



# Haar wavelets, fluctuations and structure functions: convenient choices for geophysics

S. Lovejoy<sup>1</sup> and D. Schertzer<sup>2</sup>

<sup>1</sup>Physics, McGill University, 3600 University st., Montreal, Que. H3A 2T8, Canada

<sup>2</sup>LEESU, École des Ponts ParisTech, Université Paris-Est, 6-8 Av. B. Pascal, 77455 Marne-la-Vallée Cedex 2, France

Correspondence to: S. Lovejoy (lovejoy@physics.mcgill.ca) and D. Schertzer (daniel.schertzer@enpc.fr)

Received: 2 May 2012 – Accepted: 30 July 2012 – Published: 13 September 2012

**Abstract.** Geophysical processes are typically variable over huge ranges of space-time scales. This has led to the development of many techniques for decomposing series and fields into fluctuations  $\Delta v$  at well-defined scales. Classically, one defines fluctuations as differences:  $(\Delta v)_{\text{diff}} = v(x + \Delta x) - v(x)$  and this is adequate for many applications ( $\Delta x$  is the “lag”). However, if over a range one has scaling  $\Delta v \propto \Delta x^H$ , these difference fluctuations are only adequate when  $0 < H < 1$ . Hence, there is the need for other types of fluctuations. In particular, atmospheric processes in the “macroweather” range  $\approx 10$  days to 10–30 yr generally have  $-1 < H < 0$ , so that a definition valid over the range  $-1 < H < 1$  would be very useful for atmospheric applications.

A general framework for defining fluctuations is wavelets. However, the generality of wavelets often leads to fairly arbitrary choices of “mother wavelet” and the resulting wavelet coefficients may be difficult to interpret. In this paper we argue that a good choice is provided by the (historically) first wavelet, the Haar wavelet (Haar, 1910), which is easy to interpret and – if needed – to generalize, yet has rarely been used in geophysics. It is also easy to implement numerically: the Haar fluctuation  $(\Delta v)_{\text{Haar}}$  at lag  $\Delta x$  is simply equal to the difference of the mean from  $x$  to  $x + \Delta x/2$  and from  $x + \Delta x/2$  to  $x + \Delta x$ . Indeed, we shall see that the interest of the Haar wavelet is this relation to the integrated process rather than its wavelet nature per se.

Using numerical multifractal simulations, we show that it is quite accurate, and we compare and contrast it with another similar technique, detrended fluctuation analysis. We find that, for estimating scaling exponents, the two methods are very similar, yet Haar-based methods have the advantage of being numerically faster, theoretically simpler and physically easier to interpret.

## 1 Introduction

In many branches of science – and especially in the geosciences – one commonly decomposes space-time signals into components with well-defined space-time scales. A widely used method is Fourier analysis which – when combined with statistics – can be used to quantify the contribution of structures of a given frequency and/or wavenumber to the total variance of the process. Often, however, real space analyses are preferable, not only because they are simpler to perform, but more importantly – because they are simpler to interpret. The latter are based on various definitions of fluctuations at a given scale and location, the simplest being the differences:

$$(\Delta v(\Delta x))_{\text{diff}} \equiv |\delta_{\Delta x} v|; \delta_{\Delta x} v(x + \Delta x) - v(x), \quad (1)$$

where the position is  $x$ , and the scale (“lag”)  $\Delta x$  and  $\delta$  is the difference operator (the parentheses with the “diff” subscript are only used when it is important to distinguish the difference fluctuation from others). For simplicity, we consider a scalar function ( $v$ ) of a scalar position and/or of time. We consider processes whose statistics of a given order are translationally invariant (statistically homogeneous in space, stationary in time). The resulting fluctuations can then be statistically characterized, the standard method being via the moments  $\langle \Delta v^q \rangle$ . When  $q = 2$ , these are the classical structure functions; when  $q \neq 2$ , they are the “generalized” counterparts.

The space-time variability of natural systems can often be broken up into various “scaling ranges” over which the fluctuations vary in a power law manner with respect to scale. Over these ranges, the fluctuations follow:

$$\Delta v = \varphi_{\Delta x} \Delta x^H, \quad (2)$$

where  $\Delta x$  is a random field (or series, or transects thereof at resolution  $\Delta x$ ) and the equality is a statistical one, i.e. in the sense of probability distributions. Below for simplicity, we consider spatial transects with lag  $\Delta x$ , but the results are identical for temporal lags  $\Delta t$ . In fluid systems such as the atmosphere,  $\varphi$  is a turbulent flux. For example, the Kolmogorov law corresponds to taking  $v$  as a velocity component and  $\varphi = \varepsilon^{1/3}$  with  $\varepsilon$  as the energy flux. The classical turbulence laws assumed fairly uniform fluxes, typically that  $\varphi$  is a quasi-Gaussian process.

Taking the  $q$ -th power of Eq. (2) and ensemble averaging, we see that the statistical characterization of the fluctuations in terms of generalized structure functions is

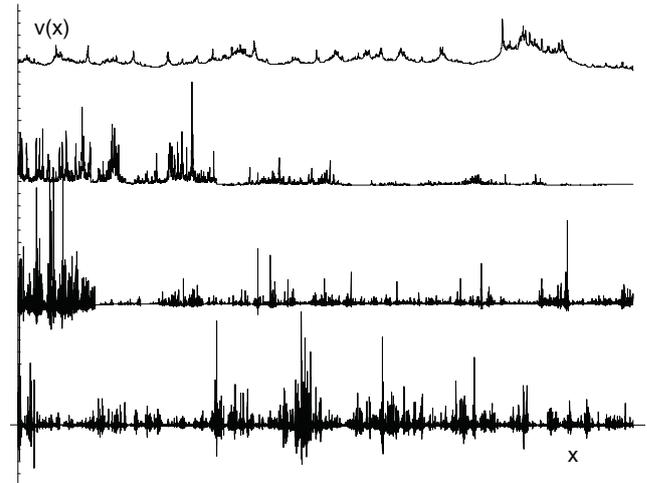
$$S_q(\Delta x) = \langle (\Delta v(\Delta x))^q \rangle = \langle \varphi_{\Delta x}^q \rangle \Delta x^{qH} \approx \Delta x^{\xi(q)}$$

$$\langle \varphi_{\Delta x}^q \rangle = \left( \frac{L}{\Delta x} \right)^{K(q)}; \quad \xi(q) = qH - K(q) \quad (3)$$

where  $S_q(\Delta x)$  is the  $q$ -th order structure function and  $\xi(q)$  is the corresponding exponent,  $K(q)$  is the moment scaling function that characterizes the flux and  $L$  is that outer scale of the scaling regime (e.g. of a cascade process). In the classical quasi-Gaussian case,  $K(q) = 0$  so that  $\xi(q)$  is linear. More generally, if the field is intermittent – for example if it is the result of a multifractal process – then the exponent  $K(q)$  is generally non linear and convex and characterizes the intermittency. Since the ensemble mean of the flux spatially averaged at any scale is the same (i.e.  $\langle \varphi_{\Delta x} \rangle$  is constant, independent of  $\Delta x$ ), we have  $K(1) = 0$  and  $\xi(1) = H$ . The physical significance of  $H$  is thus that it determines the rate at which mean fluctuations grow ( $H > 0$ ) or decrease ( $H < 0$ ) with scale  $\Delta x$ .

There is a simple and useful expression for the exponent  $\beta$  of the power spectrum  $E(k) \approx k^{-\beta}$ . It is  $\beta = 1 + \xi(2) = 1 + 2H - K(2)$  which is a consequence of the fact that the spectrum is a second-order moment (the Fourier transform of the autocorrelation function).  $K(2)$  is therefore often regarded as the spectral “intermittency correction”. Note however that, even if  $K(2)$  is not so large (it is often in the range 0.1–0.2 for turbulent fields), the intermittency is nevertheless important: firstly, because  $K(q)$  typically grows rapidly with  $q$  so that the high-order moments (characterizing the extremes) can be very large; secondly, because it is an exponent so that even when it is small, whenever the range of scales over which it acts is large enough, its effect can be large; in geo-systems, the scale ratio can be  $10^9$  or larger.

Over the last 30 years, structure functions based on differences and with concave rather than linear  $\xi(q)$  have been successfully applied to many geophysical processes; this is especially true for atmospheric processes at (weather) time scales shorter than  $\approx 10$  days which usually have  $0 > H > 1$ . In



**Fig. 1.** This shows samples of the simulations analysed in Sect. 5. From bottom to top,  $H = -7/10, -3/10, 3/10, 7/10$ . All have  $C_1 = 0.1$ ,  $\alpha = 1.8$ ; the sections are each  $2^{13}$  points long. One can clearly see the change in character of the series when  $H$  changes sign. At the bottom, fluctuations tend to cancel out, fluctuations are “stable”. At the top, they tend to reinforce each other, and fluctuations are “unstable”.

addition, in turbulence theory there are many useful theoretical results for these classical structure functions (e.g. Monin and Yaglom, 1975). However, it turns out that the use of differences to define the fluctuations is overly (and needlessly) restrictive. The problem becomes evident when we consider that the mean difference cannot decrease with increasing  $\Delta x$ : for processes with  $H < 0$ , the differences simply converge to a spurious constant depending on the highest frequencies available in the sample. Similarly, differences cannot grow faster than  $\Delta x$  – they “saturate” at a linear function – whose slope depends on the lowest frequencies present in the sample; the process no longer has stationary increments. Hence, they cannot be used when  $H > 1$ . Overall, whenever  $H$  is outside the range  $0 > H > 1$ , the exponent  $\xi(q)$  is no longer correctly estimated. The problem is that we need a definition of fluctuations such that  $\Delta v(\Delta x)$  is dominated by wavenumbers in the neighbourhood of  $\Delta x^{-1}$ . Appendix A gives more details about the relation between real and Fourier space and the ranges of  $H$  relevant for various types of fluctuation, and Fig. 1 shows the shape of typical sample series as  $H$  varies.

The need to define fluctuations more flexibly motivated the development of wavelets (e.g. Holschneider, 1995; Torrence and Compo, 1998), the classical difference fluctuation being only a special case, the “poor man’s wavelet”. To change the range of  $H$  over which fluctuations are usefully defined, one must change the shape of the defining wavelet. In the usual wavelet framework, this is done by modifying the “mother wavelet” directly, various derivatives of the Gaussian are popular, especially the second derivative “Mexican hat”. Following this, the fluctuations are calculated

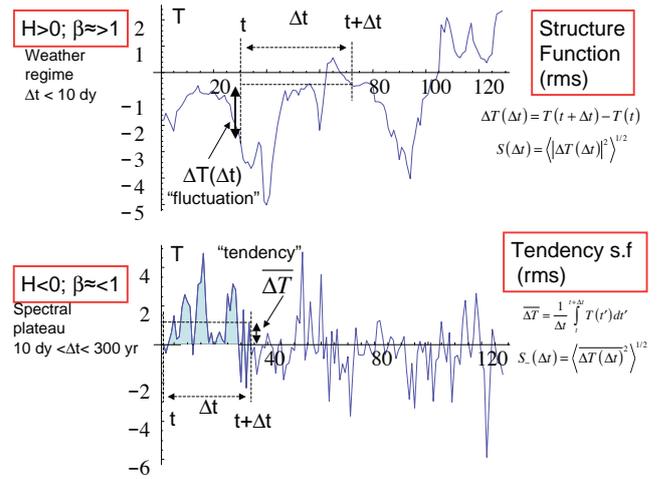
as convolutions with fast Fourier (or equivalent) numerical techniques.

A problem with this usual wavelet implementation is that not only are the convolutions numerically cumbersome, but the physical interpretation of the fluctuations is largely lost. In contrast, when  $0 < H < 1$ , the difference structure function gives direct information on the typical difference ( $q = 1$ ) and typical variations around this difference ( $q = 2$ ) and even typical skewness ( $q = 3$ ) or typical Kurtosis ( $q = 4$ ) or – if the probability tail is algebraic – of the divergence of high-order moments of differences. Similarly, when  $-1 < H < 0$ , one can define the “tendency structure function” (below) which directly quantifies the fluctuation’s deviation from zero and whose exponent characterizes the rate at which the deviations decrease when we average to larger and larger scales. These poor man’s and tendency fluctuations are also very easy to directly estimate from series with uniformly spaced data and – with straightforward modifications – to irregularly spaced data.

In Lovejoy and Schertzer (2012a, b), it is argued that, with only a few exceptions for atmospheric processes in the “weather” regime ( $\Delta t \lesssim 10$  days),  $0 < H < 1$  so that fluctuations increase with scale and the classical use of differences is adequate. However, in the low frequency “macroweather” regime ( $10 \text{ days} \lesssim \Delta t \lesssim 10\text{--}30 \text{ yr}$ ), on the contrary, quite generally (in time, but not in space!) we find  $-1 < H < 0$  so that fluctuations decrease with scale (although for larger  $\Delta t$  – the “climate” regime – it is again in the range  $0 > H > 1$ ). In order to cover the whole range (for the temperature, even millions of years), one must thus use a definition of fluctuations that is valid over the range  $1 > H > -1$ . This range is also appropriate for many solid earth applications, including topography, rock density, magnetic susceptibility, geogravity and geomagnetism; for a review, see Lovejoy and Schertzer (2007). The purpose of this paper is to argue that the Haar wavelet is particularly convenient, combining ease of calculation and of interpretation, and – if needed – it can easily be generalized to accommodate processes with  $H$  outside this range. We demonstrate this using multifractal simulations with various  $H$  exponents to systematically compare it to conventional spectral analysis and to another common – and quite similar – analysis technique, the detrended fluctuation analysis (DFA) method (Peng et al., 1994), and its extension, the multifractal detrended fluctuation analysis (MFDFA) method (Kantelhardt et al., 2001, 2002).

## 2 Defining fluctuations using wavelets

Our goal is to find a convenient definition of fluctuations valid at least over the range  $-1 < H < 1$ , but before doing so, it will be useful to first define tendency fluctuations  $(\Delta v(\Delta x))_{\text{tend}}$ . The first step is to remove the overall mean  $(\overline{v(x)})$  of the series:  $v'(x) = v(x) - \overline{v(x)}$  and then take averages over the lag  $\Delta x$ . For this purpose, we introduce the



**Fig. 2.** The top shows extracts of a multifractal temperature simulation with  $H = 0.5$ , at one hour resolution (in the weather regime). The bottom is the same simulation but at 16-day resolution (in the low frequency weather regime,  $H = -0.4$ ). One can clearly see their differing characters corresponding to  $H > 0$ ,  $H < 0$  respectively (see also Fig. 1). These simulations are used to illustrate the two different types of structure function needed when  $H > 0$  (the usual, difference) and  $H < 0$ , the “tendency” structure function. Whereas the usual structure function yields typical differences in  $T$  over an interval  $\Delta t$ , so that the mean is of no consequence, the tendency structure function uses the average of the field with the mean removed. In both cases, the  $q = 2$  moment was chosen because it is directly related to the spectrum. Taking the square root is useful since the result is then a direct measure of the “typical” fluctuations, e.g. for the  $T =$  the temperature, in units of K.

operator  $\mathcal{T}_{\Delta x}$  defining the tendency fluctuation  $(\Delta v(\Delta x))_{\text{tend}}$  as

$$(\Delta v(\Delta x))_{\text{tend}} = \mathcal{T}_{\Delta x} v = \left| \frac{1}{\Delta x} \sum_{x \leq x' \leq x + \Delta x} v'(x') \right| \quad (4)$$

or, with the help of the summation operator  $\mathcal{S}$  (not to be confused with the structure function), equivalently by

$$(\Delta v(\Delta x))_{\text{tend}} = \left| \frac{1}{\Delta x} \delta_{\Delta x} \mathcal{S} v' \right|; \quad \mathcal{S} v' = \sum_{x' \leq x} v'(x') \quad (5)$$

We can also use the suggestive notation  $(\Delta v(\Delta x))_{\text{tend}} = \overline{\Delta v}$  (see the schematic Fig. 2). When  $-1 < H < 0$ ,  $(\Delta v(\Delta x))_{\text{tend}}$  has a straightforward interpretation in terms of the mean tendency of the data to decrease with averaging. The tendency fluctuation is also easy to implement: simply remove the *overall* mean and then take the mean over intervals  $\Delta x$ : this is equivalent to taking the mean of the differences of the running sum.

It is now straightforward to define the Haar fluctuation by taking the second differences of the mean:

$$\begin{aligned}
 (\Delta v(\Delta x))_{\text{Haar}} &= \mathcal{H}_{\Delta x} v = \frac{2}{\Delta x} \delta_{\Delta x/2}^2 \mathcal{S}v \\
 &= \frac{2}{\Delta x} ((s(x) + s(x + \Delta x)) - 2s(x + \Delta x/2)) \\
 &= \frac{2}{\Delta x} \left[ \sum_{x+\Delta x/2 < x' < x+\Delta x} v(x') \right. \\
 &\quad \left. - \sum_{x < x' < x+\Delta x/2} v(x') \right]; s(x) = \mathcal{S}v, \quad (6)
 \end{aligned}$$

where  $\mathcal{H}_{\Delta x}$  is the Haar operator (not to be confused with the exponent  $H$ ). Note the use of the shorthand notation  $s(x) = \mathcal{S}v$ .

As discussed in Sect. 3, this is still a valid wavelet (see Fig. 3). However, it is almost trivial to calculate and it can be used for  $-1 < H < 1$  (Appendix A; note there is an extra factor of  $\Delta x^{-1}$  with respect to the Haar wavelet coefficient; see below). The numerical factor (2) is used so that the Haar fluctuation is close to the usual differences when  $0 < H < 1$ , (below). While this may still look a little complicated, it is actually quite simple: in words, the Haar fluctuation at lag  $\Delta x$  is simply the difference of the mean from  $x$  to  $x + \Delta x/2$  and from  $x + \Delta x/2$  to  $x + \Delta x$  i.e. it deals with fluctuations in the integrated process.

From the definitions, it is easy to obtain the relation:

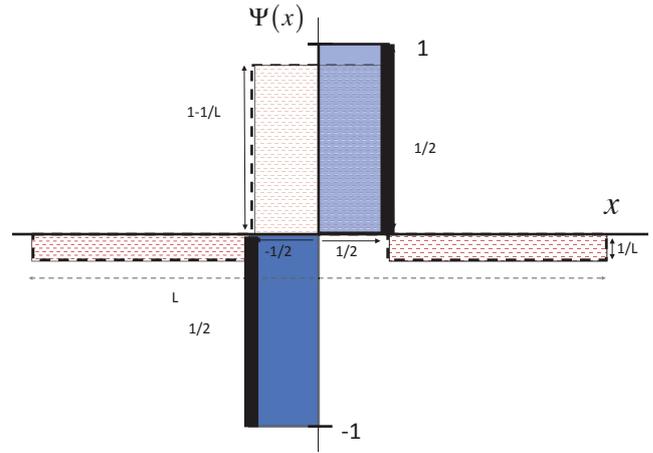
$$\mathcal{H}_{\Delta x} = 2\mathcal{T}_{\Delta x/2} \delta_{\Delta x/2} = 2\delta_{\Delta x/2} \mathcal{T}_{\Delta x/2}. \quad (7)$$

This is useful for interpreting the Haar fluctuations in terms of differences and tendencies (Sect. 5.2). Since the difference operator removes any constants, the tendency operator can be replaced by an averaging operator and is thus insensitive to the addition of constants; Eq. (7) therefore means that the Haar fluctuation is simply the difference of the series that has been degraded in resolution by averaging. Note that, for a series of length  $L$ , only a range of scales  $L/2$  is accessible; the definition Eq. (6) identifies the largest lag with scale  $L$ , the smallest with scale 2 (we could equally well have modified the definition so that this would be  $L/2$  and 1).

If needed, the Haar fluctuations can easily be generalized to higher-order  $n$  by using  $n$ -th order differences on the sum  $s$ :

$$\mathcal{H}_{\Delta x}^{(n)} = \frac{(n+1)}{\Delta x} \delta_{\Delta x/(n+1)}^{n+1} \mathcal{S}. \quad (8)$$

We note that the  $n$ -th order Haar fluctuation is valid over the range  $-1 < H < n$ ; due to the  $n+1$ -th order difference, it is insensitive to polynomials of order  $n$  in the sum and of order  $n-1$  in the original series, although, as we see below, for analyzing scaling series, this insensitivity has no particular advantage. Generalizations valid to order  $H < -1$  are



**Fig. 3.** This shows the “poor man’s wavelet” by black bars representing the amplitudes of Dirac  $\delta$  functions, (the basis of the usual difference structure function, valid for  $0 < H < 1$ ; this is only symbolic since the true  $\delta$  function amplitude is of course infinite), the Haar wavelet (the basis of the Haar structure function, uniform blue shading, the second difference of the running sum, valid for  $-1 < H < 1$ ), and the wavelet used for the “tendency” structure function (valid for  $-1 < H < 0$ ), stippled shading.

also possible by iterating the sum operator. Note further that  $\mathcal{H}_{\Delta x}^{(1)} = \mathcal{H}_{\Delta x}$  and  $\mathcal{H}_{\Delta x}^{(0)} v' = \mathcal{T}_{\Delta x} v'$ ; hence, when applied to a series with zero mean,  $\mathcal{H}_{\Delta x}^{(0)} = \mathcal{T}_{\Delta x}$ .

Consider for a moment the case  $n = 2$  which defines the “quadratic Haar fluctuation”:

$$\begin{aligned}
 (\Delta v(\Delta x))_{\text{Haarquad}} &= \frac{3}{\Delta x} (s(x + \Delta x) - 3s(x + \Delta x/3) \\
 &\quad + 3s(x - \Delta x/3) - s(x - \Delta x)). \quad (9)
 \end{aligned}$$

This fluctuation is sensitive to structures of size  $\Delta x^{-1}$  – and hence useful – over the range  $-1 < H < 2$ , and it “filters” out polynomials of order 1 (lines); this is thus essentially equivalent to the commonly used quadratic ( $n = 2$ ) MFDFA (multifractal detrended fluctuation analysis, Sect. 4) technique that we numerically investigate below.

### 3 Fluctuations and wavelets

Wavelet analysis defines fluctuations with the help of a basic “mother wavelet”  $\Psi(x)$  and a convolution:

$$\Delta v(x, \Delta x) = \int v(x') \Psi\left(\frac{x' - x}{\Delta x}\right) dx', \quad (10)$$

where we have kept the notation  $\Delta v$  to indicate “fluctuation”.  $\Delta v$  defined this way is called a “wavelet coefficient”. The basic “admissibility” condition on  $\Psi(x)$  (so that it is a valid wavelet) is that it has zero mean. If we take the fluctuation  $\Delta v$  as the symmetric (centred) difference  $\Delta v = v(x + \Delta x/2) - v(x - \Delta x/2)$ , this is equivalent to using

$$\Psi(x) = \delta(x - 1/2) - \delta(x + 1/2) \tag{11}$$

where “ $\delta$ ” is the Dirac delta function (Eq. 1, see Fig. 3; due to the statistical translational invariance, the coordinate shift by  $1/2$  is immaterial). As discussed, this “poor man’s” wavelet is adequate for many purposes. As the generality of the definition (Eq. 10) suggests, all kinds of special wavelets can be introduced. For example, orthogonal wavelets can be used which are convenient if one wishes to reconstruct the original function from the fluctuations, or to define power spectra locally at a point  $x$  rather than globally, (averaged over all  $x$ ).

The other simple to interpret fluctuation, the “tendency fluctuation” (Eq. 4), can also be expressed in terms of wavelets:

$$\Psi(x) = I_{[-1/2, 1/2]}(x) - \frac{I_{[-L/2, L/2]}(x)}{L}; L \gg 1 \tag{12}$$

where  $I$  is the indicator function:

$$I_{[a,b]}(x) = \begin{cases} 1 & a \leq x \leq b \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

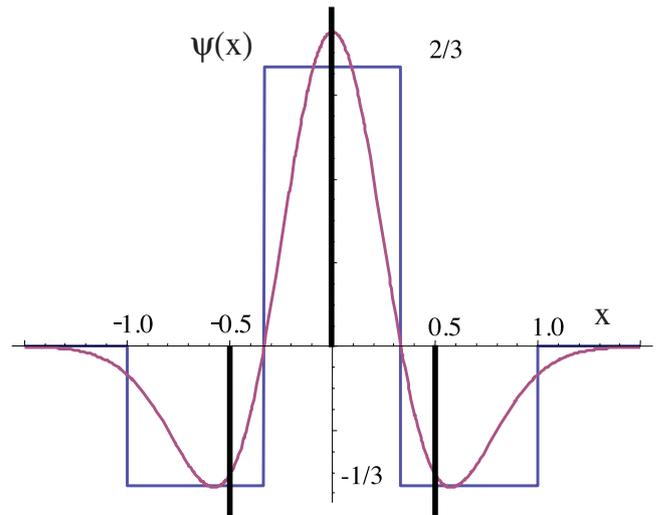
See Fig. 3 for a schematic. The first term in Eq. (12) represents the sum, whereas the second removes the mean where  $L \gg 1$  is the overall length of the data set. The removal of the mean in this way is necessary so that the wavelet satisfies the admissibility condition that its mean is zero. The only essential difference between the wavelet defined by Eq. (12) and the “tendency fluctuation” defined in Eq. (4) (where the mean is removed beforehand) is the extra normalization by  $\Delta x$  in Eq. (4) which changes the exponent by one.

Figure 3 also shows the wavelet corresponding to the Haar fluctuation:

$$\Psi(x) = \begin{cases} 1; & 0 \leq x < 1/2 \\ -1; & -1/2 \leq x < 0 \\ 0; & \text{otherwise} \end{cases} \tag{14}$$

To within the division by  $\Delta x$ , and a factor of 2, the result is the equivalent of using the Haar fluctuations (Eq. 6) as indicated in Fig. 3. It can be checked (Appendix A) that the Fourier transform of this wavelet is  $\propto k^{-1} \sin^2(k/4)$  so that compared to the difference and tendency fluctuations, whose mother wavelet has cutoffs only at high or low wavenumbers, it is better localized in Fourier space with both low and high wavenumber fall-offs ( $\approx k$  and  $\approx k^{-1}$  respectively). The Haar fluctuation is a special case of the Daubechies family of wavelets (see e.g. Holschneider, 1995 and for some applications, see Koscielny-Bunde et al., 1998, 2006; Ashok et al., 2010). It is also worth mentioning the paper by Veneziano and Furcolo (2003) who pointed out that Haar wavelets can profitably be used in theoretical calculations on multifractal cascades.

If we are interested in fluctuations valid up until  $H = 2$ , we need merely to take derivatives (or differences). For example,



**Fig. 4.** The popular Mexican hat wavelet (the second derivative of the Gaussian red, valid for  $-1 < H < 2$ ) compared with the (negative) of the second finite difference wavelet (black bars representing the relative weights of  $\delta$  functions, valid for  $0 < H < 2$ ), and the second-order “quadratic” Haar wavelet (blue) Eq. (9).

the second finite difference wavelet is

$$\Psi(x) = \frac{1}{2} \left( \delta\left(x + \frac{1}{2}\right) + \delta\left(x - \frac{1}{2}\right) \right) - \delta(x) \tag{15}$$

(corresponding to  $\Delta v = (v(x + \Delta x) + v(x - \Delta x/2))/2 - v(x)$  using centred differences, Fig. 4). Alternatively, we can compare this with the quadratic Haar wavelet, Eq. (9):

$$\psi(x) = \begin{cases} -1/3; & -1 < x < -1/3 \\ 2/3; & -1/3 < x < 1/3 \\ 1/3; & 1/3 < x < 1 \\ 0; & \text{otherwise} \end{cases} \tag{16}$$

which is very close to the Mexican hat wavelet (see Fig. 4) but is numerically much easier to implement. As with the Mexican hat, it is valid, for  $-1 < H < 2$ . It turns out that for analyzing multifractals, even over their common range of validity ( $0 < H < 1$ ), the Haar structure function is somewhat better than the poor man’s wavelet-based structure function (i.e. the usual structure function). This is demonstrated on numerical examples in Sect. 5. Just as one can use higher and higher-order finite differences to extend the range of  $H$  values, one can simply use higher and higher-order derivatives of the Mexican hat.

Wavelet mathematics are seductive – and there certainly exist areas such as speech recognition – where both frequency/wavenumber and temporal or spatial localization of statistics are necessary. However, the use of wavelets in geophysics is often justified by the existence of strong localized structures that are cited as evidence that the underlying process is statistically nonstationary/statistically inhomogeneous (i.e. in time or in space respectively). However,

multifractal cascades produce exactly such structures – the “singularities” – but their statistics are strictly translationally invariant. In this case, the structures are simply the result of the strong singularities. Nevertheless, the local statistics are usually uninteresting, and we usually average over them in order to improve statistical estimates.

Before continuing to discuss other related methods for defining fluctuations, we should mention that there have been strong claims that wavelets are indispensable for analyzing multifractals (e.g. Arneodo et al., 1999). In as much as the classical definition of fluctuations as first differences is already a wavelet, this is true (see however the DFA discussion below). Nevertheless, for most applications, one does not need the full arsenal of wavelet techniques. At a more fundamental level, however, the claim is at least debatable since, mathematically, wavelet analysis is a species of functional analysis; i.e. its objects are mathematical functions defined at mathematical points. On the contrary, the generic multifractal processes – cascades – are singular measures, they are “delocalized”, see Schertzer and Lovejoy (1992), Schertzer et al. (2002, 2010), so that strictly speaking, they are outside the scope of wavelet analysis. Wavelets may therefore not always be appropriate. A particular example where wavelets may be misleading is the popular “modulus maximum” technique where one attempts to localize the singularities (Bacry et al., 1989; Mallat and Hwang, 1992). If the method is applied to a cascade process, then, as one increases the resolution (“zooming in”), the singularity localization never converges implying that the significance of the method is not as obvious as is claimed. Similar comments apply to the more recent “wavelet leader” technique (Serrano and Figliola, 2009).

#### 4 Fluctuations as deviations from polynomial regressions: the DFA, MF DFA techniques

We now discuss a variant method that defines the fluctuations in a different way, but still quantifies the statistics in the manner of the generalized structure function by assuming that the fluctuations are stationary: multifractal detrended fluctuation analysis. The MF DFA technique (Kantelhardt et al., 2001, 2002) is a straightforward generalization of the original detrended fluctuation analysis (DFA, Peng et al., 1994). The method works for 1-D sections, so consider the transect  $v(x)$  of series on a regular grid with resolution = 1 unit,  $L$  units long. The MF DFA starts by replacing the original series by the running sum:

$$s(x) = S v' = \sum_{x' \leq x} v'(x') \quad (17)$$

(cf. Eq. 5). As discussed above, a sum is a finite difference of order  $-1$ , so that analysing  $s$  rather than  $v$  allows the treatment of transects with  $H$  down to  $-1$  (we will discuss the upper bound shortly). We now divide the range

into  $L_{\Delta x} = \text{int}(L/\Delta x)$  disjoint intervals, each indexed by  $i = 1, 2, \dots, L_{\Delta x}$  (“int” means “integer part”). For each interval starting at  $x = i \Delta x$ , one defines the “ $n$ -th order” fluctuation by the standard deviation  $\sigma_s$  of the difference of  $s$  with respect to a polynomial  $F_n(x, \Delta x) = \sigma_s$  in the running sum of  $v$  as follows:

$$\begin{aligned} F_n(x, \Delta x) &= \sigma_s \\ &= \left[ \frac{1}{\Delta x} \sum_{j=1}^{\Delta x} (s((i-1)\Delta x + j) - p_{n,i}(j))^2 \right]^{1/2}; \\ (\Delta v(\Delta x))_{\text{MF DFA}} &= \frac{F_n}{\Delta x} = \frac{\sigma_s}{\Delta x}. \end{aligned} \quad (18)$$

$x = i \Delta x$  and  $p_{n,i}(x)$  is the  $n$ -th order polynomial regression of  $s(x)$  over the  $i$ -th interval of length  $\Delta x$ ; it is the polynomial that “detrends” the running sum  $s$ ; the fluctuation  $\sigma_s$  is the root-mean-square deviation from the regression.  $F$  is the standard DFA notation for the “fluctuation function”; as indicated, it is in fact a fluctuation in the integrated series quantified by the standard deviation  $\sigma_s$ . Using  $F$ , we have  $F_n = \sigma_s = \Delta x (\Delta v)_{\text{MF DFA}}$ . So in terms of the usually cited DFA fluctuation exponent  $\alpha$  (not the Levy  $\alpha!$ ), we have  $F_n \approx \Delta x^\alpha$  and  $\alpha = 1 + H$  so that the (second-order) spectral exponent  $\beta = 2\alpha - 1 - K(2)$ , hence the oft-cited relation  $\beta = 2\alpha - 1$  ignores the intermittency correction  $K(2)$ . The fluctuation  $(\Delta v)_{\text{MF DFA}}$  is very close to the  $n$ -th order finite difference of  $v$ . It is also close to a wavelet-based fluctuation where the wavelet is of  $n$ -th order, although this is not completely rigorous (see however Kantelhardt et al., 2001; Taquu et al., 1995). As a practical matter, in the above sum, we start with intervals of length  $\Delta x = n+1$  and normalize by the number of degrees of freedom of the regression, i.e. by  $\Delta x - n$  not by  $\Delta x$ ; Eq. (18) is the approximation for large  $\Delta x$ .

In the usual presentation of the MF DFA, one considers only a single realization of the process and averages powers of the fluctuations over all the disjoint intervals  $i$ . However, more generally, we can consider a statistical ensemble and average over all the intervals and realizations. If in addition, we consider moments other than  $q = 2$ , then we obtain the multifractal DFA (MF DFA) with the following statistics:

$$\begin{aligned} \langle |F_n(\Delta x)|^q \rangle^{1/q} &= \langle |\sigma_s(\Delta x)|^q \rangle^{1/q} \\ &= \langle |\Delta x (\Delta v(\Delta x))_{\text{MF DFA}}|^q \rangle^{1/q} = \Delta x^{h(q)}, \end{aligned} \quad (19)$$

where we have used the exponent  $h(q)$  as defined in Davis et al. (1996) and Kantelhardt et al. (2002). In what follows, we will refer to the method as the MF DFA technique, even though the only difference with respect to the DFA is the consideration of the  $q \neq 2$  statistics.

From our analysis above, we see that the  $n$ -th order MF DFA exponent  $h(q)$  is related to the usual exponents by

$$h(q) = 1 + \xi(q)/q = (1 + H) - K(q)/q; \quad -1 < H < n \quad (20)$$

where we have used the relation  $\langle (\Delta v)_{\text{MF DFA}}^q \rangle \approx \Delta x^{\xi(q)}$ . Note that, in Eq. (20), the 1 appears because of the initial integration of order 1 (the sum) in the MF DFA recipe. The method works up to  $H = n$ , because the fluctuations are defined as the residues with respect to  $n$ -th order polynomials of the sum  $s$  corresponding to  $n+1$ -th order differences in  $s$  and hence  $n$ -th order differences in  $v$ . The justification for defining the MF DFA exponent  $h(q)$  in this way is that in the absence of intermittency (i.e. if  $K(q) = \text{constant} = K(0) = 0$ ), one obtains  $h(q) = H = \text{constant}$ . Unfortunately – contrary to what is often claimed – the nonconstancy of  $h(q)$  neither implies that the process is multifractal nor does its constancy imply a monofractal process. To see this, it suffices to consider the (possibly fractionally integrated, order  $H$ ) monofractal  $\beta$  model which has  $K(q) = C_1(q-1)$ . Therefore,  $h(q) = 1 + H - C_1 + C_1/q$ , which is not only non-constant but even diverges as  $q$  approaches zero (indeed the same  $q = 0$  divergence occurs for any universal multifractal with  $\alpha < 1$ ; see Eq. (21) below:  $\lim_{q \rightarrow 0} h(q) \rightarrow \infty$ ).

During the last ten years, the MF DFA technique has been applied frequently to atmospheric data; of special note is the work by A. Bunde and co-workers (e.g. Koscielny-Bunde et al., 1998; Kantelhardt et al., 2001; Bunde et al., 2002, 2005; Lennartz and Bunde, 2009) who effectively took advantage of its ability to handle the  $H < 0$ , needed for macroweather processes. In addition, several numerical studies have compared the MF DFA performance to various wavelet techniques. For example, for estimating exponents, Oswiecimka et al. (2006) find it slightly superior to the wavelet transform modulus maximum method, and Huang et al. (2011) find it somewhat advantageous when compared to the wavelet leader method. However, the exact status of the MF DFA method is not clear since – at least in the multifractal case – a completely rigorous mathematical analysis has not been done, and the literature suffers from misleading claims to the effect that the MF DFA is necessary because, by its construction, it removes nonstationarities due to trends (linear, quadratic etc. up to polynomials order  $n-1$ ) in the data. While it is true that it does remove polynomial trends, these only account for fairly trivial types of nonstationarity: beyond this, the method continues to make the standard (and strong) stationarity assumptions about the statistics of the deviations which are left over after performing the polynomial detrending. In particular, the MF DFA does nothing to remove the most common genuine type of statistical nonlinearity in the geosciences: the diurnal and annual cycles which still strongly break the scaling of the MF DFA statistics, and this for any order  $n$ . Furthermore, the corresponding wavelet or finite difference definition of fluctuations can equally well take into account these polynomial trends, and finally, the method removes trends at all scales and locations so that this emphasis on detrending is misleading. In other words, the MF DFA is a variant with respect to some of the wavelet methods, in particular with respect to the Haar wavelet and

its generalizations. While it may indeed yield accurate exponent estimates, it has the disadvantage that the interpretation of the fluctuations is not straightforward. In comparison, the usual difference and tendency fluctuations and structure functions have simple physical interpretations in terms of magnitudes of changes (when  $1 > H > 0$ ) and magnitudes of average tendencies (when  $-1 < H < 0$ ).

## 5 A comparison of different fluctuations using multifractal simulations

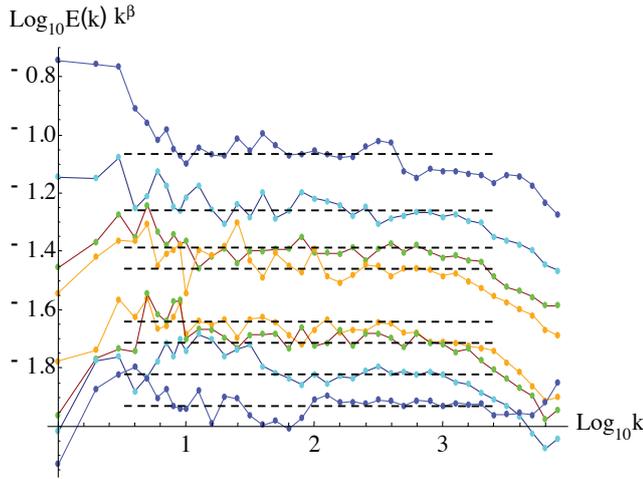
### 5.1 The simulations

To have a clearer idea of the limitations of the various fluctuations for determining the statistics of scaling functions, we shall numerically compare their performance and their corresponding structure functions when applied to the characterization of multifractals. In order to easily compare them with standard spectral analysis, we will only consider the second-order ( $q = 2$ ) structure functions and will use simulations of universal multifractals (Schertzer and Lovejoy 1987, 1997). Recall that universal multifractals are the result of stable, attractive multiplicative cascade processes and have

$$K(q) = \frac{C_1}{\alpha - 1} (q^\alpha - q) \quad (21)$$

where  $0 \leq C_1$  is the codimension of the mean and  $0 \leq \alpha \leq 2$  is the multifractality index (the Lévy index of the generator). The limit  $\alpha \rightarrow 1$  gives  $K(q) = C_1 q \log q$  and may be obtained from Eq. (21) by using l'Hôpital's rule. Recall (Eq. 3) that the corresponding structure function exponent is  $\xi(q) = qH - K(q)$ .

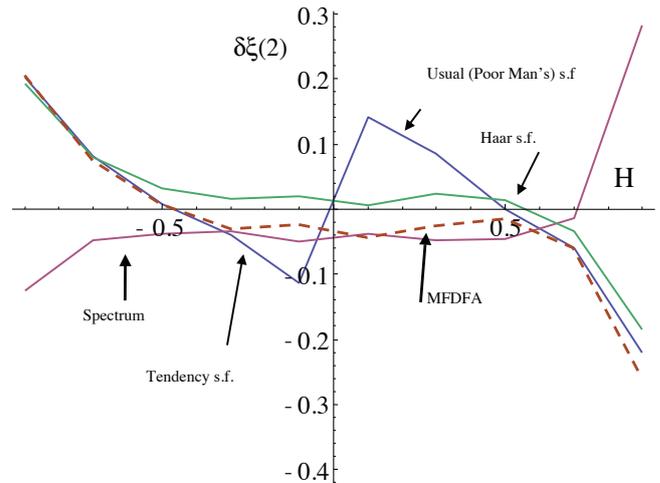
The following tests were all made on simulations with parameters  $\alpha = 1.8$ ,  $C_1 = 0.1$  with  $H$  in the range  $-7/10 < H < 7/10$  (typical parameters in meteorology and in solid earth geophysics; see the reviews Lovejoy and Schertzer, 2007, 2010b). These parameters yield an intermittency correction  $K(2) = 0.18$ ; 50 simulations were averaged to estimate the ensemble statistics. The simulations were made using the technique described in Schertzer and Lovejoy (1987) and refined in Lovejoy and Schertzer (2010a) for the flux  $\varphi$  (i.e.  $H = 0$ ). For the second fractional integration (to obtain the field  $v$  from  $\varphi$ ), we used a Fourier space power law filter  $k^{-H}$  (see Eq. 2). For this fractional integration, exponential cutoffs were used at high wavenumbers to avoid small-scale numerical instabilities (this is standard for differentiation; i.e. when filtering by  $k^{-H}$  with  $H < 0$ , it is equivalent to using the ‘‘Paul’’ wavelet; see Torrence and Compo, 1998). The raw series were  $2^{16}$  in length and were then degraded in resolution by factors of 4 to a length  $2^{14}$  pixels to avoid residual finite size effects in the simulation at the smallest scales. In order to avoid spurious correlations introduced by the periodicity of the simulations, the series were split in half so that the largest scale was  $2^{13}$ .



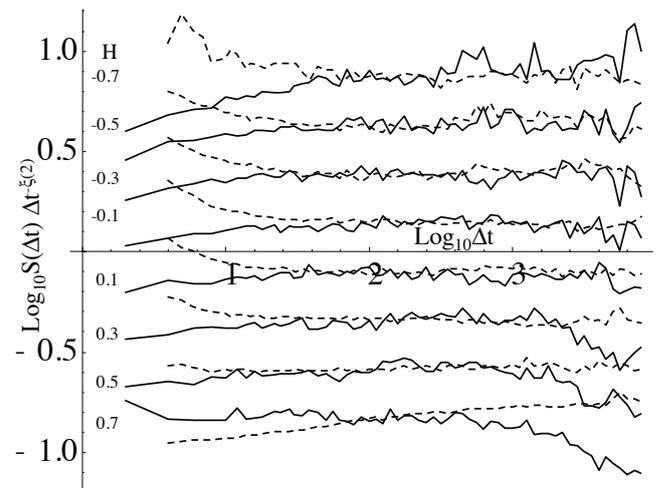
**Fig. 5.** The compensated spectra for an ensemble of 50 realizations,  $2^{14}$  each,  $\alpha = 1.8$ ,  $C_1 = 0.1$ , intermittency correction  $= K(2) = 0.18$ , with  $H$  increasing from top to bottom from  $-7/10$  to  $7/10$ ,  $\beta = 1 + 2H - K(2)$ . The dashed horizontal line is the theoretical behaviour indicated over the range used to estimate the exponent (i.e. the highest and lowest factor of  $10^{0.5}$  in wavenumber have been dropped). Each curve was offset in the vertical for clarity.

In Fig. 1 we already showed samples of the simulations visually showing the effect of increasing  $H$  from  $-7/10$  to  $7/10$ . In Fig. 5, we show the compensated spectra (obtained using standard Hann windows to avoid spectral leakage). As expected, we see that they are nearly flat, although for the lowest and highest factors of about  $10^{0.5} \approx 3$ , there were more significant deviations. The slopes of the central portions were thus determined by regression, and the corresponding exponents are shown in Fig. 6. In Fig. 5, the spectra were averaged over logarithmically spaced bins, 10 per order of magnitude (with the exception of the lowest factor of 10 where all the values were used). We also performed the regressions using log-log fits using all the Fourier components (rather than just the averaged values) in the same central range; these gave virtually identical exponent estimates (the mean absolute differences in the estimated spectral exponent  $\beta$  were  $\approx \pm 0.007$ ), and were thus not shown. Similarly, if instead of using log-log linear regression, we use nonlinear regressions over the same range, we obtain identical exponents to within  $\pm 0.009$ , and again these estimates were too similar to be worth showing. From Fig. 6, one can see that there seems to be a residual small bias of about  $-0.04$ , the origin of which is not clear.

We next consider the Haar structure function which we compare to the second-order ( $q = 2$ ), quadratic ( $n = 2$ ) MFDFA structure function analogue (based on the MFDFA fluctuation  $(\Delta v)_{\text{MFDFA}} = F/\Delta x$  where  $F$  is the usual MFDFA scaling function; see Eq. 18). The former is valid over the range  $-1 < H < 1$ , the latter over the range  $-1 < H < 2$ . We can see (Fig. 7) that the Haar structure

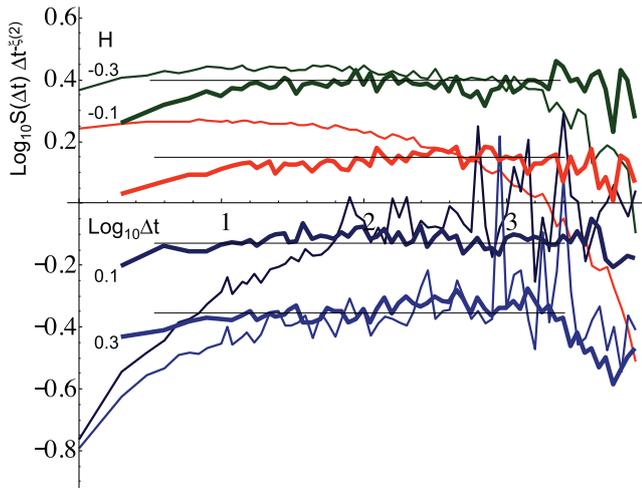


**Fig. 6.** This shows the regression estimates of the compensated exponents for the spectra  $\delta\xi(2) = \xi(2)_{\text{numerics}} - \xi(2)_{\text{theory}}$  (red; perfect methods give  $\delta\xi(2) = 0$ ), Haar structure function ( $q = 2$ ; green), quadratic,  $q = 2$  MFDFA (dashed), the usual difference (poor man's) structure function ( $q = 2$ , blue, for  $H > 0$ ), and the tendency structure function ( $q = 2$ , same line, blue for  $H < 0$ ).



**Fig. 7.** This shows the compensated Haar structure functions  $S_{\text{Haar}}(\Delta x) = \langle \Delta v_{\text{Haar}}(\Delta x)^2 \rangle^{1/2}$  (solid lines) and the second-order MFDFA technique (dashed lines), again for  $H = -7/10$  (top) to  $H = 7/10$  (bottom), every  $1/5$ . Each curve was offset in the vertical for clarity by  $1/4$ . Flat regions correspond to the theoretically predicted behaviours.

function does an excellent job with overall bias mostly around  $+0.01$  to  $0.02$ , (Fig. 6), and that the MFDFA method is nearly as good (overall bias  $\approx -0.02$  to  $-0.04$ ). It is interesting to compare this in the same figure with the results of the tendency structure function ( $H < 0$ ) and the usual difference (poor man's) structure function ( $H > 0$ ), (Fig. 8). The tendency structure function turns out to be quite accurate,



**Fig. 8.** This compares the compensated Haar structure function (thick), the difference structure function (thin, below the axis, for  $H = 3/10$  (bottom) and  $1/10$  black, second up), and the tendency structure function (third from bottom, thin red,  $H = -1/10$ ), and top, thin green,  $H = -3/10$ ). It can be seen that the standard difference structure function has poor scaling for nearly two orders of magnitude when  $H = 1/10$ , and one order of magnitude for  $H = 3/10$  (see Fig. 6 for quantitative estimates).

except near the limiting value  $H = 0$  with the large scales showing the largest deviations. In comparison, for the difference structure function, the scaling is poorest at the smaller scales requiring a range factor of  $\approx 100$  convergence for  $H = 1/10$  and a factor  $\approx 10$  for  $H = 3/10$ . The overall regression estimates (Fig. 6) show that the biases for both increase near their limiting values  $H = 0$  to accuracies  $\approx \pm 0.1$  (note that they remain quite accurate if one only makes regressions around the scaling part). This can be understood since the process can be thought of as a statistical superposition of singularities of various orders, and only those singularities with large enough orders (difference structure functions) or small enough orders (tendency structure functions) will not be affected by the theoretical limit  $H = 0$ . The main advantage of the Haar structure function with respect to the usual (tendency or difference) structure functions is precisely that it is valid over the whole range  $-1 < H < 1$ , so that it is not sensitive to the  $H = 0$  limit.

Before proceeding, we could make a general practical remark about these real space statistics: most lags  $\Delta x$  do not divide the length of the series ( $L$ ) exactly so that there is a “remainder” part. This is not important for the small  $\Delta x \ll L$  since  $f$ . For each realization, there will be many disjoint intervals of length  $\Delta x$ . However, when  $L/3 \approx \Delta x < L$ , the statistics can be sensitive to this since, from each realization, there will only be one or two segments and hence poor statistics. A simple expedient is to repeat the analysis on the reversed series and average the two results. This can indeed improve the statistics at large  $\Delta x$  when only one or a small

number of realizations are available; it has been done in the analyses presented here.

**5.2 The theoretical relation between poor man’s, tendency and Haar fluctuations, hybrid structure functions**

We have seen that, although the difference and tendency structure functions have the advantage of having simple interpretations in terms of respectively the average changes in the value of the process and its mean value over an interval, this simple interpretation is only valid over a limited range of  $H$  values. In comparison, the Haar and MF DFA fluctuations give structure functions with valid scaling exponents over wider ranges of  $H$ . But what about their interpretations? We now briefly show how to relate the Haar, tendency and poor man’s structure functions thus giving it a simple interpretation as well.

In order to see the connection between the fluctuations, we use the “saturation” relations:

$$\begin{aligned} \delta_{\Delta x} v &\stackrel{d}{=} C_{\text{tend}} v; H < 0 \\ \mathcal{T}_{\Delta x} v &\stackrel{d}{=} C_{\text{diff}} v; H > 0 \end{aligned} \tag{22}$$

where  $C_{\text{tend}}$  and  $C_{\text{diff}}$  are proportionality constants and  $\stackrel{d}{=}$  indicates equality in the random variables in the sense of probability distributions ( $a \stackrel{d}{=} b$  if  $\Pr(a > \zeta) = \Pr(b > \zeta)$  where “Pr” means “probability” and  $\zeta$  is an arbitrary threshold). These relations were easily verified numerically and arise for the reasons stated above; they simply mean that, for series with  $H < 0$ , the differences are typically of the same order as the function itself and that for  $H > 0$ , the tendencies are of the same order: the fluctuations “saturate”. By using Eqs. (7) and (22), we now obtain

$$\begin{aligned} \mathcal{H}_{\Delta x} v &= 2\mathcal{T}_{\Delta x/2} \delta_{\Delta x/2} v \stackrel{d}{=} 2\mathcal{T}_{\Delta x/2} v \stackrel{d}{=} C'_{\text{tend}} \mathcal{T}_{\Delta x} v; H < 0 \\ \mathcal{H}_{\Delta x} v &= 2\delta_{\Delta x/2} \mathcal{T}_{\Delta x/2} v \stackrel{d}{=} 2\delta_{\Delta x/2} v \stackrel{d}{=} C'_{\text{diff}} \delta_{\Delta x} v; H > 0 \end{aligned} \tag{23}$$

where  $C'_{\text{tend}}$  and  $C'_{\text{diff}}$  are “calibration” constants (only a little different from the unprimed quantities – they take into account the factor of two and the change from  $\Delta x/2$  to  $\Delta x$ ). Taking the  $q$ -th moments of both sides of Eq. (23), we obtain the results for the various  $q$ -th order structure functions:

$$\begin{aligned} \frac{\langle (\Delta v)_{\text{Haar}}^q \rangle}{\langle (\Delta v)_{\text{diff}}^q \rangle} &= C'^q_{\text{diff}}; H > 0 \\ \frac{\langle (\Delta v)_{\text{Haar}}^q \rangle}{\langle (\Delta v)_{\text{tend}}^q \rangle} &= C'^q_{\text{tend}}; H < 0 \end{aligned} \tag{24}$$

This shows that, at least for scaling processes, the Haar structure functions will be the same as the difference ( $H > 0$ ) and tendency structure functions ( $H < 0$ ), as long as these are “calibrated” by determining  $C'_{\text{diff}}$ , and  $C'_{\text{tend}}$ . In other words, by comparing the Haar structure functions with the usual structure functions, we can develop useful correction factors which will enable us to deduce the usual fluctuations

given the Haar fluctuations. Although the MF DFA fluctuations are not wavelet coefficients, at least for scaling processes, the same basic argument applies. Since the MF DFA and usual fluctuations scale with the same exponents, they can only differ in their prefactors, the calibration constants.

To see how this works on a scaling process, we confined ourselves to considering the root-mean-square (RMS)  $S(\Delta t) = \langle \Delta v^2 \rangle^{1/2}$ ; see Fig. 9 which show the ratio of the usual  $S(\Delta t)$  to those of the RMS Haar and MF DFA fluctuations. We see that, at small  $\Delta x, s$ , the ratio stabilizes after a range of about a factor 10–20 in scale and that it is quite constant up to the extreme factor of 3 or so (depending a little on the  $H$  value). The deviations are largely due to the slower convergence of the usual RMS fluctuations to their asymptotic scaling form. From the figure, we can see that the ratios  $C'_{\text{Haar}} = S/S_{\text{Haar}}$  (Fig. 9 upper left) and corresponding  $C'_{\text{MF DFA}} = S/S_{\text{MF DFA}}$  (Fig. 9, upper right) are well defined in the central region, and are near unity, more precisely in the range  $1/2 < C'_{\text{Haar}} < 2$  for the most commonly encountered range of  $H$ :  $-4/10 < H < 4/10$ . In comparison (Fig. 9, bottom) over the same range of  $H$ , we have  $0.23 > C'_{\text{MF DFA}} > 0.025$  so that the MF DFA fluctuation is quite far from the usual ones and varies strongly with  $H$ . For applications, one may use the semi-empirical formulae  $C'_{\text{Haar}} \approx 1.1e^{-1.65H}$  and  $C'_{\text{MF DFA}} \approx 0.075e^{-2.75H}$  which are quite accurate over the entire range  $-7/10 < H < 7/10$ . Although these factors in principle allow one to deduce the usual RMS structure function statistics from the MF DFA and Haar structure functions, this is only true in the scaling regime. Since the corrections for Haar fluctuations are close to unity, for many purposes they can be used directly.

### 5.3 Hybrid fluctuations and structure functions: an empirical example with both $H < 0$ , and $H > 0$ regimes

For pure scaling functions, the difference ( $1 > H > 0$ ) or tendency ( $-1 < H < 0$ ) structure functions are adequate; the real advantage of the Haar structure function is apparent for functions with two or more scaling regimes: one with  $H > 0$ , one with  $H < 0$ . Can we still “calibrate” the Haar structure function so that the amplitude of typical fluctuations can still be easily interpreted? For these, consider Eq. (23) which motivates the definition,

$$\mathcal{H}_{\text{hybrid}, \Delta x} v = \max(\delta_{\Delta x} v, \mathcal{T}_{\Delta x} v), \quad (25)$$

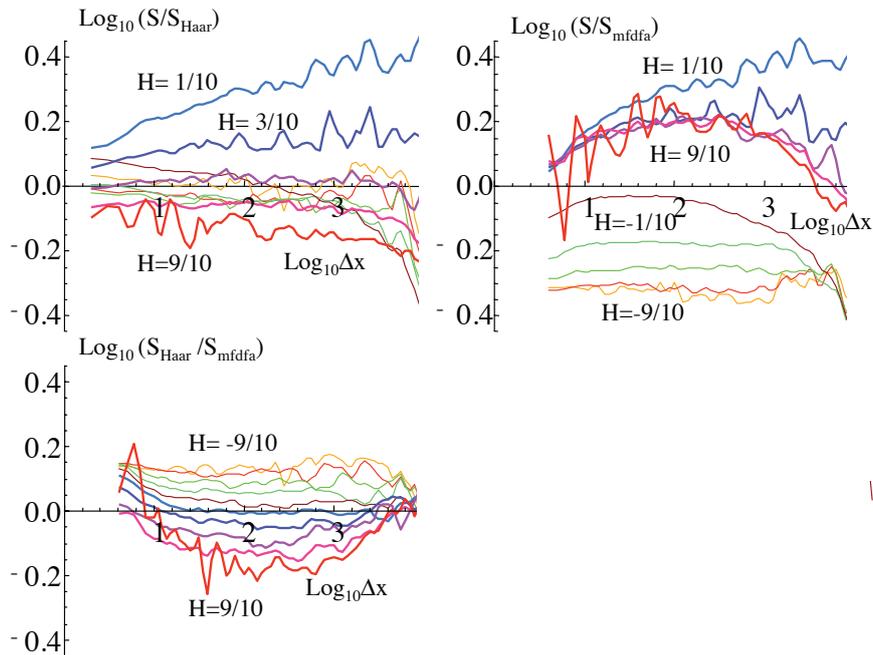
of a “hybrid” fluctuation as the maximum of the difference and tendency fluctuations; the “hybrid structure function” is thus the maximum of the corresponding difference and tendency structure functions and therefore has a straightforward interpretation. The hybrid fluctuation is useful if a unique calibration constant  $C_{\text{hybrid}}$  can be found:

$$\mathcal{H}_{\text{hybrid}, \Delta x} v \approx C_{\text{hybrid}} \mathcal{H}_{\Delta x} v. \quad (26)$$

To get an idea of how the different methods can work on real data, we consider the example of the global monthly averaged surface temperature of the earth. This is obviously highly significant for the climate, but the instrumental temperature estimates are only known over a small number of years. For this purpose, we chose three different series whose period of overlap was 1880–2008; 129 yr. These were the NOAA NCDC (National Climatic Data Center) merged land air and sea surface temperature data set (abbreviated NOAA NCDC below, from 1880 on a  $5^\circ \times 5^\circ$  grid; see Smith et al., 2008 for details), the NASA GISS (Goddard Institute for Space Studies) data set (from 1880 on a  $2^\circ \times 2^\circ$  Hansen et al., 2010) and the HadCRUT3 data set (from 1850 to 2010 on a  $5^\circ \times 5^\circ$  grid). HadCRUT3 is a merged product created out of the HadSST2 (Rayner et al., 2006) Sea Surface Temperature (SST) data set and its companion data set CRUTEM3 of atmospheric temperatures over land, see Brohan et al. (2006). Both the NOAA NCDC and the NASA GISS data were taken from <http://www.esrl.noaa.gov/psd/>, the others from <http://www.cru.uea.ac.uk/cru/data/temperature/>. The NOAA NCDC and NASA GISS are both heavily based on the Global Historical Climatology Network (Peterson and Vose, 1997), and have many similarities including the use of sophisticated statistical methods to smooth and reduce noise. In contrast, the HadCRUTM3 data are less processed with corresponding advantages and disadvantages. We could note that detailed analysis has shown that in the  $H < 0$  “macroweather” regime ( $\approx 10$  yr in Fig. 10), the intermittency is typically low; i.e.  $C_1 \approx 0.01$  (Lovejoy and Schertzer, 2012b), whereas it is  $\approx 0.07$  in the lower frequency (climate) regime which is a value comparable to that in the weather regime, i.e. of turbulence.

Figure 10 shows the comparison of the difference, tendency, hybrid and Haar RMS structure functions; the latter increased by a factor  $C_{\text{hybrid}} = 10^{0.35} \approx 2.2$ . It can be seen that the hybrid structure function does extremely well; the deviation of the calibrated Haar structure function from the hybrid one is  $\pm 14\%$  over the entire range of near a factor  $10^3$  in time scale. This shows that, to a good approximation, the Haar structure function can preserve the simple interpretation of the difference and tendency structure functions: in regions where the logarithmic slope is between  $-1$  and  $0$ , it approximates the tendency structure function, whereas in regions where the logarithmic slope is between  $0$  and  $1$ , the calibrated Haar structure function approximates the difference structure function.

Before embracing the Haar structure function, let us consider its behaviour in the presence of non-scaling perturbations. It is common to consider the sensitivity of statistical scaling analyses to the presence of non-scaling external trends superposed on the data which therefore break the overall scaling. Even when there is no reason to suspect such trends, the desire to filter them out is commonly invoked to justify the use of quadratic MF DFA or high-order wavelet techniques that eliminate linear or higher-order polynomial



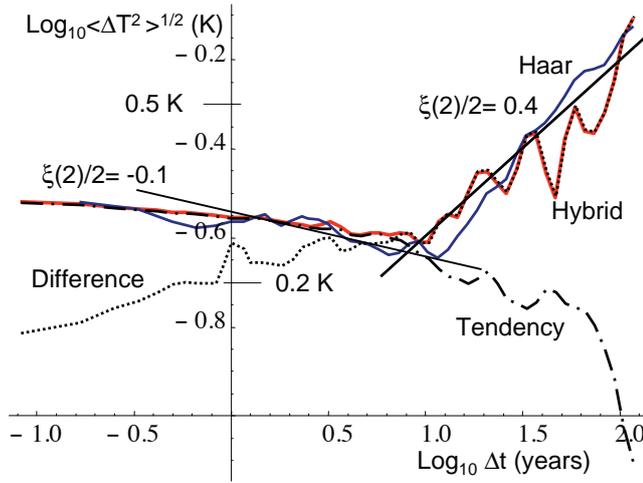
**Fig. 9.** These graphs compare the ratios of RMS structure functions ( $S(\Delta x) = \langle \Delta v^2 \rangle^{1/2}$ ) for the simulations discussed in the text. Upper left shows the ratio of the RMS usual structure function (i.e. difference when  $H > 0$ , tendency when  $H < 0$ ) to the Haar structure function for  $H = 1/10, 3/10, \dots, 9/10$  (thick, top to bottom), and  $-9/10, -7/10, \dots, -1/10$  (thin, top to bottom). The Haar structure function has been multiplied by  $2^{\xi(2)/2}$  to account for the difference in effective resolution. The central flat regions where the scaling is accurate indicate a constant ratio  $C'_{\text{Haar}} = S/S_{\text{Haar}}$  which is typically less than a factor of two. Upper right is the same but for the ratio of the usual to MFDFA RMS structure functions after the MFDFA was normalized by a factor of 16 and the resolution correction  $4^{\xi(2)/2}$  (its smallest scale is 4 pixels). Bottom left shows the ratio of the MFDFA structure function to the Haar structure function both with the normalization and resolution corrections indicated above (note the monotonic ordering with respect to  $H$ , which increases from top to bottom).

trends. However, for this purpose, these techniques are not obviously appropriate since, on the one hand, they only filter out polynomial trends (and not for example the more geophysically relevant periodic trends), while, on the other hand, even for this, they are “overkill” since the trends they filter are filtered at all scales – not just the largest. The drawback is that, with these higher-order fluctuations, we lose the simplicity of interpretation of the Haar fluctuation while obtaining few advantages. Figure 11 shows the usual (linear) Haar RMS structure function compared to the quadratic Haar and quadratic MFDFA structure functions. It can be seen that, while the latter two are close to each other (after applying different calibration constants; see the figure caption), and that the low and high frequency exponents are roughly the same, the transition time scale has shifted by a factor of about 3 so that overall they are quite different from the Haar structure function. It is therefore not possible to simultaneously calibrate the high and low frequencies.

If there are indeed external trends that perturb the scaling, these will only exist at the largest scales; it is sufficient to remove a straight line (or if needed second-, third- or higher-order polynomials) from the entire data set (i.e. at the largest scales only). Figure 12 shows the result when the

Haar structure function is applied to scaling data that have been detrended in two slightly different ways: by removing a straight line through the first and last points of the series; and by removing a regression line. Since these changes essentially affect the low frequencies only, they mostly affect the extreme factor of two in scale. If we discount this extreme factor 2, then in the figure we see that there are again two scaling regimes; the low frequency one is slightly displaced with respect to the Haar structure function, but not nearly as much as for the quadratic Haar. In other words, removing the linear trends at all scales as in the quadratic Haar or the quadratic MFDFA is too strong to allow a simple interpretation of the result and it is unnecessary if one wishes to eliminate external trends.

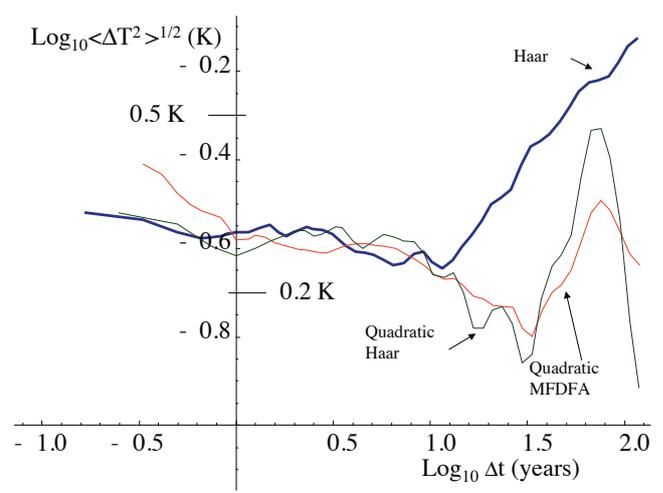
In summary, if all that is required are the scaling exponents, the Haar structure function and quadratic MFDFA seem to be excellent techniques. However, the Haar structure function has the advantage of numerical efficiency, simplicity of implementation, and simplicity of interpretation.



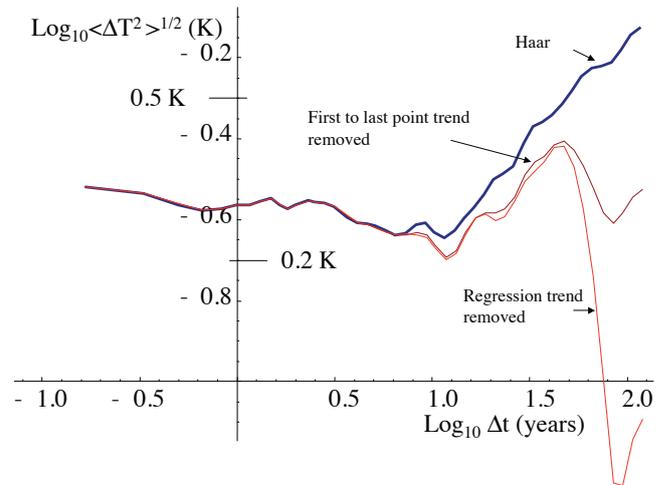
**Fig. 10.** A comparison of the different structure function analyses (root-mean-square, RMS) applied to the ensemble of three monthly surface series (NASA GISS, NOAA CDC, HADCRUT3; see Lovejoy and Schertzer, 2012a), each globally and annually averaged, from 1880–2008 (1548 points each). The usual (difference, poor man’s) structure function is shown (dotted, lower left), the tendency structure function (dot-dashed, lower right), the maximum of the two (“hybrid”, thick, red), and the Haar in blue (as indicated); it has been increased by a factor  $C = 10^{0.35} = 2.2$ , and the RMS deviation with respect to the hybrid is  $\pm 14\%$ . Reference slopes with exponents  $\xi(2)/2 \approx 0.4, -0.1$  are also shown (black corresponding to spectral exponents  $\beta = 1 + \xi(2) = 1.8, 0.8$ , respectively). In terms of difference fluctuations, we can use the global root-mean-square  $\langle \Delta T(\Delta t)^2 \rangle^{1/2}$  annual structure functions (fitted for  $129 \text{ yr} > \Delta t > 10 \text{ yr}$ ), obtaining  $\langle \Delta T(\Delta t)^2 \rangle^{1/2} \approx 0.08 \Delta t^{0.33}$  for the ensemble. In comparison, Lovejoy and Schertzer (1986) found the very similar  $\langle \Delta T(\Delta t)^2 \rangle^{1/2} \approx 0.077 \Delta t^{0.4}$  using Northern Hemisphere data (these correspond to  $\beta_c = 1.66, 1.8$  respectively). Reproduced from Lovejoy and Schertzer (2012a).

**6 Conclusions**

Geosystems are commonly scaling over large ranges of scale in both space and time. In order to characterize their properties, one decomposes them into elementary components with well-defined scales and characterizes the variability with the help of various statistics. Their scaling behaviours can then be characterized by exponents that are obtained by systematically changing the scale. The classical analysis techniques are spectral analysis (Fourier decomposition) and structure functions based on fluctuations defined as differences. However, over the last 20 yr, the development of wavelets has provided a mathematical basis for defining fluctuations. As far as geophysical applications are concerned, this generality has not proved to be so helpful. This is because different workers use different types of wavelets and these are often chosen for various mathematical properties which are not necessarily



**Fig. 11.** The same temperature data as Fig. 10: a comparison of the RMS Haar structure function (multiplied by  $10^{0.35} = 2.2$ ), the RMS quadratic Haar (multiplied by  $10^{0.15} = 1.4$ ) and the RMS quadratic MF DFA (multiplied by  $10^{1.5} = 31.6$ ). Reproduced from Lovejoy and Schertzer (2012a).



**Fig. 12.** The same temperature data as Fig. 10: a comparison of the RMS Haar structure function applied to the raw data, after removing linear trends in two ways. The first defines the linear trend by the first and last points, while the second uses a linear regression; all were multiplied by  $C_{\text{hybrid}} = 10^{0.35} = 2.2$ .

important for applications. In addition, the numerical determination of the wavelet coefficients is often cumbersome. Finally, and most importantly, the interpretation of the corresponding wavelet coefficients is often obscure and attention has been almost exclusively focused on the exponents.

These features of wavelet analysis have led to the popularity of a slightly different wavelet-like technique – the detrended fluctuation analysis (DFA) method and its extension, the multifractal DFA (MF DFA). An unfortunate consequence

is that the criterion for the applicability of the different techniques has largely been ignored (the range of  $H$  exponents), and the advantages and disadvantages of different approaches have not been critically (and quantitatively) evaluated. In this paper, we argue that the historically first wavelet – the Haar wavelet – has several advantages. Ironically, these advantages are not related to their wavelet nature; the wavelet framework is not especially helpful nor is it needed. These advantages are not so much at the level of the accuracy of the exponents – which turn out to be about the same as for spectral analysis and the DFA technique – but rather in the ease of implementation and interpretation of the results. This is demonstrated both with the help of multifractal simulations (exponents) and the ease of interpretation with an atmospheric example where we show that, with “calibration”, the Haar structure functions can be made very close to both the usual difference structure functions (in regions where  $1 > H > 0$ ) and to simple “tendency” structure functions (in regions where  $-1 < H < 0$ ). This simplicity is already lost when using either the DFA or even higher-order wavelet-based generalizations of Haar fluctuations. The Haar structure functions have recently been applied systematically to atmospheric data where they have contributed to clarifying the fact that there are not two regimes (weather and climate) but rather three: weather ( $\lesssim 10$  day); macroweather (between  $\approx 10$  days and  $\approx 10\text{--}30$  yr,  $H < 0$ ); and the climate ( $\gtrsim 10\text{--}30$  yr,  $H > 0$ ) (Lovejoy and Schertzer, 2012a) (see <http://www.physics.mcgill.ca/> for Haar software (Mathematica and MathLab)).

### Appendix A

#### Relations between real and Fourier space analyses and the limiting $H$ exponents

We repeatedly underlined the limiting  $H$  values associated with specific types of fluctuations. In this appendix we clarify the origin of these limits by recalling the classical relation between the fluctuations and the spectra.

Consider the statistically translationally invariant process  $v(x)$  in 1-D: the statistics are thus independent of  $x$  and this implies that the Fourier components are “ $\delta$  correlated”:

$$\langle \tilde{v}(k) \tilde{v}(k') \rangle = \delta_{k+k'} \langle |\tilde{v}(k)|^2 \rangle; \quad \tilde{v}(k) = \int e^{-ikx} v(x) dx \quad (A1)$$

If it is also scaling, then the spectrum  $E(k)$  is a power law:  $E(k) \approx \langle |\tilde{v}(k)|^2 \rangle \approx k^{-\beta}$  (where here and below, we ignore constant terms such as factors of  $2\pi$  etc.). Let us first consider the difference fluctuation. In terms of its Fourier components, this fluctuation is thus

$$\begin{aligned} (\Delta v(x, \Delta x))_{\text{diff}} &= v(x + \Delta x) - v(x) \\ &= \int e^{ikx} \tilde{v}(k) (e^{ik\Delta x} - 1) dk \end{aligned} \quad (A2)$$

so that the F.T. of  $(\Delta v(x, \Delta x))_{\text{diff}}$  is  $\tilde{v}(k) (e^{ik\Delta x} - 1)$ . We first consider the statistics of quasi-Gaussian processes for which  $C_1 = 0$ ,  $\xi(q) = Hq$ . Assuming statistical translational invariance, we drop the  $x$  dependence and obtain the relation of second-order structure function to the spectrum:

$$\begin{aligned} \langle |\Delta v(\Delta x)|^2 \rangle &= 4 \int e^{ikx} \langle |\tilde{v}(k)|^2 \rangle \sin^2\left(\frac{k\Delta x}{2}\right) dk \\ &\approx \int e^{ikx} k^{-\beta} \sin^2\left(\frac{k\Delta x}{2}\right) dk. \end{aligned} \quad (A3)$$

As long as the integral on the right converges, then using the transformation of variables,  $x \rightarrow \lambda x$ ,  $k \rightarrow \lambda^{-1}k$ , we find

$$\begin{aligned} \langle |\Delta v(\Delta x)|^2 \rangle &\propto \Delta x^{\xi(2)} \\ &\approx \int e^{ikx} k^{-\beta} \sin^2\left(\frac{k\Delta x}{2}\right) dk \propto \Delta x^{\beta-1} \end{aligned} \quad (A4)$$

so that  $\beta = \xi(2) + 1 = 2H + 1$  ( $C_1 = K(q) = 0$  here, Eq. 21). However, for large  $k$ , the integrand  $\approx k^{-\beta}$ , which has a large wavenumber divergence whenever  $\beta \leq 1$ . However, since for small  $k$ ,  $\sin^2(k\Delta x/2) \propto k^2$ , there will be a low wavenumber divergence only when  $\beta \geq 3$ . Although real world (finite) data will not diverge, in the theoretically divergent cases, the structure functions will no longer characterize the local fluctuations, but rather those on either the highest or lowest wavenumbers present in the data. In the quasi-Gaussian case, or when  $C_1$  is small, we have  $\xi(2) \approx 2H$  and we conclude that the use of first-order differences to define the fluctuations leads to second-order structure functions being meaningful in the sense that they adequately characterize the fluctuations whenever  $1 < \beta < 3$  (i.e.  $0 < H < 1$ ).

Since  $0 < H < 1$  is the usual range of geophysical  $H$  values, and the difference fluctuations are very simple, they are commonly used. However, we can see that there are limitations; in order to extend the range of  $H$  values, it suffices to define fluctuations using finite differences of different orders. To see how this works, consider using second (centred) differences:

$$\begin{aligned} (\Delta v(x, \Delta x))_{\text{second}} &= v(x) - \frac{1}{2}(v(x + \Delta x/2) - v(x - \Delta x/2)) \\ &= \int e^{ikx} \tilde{v}(k) \left[ 1 - \frac{1}{2}(e^{ik\Delta x/2} + e^{-ik\Delta x/2}) \right] dk \\ &= 2 \int e^{ikx} \tilde{v}(k) \sin\left(\frac{k\Delta x}{4}\right) dk. \end{aligned} \quad (A5)$$

Repeating the above arguments, we can see that the relation  $\beta = \xi(2) + 1$  holds now for  $1 < \beta < 5$ , or (with the same approximation)  $0 < H < 2$ . Similarly, by replacing the original series by its running sum (a finite difference of order  $-1$ ) – as done in the MFDFA technique – we can extend the range of  $H$  values down to  $-1$ .

More generally, going beyond Gaussian processes, we can consider intermittent fractionally

integrated flux processes (Sect. 5.1), which have  $\tilde{v}(k) \approx \tilde{\varphi}(k) |k|^{-H}$ . We see that the F.T. of  $(\Delta v(x, \Delta x))_{\text{diff}}$  is  $(e^{ik\Delta x} - 1) |k|^{-H} \tilde{\varphi}(k)$ . This implies that, for low wavenumbers,  $\tilde{v}(k) \approx |k|^{1-H} \tilde{\varphi}(k)$  ( $k \ll 1/\Delta x$ ), whereas at high wavenumbers,  $\tilde{v}(k) \approx |k|^{-H} \tilde{\varphi}(k)$  ( $k \gg 1/\Delta x$ ). Hence, since the mean of  $\varphi$  has no scale dependence, we find that, for  $0 < H < 1$ , the fluctuations  $\Delta v(\Delta x)$  are dominated by wavenumbers  $k \approx 1/\Delta x$ . So for this range of  $H$ , fluctuations defined as differences capture the variability of  $\Delta x$ -sized structures, not structures much smaller or much larger than  $\Delta x$ . More generally, since the F.T. of the  $n$ -th derivative  $d^n v/dx^n$  is  $(ik)^n \tilde{v}(k)$  and the finite derivative is the same for small  $k$  but “cut-off” at large  $k$ , we find that  $n$ -th order fluctuations are dominated by structures with  $k \approx 1/\Delta x$  as long as  $0 < H < n$ . This means that  $\Delta v(\Delta x)$  does indeed reflect the  $\Delta x$  scale fluctuations.

*Acknowledgements.* We thank A. Bunde for helpful correspondence on the MFDFA technique.

Edited by: N. S. Erokhin

Reviewed by: four anonymous referees

## References

- Arneodo, A., Decoster, N., and Roux, S. G.: Intermittency, Log-Normal Statistics, And Multifractal Cascade Process In High-Resolution Satellite Images Of Cloud Structure, *Phys. Rev. Lett.*, 83, 1255–1258, 1999.
- Ashok, V., Balakumaran, T., Gowrishankar, C., Vennila, I. L. A., and Nirmalkumar, A.: The Fast Haar Wavelet Transform for Signal & Image Processing, *Int. J. Comp. Sci. Info.*, 7, 126–130, 2010.
- Bacry, A., Arneodo, A., Frisch, U., Gagne, Y., and Hopfinger, E.: Wavelet analysis of fully developed turbulence data and measurement of scaling exponents, in: *Turbulence and coherent structures*, edited by: Lessieur, M. and Metais, O., Kluwer, 703–718, 1989.
- Brohan, P., Kennedy, J. J., Harris, I., Tett, S. F. B., and Jones, P. D.: Uncertainty estimates in regional and global observed temperature changes: a new dataset from 1850, *J. Geophys. Res.*, 111, D12106, doi:10.1029/2005JD006548, 2006.
- Bunde, A., Kropp, J., and Schellnhuber, H. J. (Eds.): *The science of disasters: Climate disruptions, heart attacks and market crashes*, Springer, 2002.
- 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, 048701, doi:10.1103/PhysRevLett.94.048701, 2005.
- Davis, A., Marshak, A., Wiscombe, W., and Cahalan, R.: Multifractal characterization of intermittency in nonstationary geophysical signals and fields, in: *Current topics in nonstationary analysis*, edited by: Harding, J., Douglas, B., and Andreas, E., World Scientific, Singapore, 97–158, 1996.
- Haar, A.: *Zur Theorie des orthogonalen Funktionensysteme*, *Mathematische Annalen*, 69, 331–371, 1910.
- Hansen, J., Ruedy, R., Sato, M., and Lo, K.: Global surface temperature change, *Rev. Geophys.*, 48, RG4004, doi:10.1029/2010RG000345, 2010.
- Holschneider, M.: *Wavelets: an analysis tool*, Oxford, Clarendon Press, New York, 1995.
- Huang, Y. X., Schmitt, F. G., Hermand, J.-P., Gagne, Y., Lu, Z. M., and Liu, Y. L.: Arbitrary-order Hilbert spectral analysis for time series possessing scaling statistics, Comparison study with detrended fluctuation analysis and wavelet leaders, *Phys. Rev. E*, 84, 016208, doi:10.1103/PhysRevE.84.016208, 2011.
- Kantelhardt, J. W., Koscielny-Bunde, E., Rego, H. H. A., Havlin, S., and Bunde, S.: Detecting long range correlations with detrended fluctuation analysis, *Physica A*, 295, 441–454, 2001.
- Kantelhardt, J. W., Zschiegner, S. A., Koscielny-Bunde, K., Havlin, S., Bunde, A., and Stanley, H. E.: Multifractal detrended fluctuation analysis of nonstationary time series, *Physica A*, 316, 87–114, 2002.
- 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.
- Koscielny-Bunde, E., Kantelhardt, J. W., Braund, P., Bunde, A., and Havlin, S.: Long-term persistence and multifractality of river runoff records: Detrended fluctuation studies, *J. Hydrol.*, 322, 120–137, 2006.
- Lennartz, S. and Bunde, A.: Trend evaluation in records with long term memory: Application to global warming, *Geophys. Res. Lett.*, 36, L16706, doi:10.1029/2009GL039516, 2009.
- Lovejoy, S. and Schertzer, D.: Scale invariance in climatological temperatures and the spectral plateau, *Ann. Geophys.*, 4B, 401–410, 1986.
- Lovejoy, S. and Schertzer, D.: Scaling and multifractal fields in the solid earth and topography, *Nonlin. Processes Geophys.*, 14, 465–502, doi:10.5194/npg-14-465-2007, 2007.
- Lovejoy, S. and Schertzer, D.: On the simulation of continuous in scale universal multifractals, part I: spatially continuous processes, *Comput. Geosci.*, 36, 1393–1403, doi:10.1016/j.cageo.2010.04.010, 2010a.
- 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, 2010b.
- Lovejoy, S. and Schertzer, D.: Low frequency weather and the emergence of the Climate, in: *Complexity and Extreme Events in Geosciences*, edited by: Sharma, A. S., Bunde, A., Baker, D., and Dimri, V. P., AGU monographs, 2012a.
- Lovejoy, S. and Schertzer, D.: *The Weather and Climate: Emergent Laws and Multifractal Cascades*, Cambridge University Press, Cambridge, 660 pp., 2012b.
- Mallat, S. and Hwang, W.: Singularity detection and processing with wavelets, *IEEE Trans. Info. Theory*, 38, 617–643, 1992.
- Monin, A. S. and Yaglom, A. M.: *Statistical Fluid Mechanics*, MIT press, Boston MA, 1975.
- Oswiecimka, P., Kwapien, J., and Drozd, S.: Wavelet versus detrended fluctuation analysis of multifractal structures, *Phys. Rev. E*, 74, 16103, doi:10.1103/PhysRevE.74.016103, 2006.
- 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.

- Peterson, T. C. and Vose, R. S.: An overview of the Global Historical Climatology Network temperature database, *B. Am. Meteorol. Soc.*, 78, 2837–2849, 1997.
- Rayner, N. A., Brohan, P., Parker, D. E., Folland, C. K., Kennedy, J. J., Vanicek, M., Ansell, T., and Tett, S. F. B.: Improved analyses of changes and uncertainties in marine temperature measured in situ since the mid-nineteenth century: the HadSST2 dataset, *J. Climate*, 19, 446–469, 2006.
- 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.
- Schertzer, D. and Lovejoy, S.: Hard and Soft Multifractal processes: *Physica A*, 185, 187–194, 1992.
- Schertzer, D. and Lovejoy, S.: Universal Multifractals do Exist!, *J. Appl. Meteor.*, 36, 1296–1303, 1997.
- Schertzer, D., Lovejoy, S., and Hubert, P.: An introduction to stochastic multifractal fields, *Mathematical Problems in Environmental Science and Engineering*, edited by: Ern, A. and Weiping, L., Series in Contemporary Applied Mathematics, vol.4, Higher Education Press, Beijing, 106–179, 2002.
- Schertzer, D., Tchiguirinskaia, I., Lovejoy, S. and Hubert, P.: No monsters, no miracles: in nonlinear sciences hydrology is not an outlier!, *Hydrol. Sci. J.*, 55, 965–979, 2010.
- Serrano, E. and Figliola, A.: Wavelet Leaders: A new method to estimate the multifractal singularity spectra, *Physica A*, 388, 2793–2805, 2009.
- Smith, T. M., Reynolds, R. W., Peterson, T. C., and Lawrimore, J.: Improvements to NOAA's Historical Merged Land-Ocean Surface Temperature Analysis (1880–2006), *J. Climate*, 21, 2283–2293, 2008.
- Taqqu, M., Teverovsky, V., and Willinger, W.: Estimators for long-range dependence: an empirical study, *Fractals*, 3, 785–798, 1995.
- Torrence, T. and Compo, G. P.: A practical guide to wavelet analysis, *B. Am. Meteorol. Soc.*, 79, 61–78, 1998.
- Veneziano, D. and Furcolo, P.: Marginal Distribution Of Stationary Multifractal Measures And Their Haar Wavelet Coefficients, *Fractals*, 11, 253–270, doi:10.1142/S0218348X03002051, 2003.