

Supplementary Information (SI)

S1: Discrete Wavelet Transform

Suppose a data signal (Y) is a vector of size $N \times 1$. The DWT of Y , denoted as d , is represented as

$$d = WY, \quad (1)$$

where W is an orthogonal matrix of size $N \times N$. The elements in W are determined by selecting a particular wavelet basis, such as Haar, Daubechies, and Symmlet. Although N can be arbitrary, it is usually selected to be a power of two, i.e., $N = 2^J$, $J \in \mathbb{Z}^+$ for the ease of calculating the DWT.

When N is large, and calculating d as in (1) becomes computationally expensive, a fast algorithm based on filtering proposed by Mallat is used for computational efficiency. With this algorithm the DWT is obtained by performing a series of successive convolutions that involve a wavelet-specific low-pass filter h and its quadrature-mirror counterpart, high-pass filter g . These repeated convolutions using the two filters accompanied with the operation of decimation (keeping every second coefficient of the convolution) generate a multiresolution representation of a signal, consisting of a smooth approximation (c -coefficients) and a hierarchy of detail coefficients d_{jk} at different resolutions (indexed by a scale index j) and different locations within the same resolution, indexed by k . The convolutions with filters h and g are repeated until a desired decomposition level $j = J_0$ is reached ($1 \leq J_0 \leq J - 1$ and $J = \log_2 N$).

Thus, the vector $\underline{d}$ in (1) has a structure,

$$\underline{d} = (\underline{c}_{J_0}, \underline{d}_{J_0}, \dots, \underline{d}_{J-2}, \underline{d}_{J-1}), \quad (2)$$

where $\underline{c}_{J_0}$ is a vector of coefficient corresponding to a smooth trend in signal, and $\underline{d}_j$ are detail coefficients at different resolutions j where $J_0 \leq j \leq J - 1$. It is the logarithm of the variance of coefficients in $\underline{d}_j$ that is used to define wavelet spectra.

S2: Monofractality

A deterministic function $f(t)$ of d -dimensional argument t is said to be self-similar, if $f(\lambda t) = \lambda^{-H} f(t)$, for some choice of the Hurst exponent H , and for all dilation factors λ . The notation of self-similarity has been extended to random processes. Specifically, a stochastic process $\{X(t), t \in \mathbb{R}^d\}$ is self-similar with a scaling parameter (or Hurst exponent H) if, for any $\lambda \in \mathbb{R}^+$,

$$X(\lambda t) \stackrel{d}{=} \lambda^H X(t), \quad (3)$$

where the relation “ $\stackrel{d}{=}$ ” denotes the equality in all finite-dimensional distributions.

The scaling exponent (H), is a statistical parameter in a well-defined statistical model when the signal follows stochastic behavior (e.g., Gaussianity and stationary increments).

Wavelet transform-based methods have been proven to be suitable for self-similar modeling processes with stationary increments. For example, models such as fractional Brownian motion (fBm) in one dimension and fractional Brownian fields (fBf) in d dimensions are frequently used to explain data that scale. For a comprehensive description, see [1].

S3: Standard Wavelet Spectra

For data that scale, wavelet transforms have also been a widely used tool for estimating Hurst exponents. While several methods are available in the wavelet domain, most estimations rely on the properties of the wavelet spectrum. The wavelet spectrum can be described theoretically as follows. Considering the self-similar process given in (3), the detail wavelet coefficients satisfy

$$d_{jk} \stackrel{d}{=} 2^{-j(H+1/2)} d_{0k}, \quad (4)$$

for a fixed multiresolution level j and under L_2 normalization. Also, if the process in (3) is a zero-mean stochastic process and possesses stationary increments (i.e., $\mathbb{E}(X) = 0$ and $X(t+h) - X(t)$ is independent of t), then $\mathbb{E}(d_{0k}) = 0$ and $\mathbb{E}(d_{0k}^2) = \mathbb{E}(d_{00}^2)$. As a result,

$$\mathbb{E}(d_{jk}^2) \propto 2^{-j(2H+1)}. \quad (5)$$

Equation (5) provides a basis for estimating the Hurst exponent H by taking logarithms of both sides. The set of pairs $(j, S(j)) = (j, \log_2 \mathbb{E}(d_{jk}^2))$, is called the wavelet spectrum. For the processes that scale, the logarithms of average squared of detail wavelet coefficients (log energies) at different resolution scales exhibit a linear relationship with respect to the scale index, and this linear relationship, in particular the slope of the linear fit, is used to find an estimator for H . More comprehensive information is available in [2].

In practice, wavelet spectrum is computed as follows. Let $\underline{d}_j = \{d_1, d_2, \dots, d_n\}$ represents detail wavelet coefficients in the j th multiresolution level of wavelet transform of a signal X . Then, the wavelet spectrum of X is computed as

$$S(j) = \log_2 \left(\overline{d_j^2} \right), \quad J_0 \leq j \leq J-1, \quad (6)$$

where $\overline{d_j^2} (= \frac{1}{n} \sum_{i=1}^n d_i^2)$ is the average of squared wavelet coefficients (also known as log-energy) in level j , $J = \log_2(N)$, N is the signal length, and J_0 ($1 \leq J_0 \leq J-1$) is the coarsest level of detail used to define the wavelet spectrum.

In order to estimate H , wavelet log-energies are regressed on the scale indexes. Based on the slope of the estimated regression model, the estimated value of H is calculated as $H = -(slope + 1)/2$.

S4: Multifractality

According to (Reidi et al., 1999) and (Reidi (2002)), a multifractal spectrum is defined by a local singularity strength measure based on wavelets:

$$\alpha(t) = \lim_{k2^j \rightarrow t} \left(-\frac{1}{j} \log_2 |d_{j,k}| \right), \quad (7)$$

where $d_{j,k}$ is the normalized wavelet coefficient corresponding to the basis that is L_1 normalized, that is $\phi_{jk}(t) = 2^j \phi(2^j t - k)$ and $\psi_{jk}(t) = 2^j \psi(2^j t - k)$ at scale j and location k . Smaller values of $\alpha(t)$ correspond to larger oscillation in the signal and thus more singularity strengths. The frequency (in t) of occurrence of a given singularity strength α is measured by the multifractal spectrum, denoted as $f(\alpha)$ analogously to Reidi (2002). That is, any inhomogeneous process has a set of local singularity measures and their distribution ($f(\alpha)$) forms the multifractal spectra.

A practical approach to compute multifractal spectrum makes use of the theory of large deviations, where f would be interpreted as the rate function in a *Large Deviation Principle* (f measures how frequently (in k) the observed $(-1/j) \log_2 |d_{j,k}|$ deviate from the expected value for α in scales defined by j). More specifically, the multifractal spectrum depends on the concepts of *partition function*, $T(q)$ and *Legendre transform*.

The partition function, $T(q)$, is defined as

$$T(q) = \lim_{j \rightarrow \infty} \left(-\frac{1}{j} \log_2 \mathbb{E}(|d_{j,k}|^q) \right), \quad (8)$$

$T(q)$ describes the limiting behavior of q th moment of a typical wavelet coefficient $d_{j,k}$ at the scale j and location k . The moment order $q \in \mathbb{R}$ is within a certain range, covering negative numbers as well. Then, the multifractal formalism points that the multifractal spectrum can be calculated by taking the *Legendre transform* of the log moment generating function:

$$f(\alpha) \simeq f_L(\alpha) := \inf_q [q\alpha - T(q)]. \quad (9)$$

It can be shown that $f(\alpha) = q\alpha - T(q)$ at $\alpha = T'(q)$ given that $T''(q) < 0$, and also $f_L(\alpha)$ converges to the true multifractal spectrum using the theory of large deviations.

To compute multifractal spectrum, we need a good estimator of the partition function. This estimator can be derived by rearranging (8) as follows.

$$\mathbb{E}|d_{j,k}|^q \sim 2^{-jT(q)} \quad \text{as } j \rightarrow \infty. \quad (10)$$

On the other hand, the q th moment of the wavelet coefficients of the power law process, holds the following expression:

$$\mathbb{E}|d_{j,k}|^q = C_q 2^{-jqH}, \quad (11)$$

where H is the irregularity index (scaling exponent) and C_q is a constant depending only on q .

Comparing (10) and (11), we can easily connect the partition function with the scaling exponent estimation procedure explained in section 3.3 in the main document. The use of linear regression has been a standard practice to identify H since the values $\mathbb{E}(|d_{j,k}|^q)$ could be easily computed by moment matching method, estimating the partition function, $T(q)$ easy. We can express this more formally as follows:

$$\log_2 \widehat{S_j(q)} = -jT(q) + \varepsilon_j, \quad (12)$$

where $\widehat{S_j(q)} = \frac{1}{2^j} \sum_{k=1}^{N2^{-j}} |d_{jk}|^q$ is the empirical q th moment of the wavelet coefficients, N is the length of the signal, and ε is the error term introduced from the moment matching method when replacing the true moment with the empirical ones. The ordinary least square is a convenient method for estimating the partition function.

After estimating the transfer function $T(q)$, the *Legendre transform* is performed on the estimated the partition function. Since $\frac{\partial}{\partial q}(\alpha q - T(q)) = \alpha - T'(q)$ and $T''(q) < 0$, the maximum value of $\alpha q - T(q)$ is achieved at $a = T'^{(-1)}(\alpha)$. This indicates that performing Legendre transformation is divided into two steps: computing numerical derivative of $T(q)$ using finite difference and then evaluating the value of Legendre spectrum at $\alpha = \widehat{T'(q)}$. It is important to note that the *Legendre transform* cannot estimate the multifractal spectrum at arbitrary singularity strength α . A set of multifractal spectrum values is determined for a range of q values. The more q values adopted, the smoother the multifractal spectrum. For example, a representative example multifractal spectrum can be found in Figure 2b.

References

1. Goncalves, P., Riedi, R. & Baraniuk, R. A simple statistical analysis of wavelet-based multifractal spectrum estimation. In *Conference Record of Thirty-Second Asilomar Conference on Signals, Systems and Computers (Cat. No.98CH36284)*, vol. 1, 287–291 vol.1 (1998).
2. Vidakovic, B. *Statistical modeling by wavelets* (John Wiley & Sons, 2009).