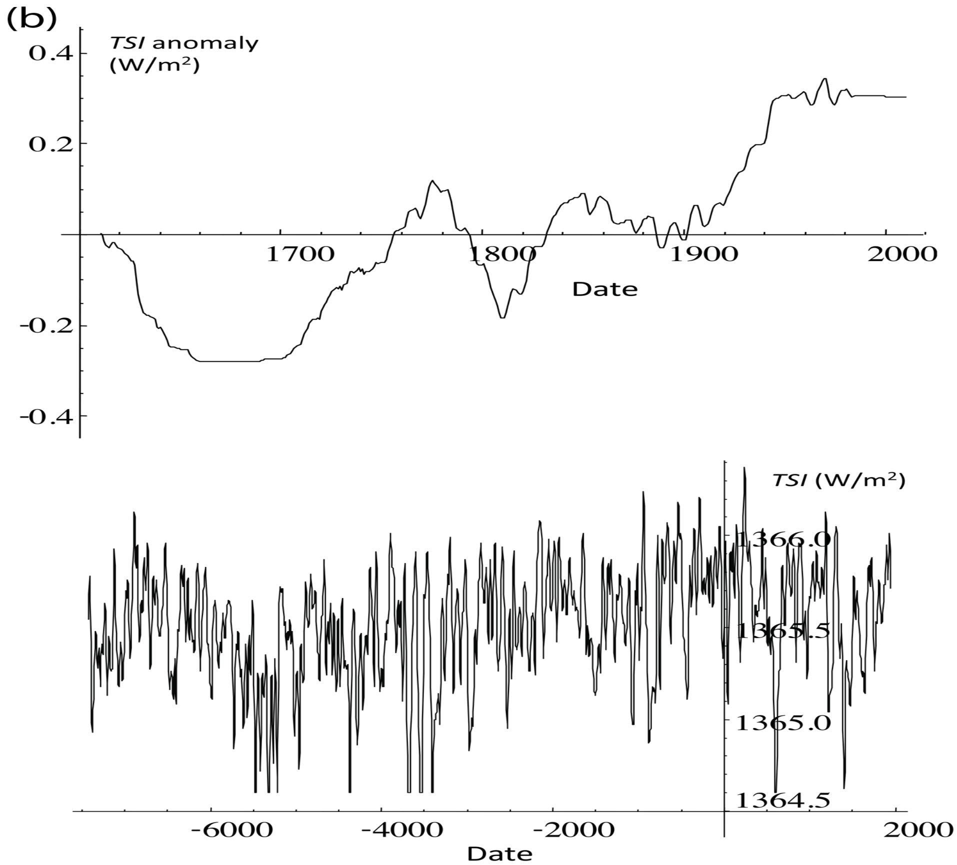

Figure 2b. A comparison of the sunspot derived Total Solar Irradiance (TSI) anomaly (top, used in the ZC and GISS simulations back to 1610, $H \approx 0.4$) with a recent ^{10}Be reconstruction (bottom, total TSI - mean plus anomaly - since 7362 BC, see Fig. 2a for a fluctuation analysis, $H \approx -0.3$) similar to that "spliced" onto the sunspot reconstruction for the period 1000-1610. We can see that the statistical characteristics are totally different with the sunspot variations "wandering" $(H > 0)$ whereas the ^{10}Be reconstruction is "cancelling" $(H < 0)$. The sunspot data were for the "background" (i.e. with no 11 year cycle, see Wang et al., 2005 for details), the data for the ^{10}Be curve were from Shapiro et al., (2011).

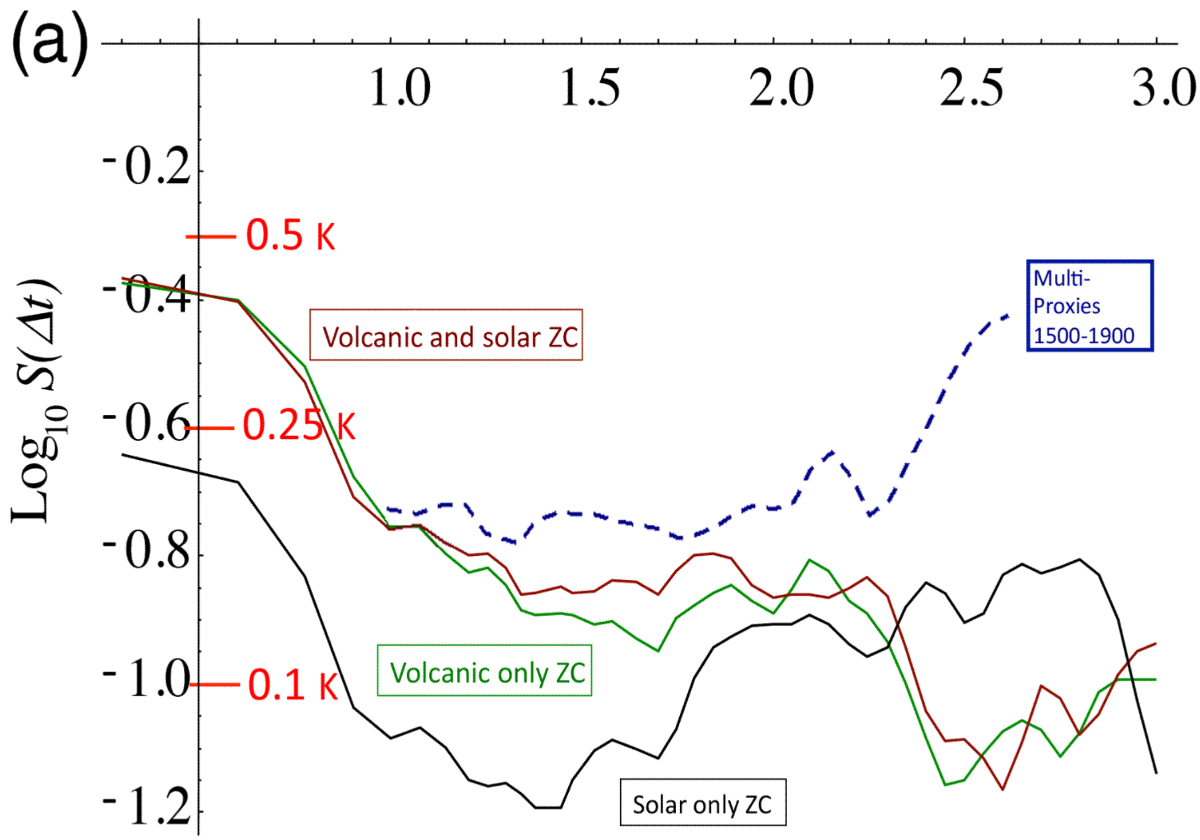

Figure 3a. The RMS Haar fluctuations of the Zebiak−Cane (ZC) model responses (from an ensemble of 100 realizations) with volcanic only (green, from the updated Crowley reconstruction), solar only (black, using the sunspot based background (Wang et al., 2005), and both (brown). No anthropogenic effects were modelled. Also shown for reference are the fluctuations for three multiproxy series (blue, dashed, from 1500-1900, pre-industrial, the fluctuations statistics from the three series were averaged, this curve was taken from Lovejoy and Schertzer, 2012b). We see that all the combined volcanic and solar response of the model reproduces the statistics until scales of ≈ 50-100 years; however at longer time scales, the model fluctuations are substantially too weak – roughly 0.1K (corresponding to ±0.05K) and constant or falling, whereas at 400 yr scales, the temperature fluctuations are ≈0.25K (±0.125) and rising.

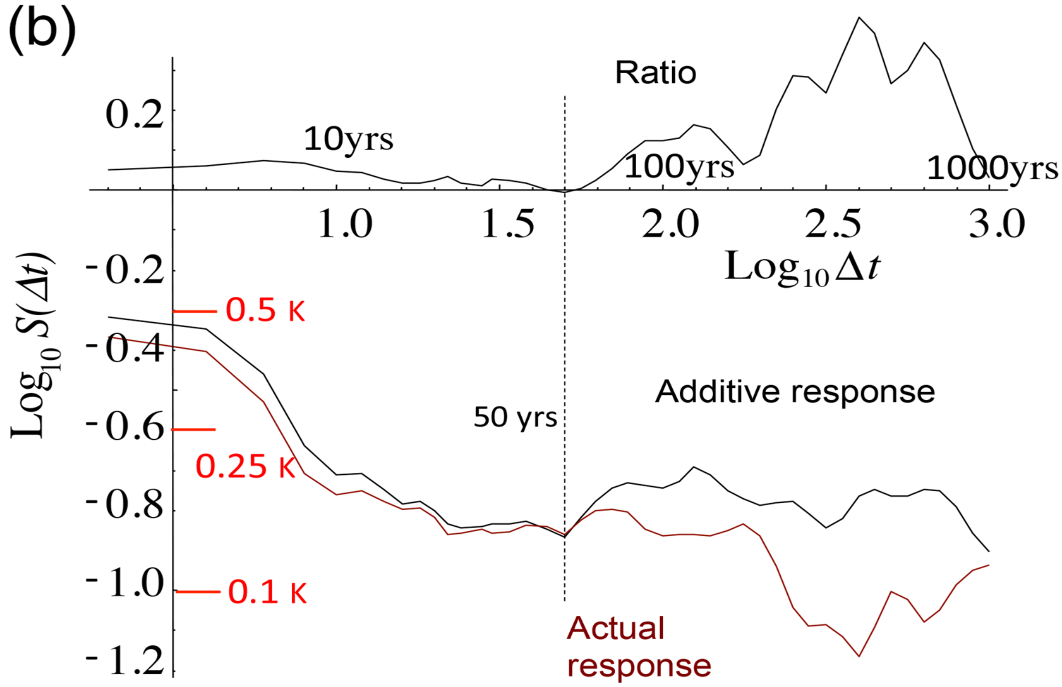

Figure 3b. A comparison of the RMS fluctuations of the ZC model response to combined solar and volcanic forcings (brown, bottom, from Fig. 3a), with the theoretical additive responses (black, bottom) as well as their ratio ($S_{additive} / S_{actual}$ black, top). The additive response was determined from the root mean square of the solar only and volcanic only response variances (from Fig. 3a): additivity implies that the fluctuation variances add (assuming that the solar and volcanic forcings are statistically independent). We can see that after about 50 years, there are strong negative feedbacks, the solar and volcanic forcings are subadditive, see Fig. 3c for a blow up of the ratio.

(c)

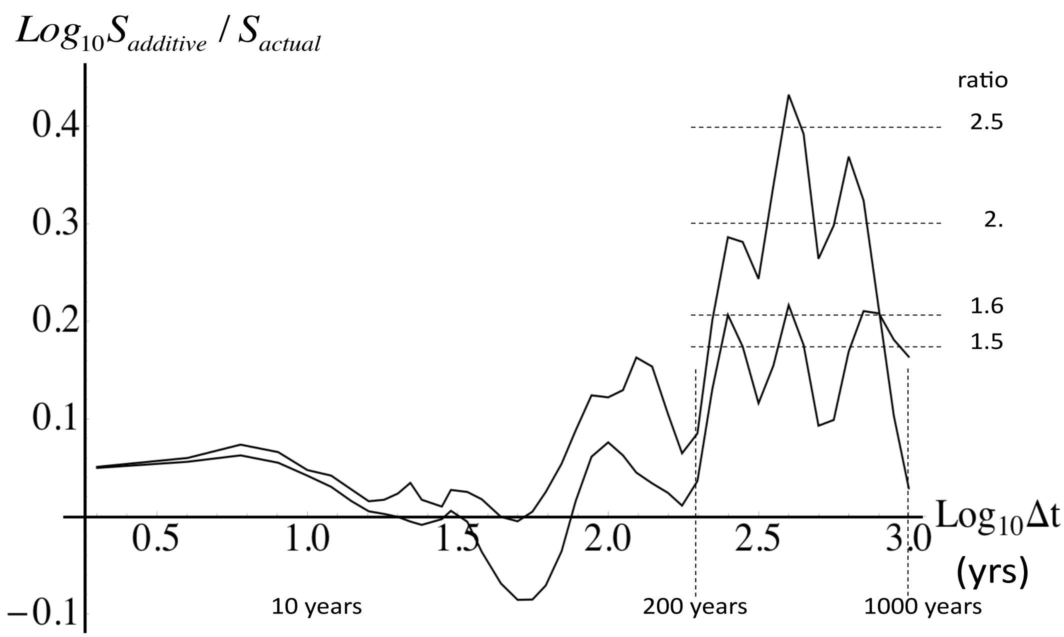

Figure 3c. An enlarged view of the ratio of the linear to nonlinear responses (from Fig. 3b). The top curve assumes for the combined forcing, the linearity of the response and statistical independence of the solar and volcanic forcings, whereas the bottom curve assumes only that the combined response to the forcing is linearuses the actual response to the combined forcings. The maximum at around 400 yrs (top curve) corresponds to a factor $\approx$ 2 ($\approx$1.5, bottom curve) of negative feedback between the solar and volcanic forcings. The decline at longer durations (Δt’s the single 1000 yr fluctuation) is likely to be an artefact of the limited statistics at these scales.

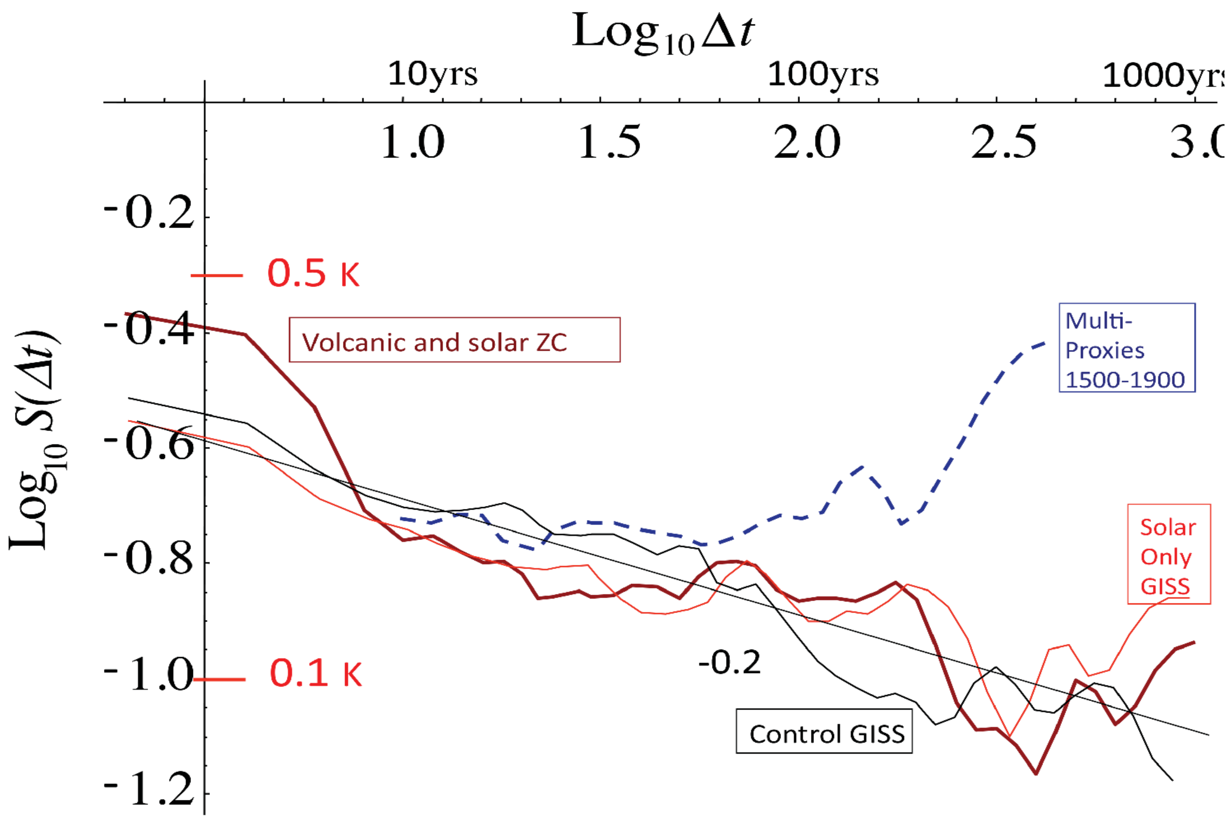

Figure 4. A comparison of the Zebiak-Cane (ZC) model combined (volcanic and solar forcing) response (thick brown) with GISS-E2-R simulations with solar only forcing (red) and a control run (no forcings, black), the GISS structure functions are for land, northern hemisphere, reproduced from Lovejoy et al., (2013).

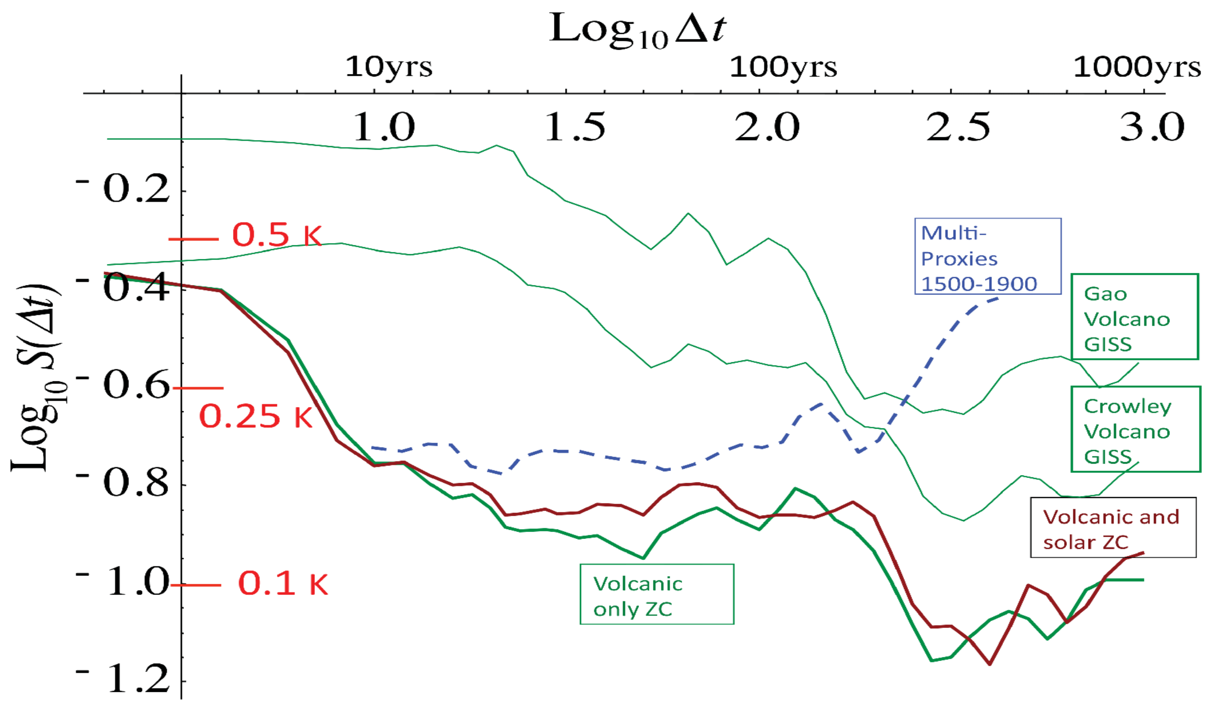

Figure 5. A comparison of the volcanic forcings for the ZC model (bottom green) and for the GISS-E2-R GCM for two different volcanic reconstructions (Gao et al., 2008, and Crowley, 2000) (top green curves, reproduced from Lovejoy et al., 2013). Also shown is the combined response (ZC, brown) and the preindustrial multiproxies (dashed blue).

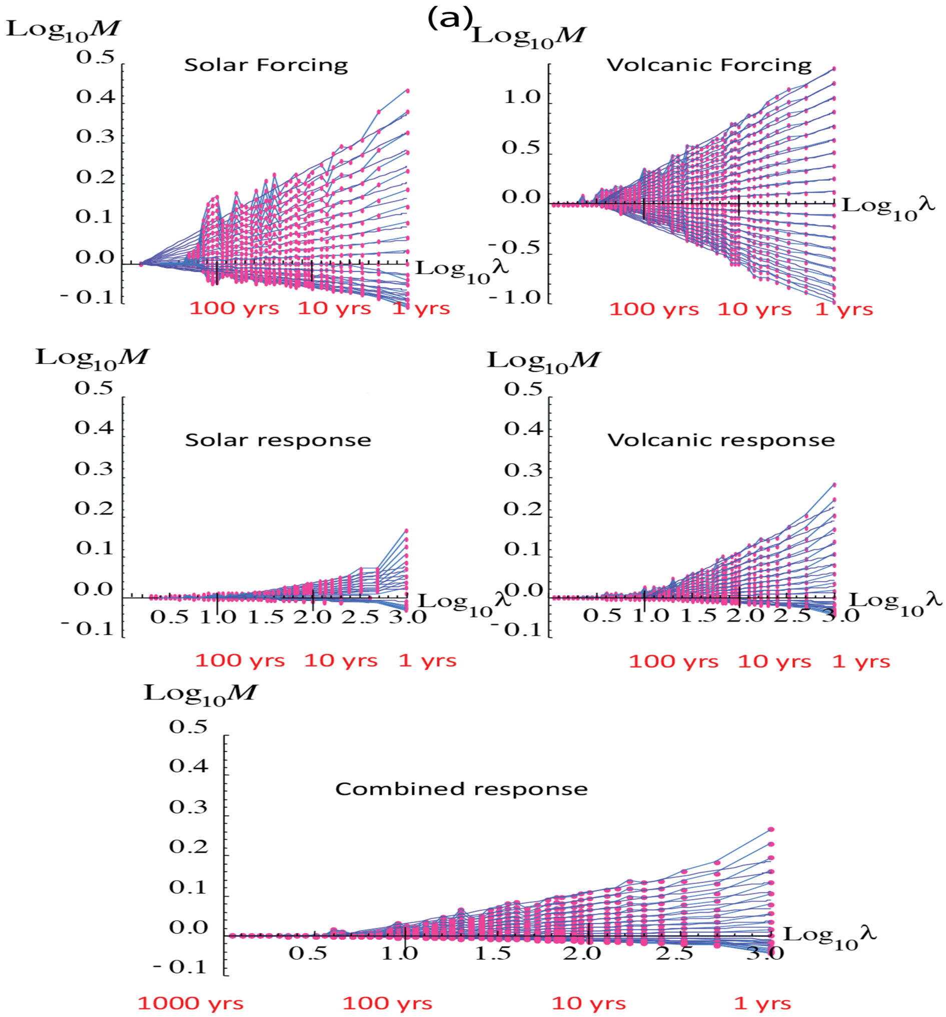

Figure 6a. Analysis of the fluxes/cascade structures of the ZC forcings (top row) and ZC temperature responses (middle, bottom rows); the normalized trace moments (Eq. (11)) are plotted for q = 2, 1.9, 1.8, 1.7, 1.6, …0.1. Upper left is solar forcing (last 400 yrs only, mostly sunspot based), upper right is volcanic, middle left, solar response (last 400 yrs), middle right (volcanic response), lower left, response to combined forcings (last 1000 yrs). Note that all axes are the same except for volcanic. For the solar, only the last 400 yrs were used since this was reconstructed using the more reliable sunspot based method. The earlier ^{10}Be based reconstruction had relatively poor resolution and is not shown. Since the volcanic variability was so dominant, for the combined response (bottom left) the entire series was used. The red points and lines are the empirical values, the blue lines are regressions constrained to go through a single outer scale point, see eq. (11). In comparing the different parts of the figure, note in particular i) the log-log linearity for different statistical moments, ii) the fact that the lines for different moments reasonably cross at a single outer scale,

and iii) the overall amplitude of the fluctuations – for example by visually comparing the range of the $q = 2$ moments (the top series) as we move from one graph to another.

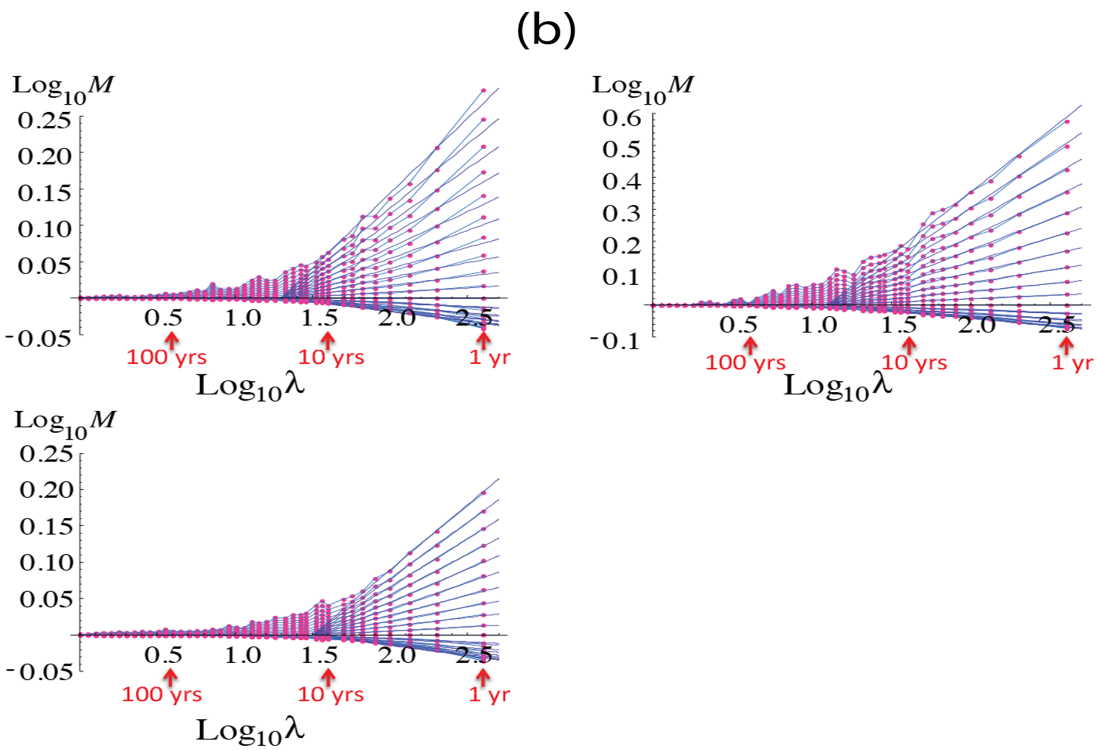

Figure 6b. The above shows the responses for the GISS-E2-R simulations (northern hemisphere, land, 1500-1900), $\lambda=1$ corresponds to 400 yrs. The upper left is for the response to the Crowley reconstructed volcanic forcings (same as used in the ZC simulations, not the change in the vertical scale), the upper right for the Gao reconstructed volcanic forcings and the lower left is for the solar only (mostly sunspot based, same as used in the ZC simulations).

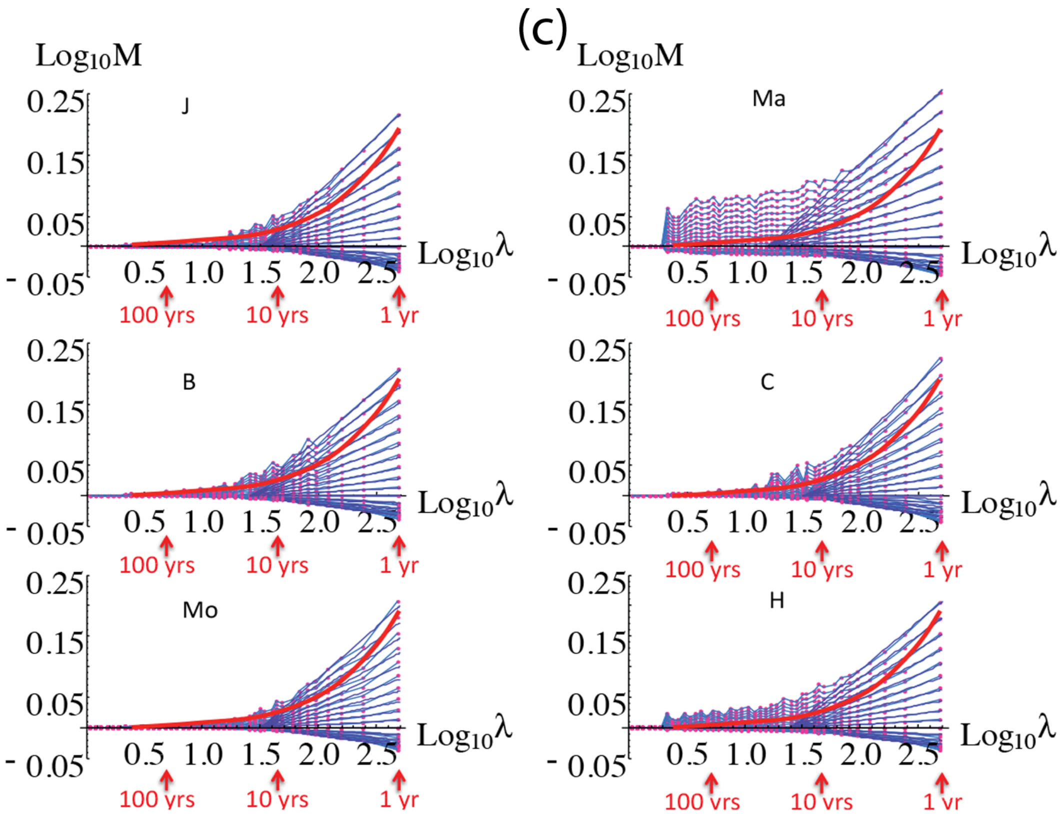

Figure 6c. Trace moment analysis of six annual resolution multiproxies, J = Jones, Ma = Mann 98, B = Briffa, C = Crowley, Mo = Moberg, H = Huang, the curves are reproduced with permission from figure 11.8, of Lovejoy and Schertzer, (2013), where full details and references are given. All were for the pre-industrial period 1500-1900 AD; $\lambda=1$ corresponds to 400 yrs. The curve shows the generic convergence of the envelope of curves to a quasi-Gaussian process, the proximity of the curve to the envelope indicates that with the possible exception of the Mann curve, the intermittency is low.