

Wavelet Transform Based Multifractal Formalism in Outlier Detection and Localisation for Financial Time Series

Zbigniew R. Struzik, Arno P.J.M. Siebes

CWI

P.O. Box 94079, 1090 GB Amsterdam, The Netherlands

ABSTRACT

We present a method of detecting and localising outliers in financial time series and other stochastic processes. The method checks the internal consistency of the scaling behaviour of the process within the paradigm of the multifractal spectrum. Deviation from the expected spectrum is interpreted as the potential presence of outliers. The detection part of the method is then supplemented by the localisation analysis part, using the local scaling properties of the time series. Localised outliers can then be removed one by one, with the possibility of dynamic verification of spectral properties. Both the multifractal spectrum formalism and the local scaling properties of the time series are implemented on the wavelet transform modulus maxima tree.

Keywords and Phrases: multifractal analysis, wavelet transform, Hölder exponent, outlier detection

1. INTRODUCTION

All real life stochastic signals are subject to contamination by noise. This can be relatively low level random noise or some systematic bias. But it can also be a sudden shoot off from the current value of the sensoric information - the outlier. The main difference between the noise or bias in the traditional sense and an outlier, which of course is also 'noise', is the inherently isolated and local character of the outlier. In addition, outliers will usually have high amplitudes and generally there will be relatively few of them. Therefore, non-stationarity and a highly erratic character is the main characteristic which contrasts with stationary (white) noise or linear or polynomial bias.

Financial time series, which are gathered by banks on a daily basis in large volumes (to the order of a million different indicators per day), are sometimes contaminated with outliers. All financial records are hand typed and are prone to exactly the outlier type of contamination/noise. A highly desirable possibility would be to be able to detect such errors at an early stage in the newly collected data. In a more general formulation, it is of fundamental importance to determine whether a particular extreme observation belongs to the bulk of the data or should be treated as an outlier. The issue of outliers in financial time series has recently been extensively discussed by Johansen et al [1, 2] and by L'vov et al [3].

In figure 1, we present examples of the time series obtained from a Dutch bank with suspected outliers in them. The reader may say, 'It is very easy to see outliers in these plots; one can determine them with the naked eye.' We agree. The human eye (or visual system) is an outstanding pattern recognition system, very good in detecting non-stationarities. It is capable of fitting models and performing approximations in real time using various acquired knowledge. It is, however, also an adaptive system: different runs of the same experiment will likely result in different answers. Varied magnification, resolution or normalisation will also most likely affect the results. In this paper, we would like to provide some *objective* criteria which should also be suitable to be exercised automatically by the computer.

For this purpose we will need a methodology which is capable of determining the statistical nature of the non-stationary process in both a global and a local sense. The global sense will be primarily used to detect the presence of the outliers in the process analysed. The local properties will next be used to

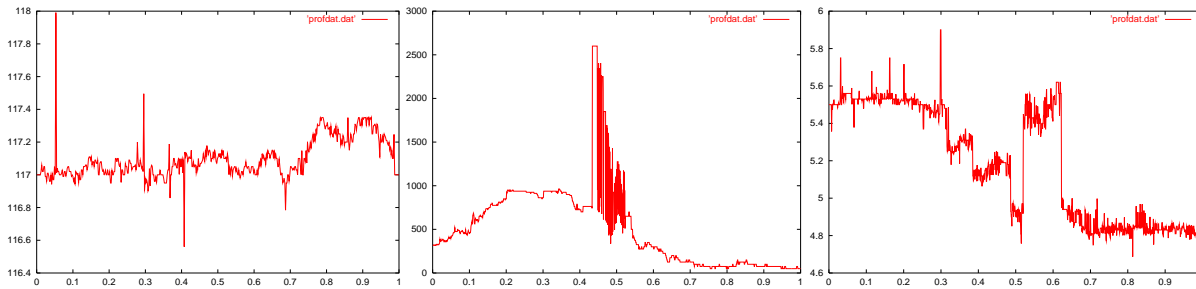


Figure 1: Example financial time series with suspected outlier presence.

identify the offenders. Multifractal formalism has recently been suggested for analysis and modelling of financial time series [4, 5, 6, 7]. It seems therefore natural to build our methodology around the generic wavelet based multifractal analysis first introduced by Arneodo, Bacry and Muzy [8, 9]. The formalism has been extensively used to test many natural phenomena and has contributed to substantial progress in several domains in which it has been applied [10, 11, 12, 13]. Nevertheless its use has been hampered by its sensitivity to outliers. Ironically, this is, of course, a highly required property in our application.

We will first introduce the wavelet transform, paying special attention to detection and analysis of singular events. Then we will briefly describe the global partition function based multifractal formalism, showing how it can be applied to test the statistical integrity of the analysed process. Next we will introduce the local version of multifractal singularity strength analysis, again showing how the statistical integrity of the analysed process can be tested, now in a local fashion, allowing not only the detection but also the localisation of outliers.

Section 5 closes the paper with conclusions and suggestions for future developments.

2. CONTINUOUS WAVELET TRANSFORM AND ITS MAXIMA USED TO REVEAL THE STRUCTURE OF SINGULARITIES IN THE TIME SERIES

Conceptually, the wavelet transformation [14, 15] is a convolution product of the time series with the scaled and translated kernel - the wavelet $\psi(x)$, usually a n -th derivative of a smoothing kernel $\theta(x)$. Usually, in the absence of other criteria, the preferred choice is the kernel, which is well localised both in frequency and position. In this paper, we chose the Gaussian $\theta(x) = \exp(-x^2/2)$ as the smoothing kernel, which has optimal localisation in both domains.

The scaling and translation actions are performed by two parameters; the scale parameter s ‘adapts’ the width of the wavelet kernel to the *microscopic resolution* required, thus changing its frequency contents, and the location of the analysing wavelet is determined by the parameter b :

$$Wf(s, b) = \frac{1}{s} \int_{-\infty}^{\infty} dx f(x) \psi\left(\frac{x-b}{s}\right),$$

where $s, b \in \mathbf{R}$ and $s > 0$ for the continuous version (CWT).

The 3D plot in figure 2 shows how the wavelet transform reveals more and more detail while going towards smaller scales, i.e. towards smaller $\log(s)$ values. Therefore, the wavelet transform is sometimes referred to as the ‘mathematical microscope’, due to its ability to focus on weak transients and singularities in the time series. The wavelet used determines the optics of the microscope; its magnification varies with the scale factor s .

It can be shown [16] that for *isolated* cusp singularities, the location of the singularity can be detected, and the related exponent can be recovered from the scaling of the Wavelet Transform, along the so-called *maxima line*, converging towards the singularity. This is a line where the wavelet trans-

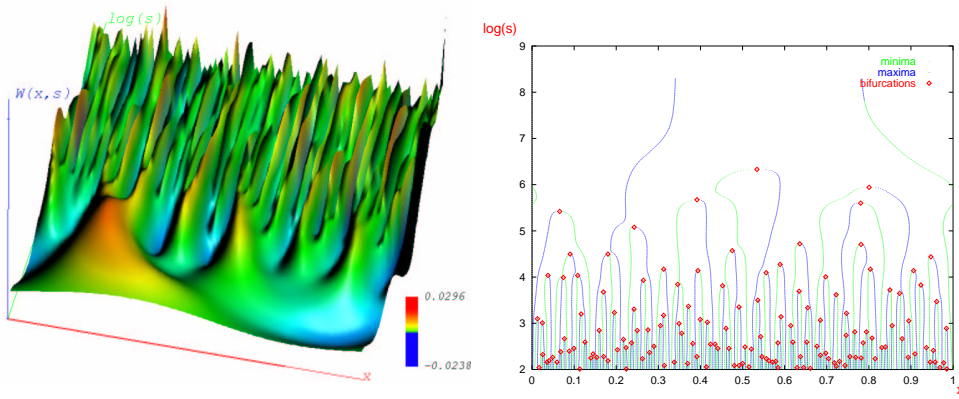


Figure 2: Left: Continuous Wavelet Transform representation of the random walk (Brownian process) time series. The wavelet used is the Mexican hat - the second derivative of the Gaussian kernel. The coordinate axes are: position x , scale in logarithm $\log(s)$, and the value of the transform $W(s, x)$. Right: the related WTMM representation.

form reaches local maximum (with respect to the position coordinate). Connecting such local maxima within the continuous wavelet transform ‘landscape’ gives rise to the entire tree of maxima lines. Restricting oneself to the collection of such maxima lines provides a particularly useful representation [17] (WTMM) of the entire CWT. It incorporates the main characteristics of the WT: the ability to reveal the *hierarchy* of (singular) features, including the scaling behaviour [9]. Restricting oneself to the collection of such maxima lines provides a particularly useful representation of the entire CWT. In particular, we have the following power law proportionality for the wavelet transform of the *isolated* cusp singularity in $f(x_0)$:¹

$$W^{(n)}f(s, x_0) \sim |s|^{h(x_0)}.$$

This is on condition that the wavelet has at least n vanishing moments, i.e. it is orthogonal to polynomials up to degree n : $\int_{-\infty}^{+\infty} x^m \psi(x) dx = 0 \quad \forall m, 0 \leq m < n$.

Moreover, the wavelet transform and its WTMM representation can also be shown to be invariant with respect to the rescaling/renormalisation operation [25, 9, 21, 20]. This property makes it an ideal tool for revealing the renormalisation structure of the (hypothetical) multiplicative process underlying the analysed time series.

3. MULTIFRACTAL FORMALISM ON THE WTMM TREE

The WTMM tree lends itself very well to defining the partition function based multifractal formalism (MF) [9]. The MF takes the moments q of the measure distributed on the WTMM tree to obtain the dependence of the scaling function $\tau(q)$ on the moments q :

$$\mathcal{Z}(s, q) \sim s^{\tau(q)}.$$

The $\mathcal{Z}(s, q)$ is the partition function of the q -th moment of the measure distributed over the wavelet transform maxima at the scale s considered:

¹One should bear in mind that the above relation is an approximate case for which exact theorems exist [18]. In particular, we will restrict the scope of this paper to Hölder singularities for which the local and point-wise Hölder exponents are equal [19]. Thus we will not take into consideration the ‘oscillating singularities’ (e.g. $x^\alpha \sin(1/x^\beta)$) requiring two exponents [18, 20]. Nevertheless, it is sufficient for our purpose to state that the continuous wavelet transform can be used for characterising the Hölder singularities in the time series even if masked by the polynomial bias.

$$\mathcal{Z}(s, q) = \sum_{\Omega(s)} (W f \omega_i(s))^q, \quad (3.1)$$

where $\Omega(s) = \{\omega_i(s)\}$ is the set of all maxima $\omega_i(s)$ at the scale s , satisfying the constraint on their local logarithmic derivative in scale [22]. (The local slope bound used throughout this paper is $|\check{h}| \leq 2$.)

Since the moment q has the ability to select a desired range of values: small for $q < 0$, or large for $q > 0$, the scaling function $\tau(q)$ globally captures the distribution of the exponents $h(x)$ - weak exponents are addressed with large negative q , while strong exponents are suppressed and effectively filtered out. For the large positive q , the opposite takes place (and strong exponents are addressed while weak exponents are effectively filtered out).

This dependence may be linear, indicating that there is only one class of singular structures and related exponents, or it may have a slope non-linearly changing with q . In the latter case, the local tangent slope to $\tau(q_*)$ will give the corresponding exponent, i.e. $h(q_*)$, with its related dimension marked on the ordinate axis $C = D(h(q_*))$, where $\tau(q_*) = h(q_*)q_* + C$. The set of values C , i.e. dimensions $D(h(q_*))$ for each value of h selected with q_* is the so-called spectrum of the singularities $D(h)$ of the fractal signal. Formally, the transformation from $\tau(q)$ to $D(h)$ is referred to as the Legendre transformation:

$$\begin{aligned} \frac{d\tau(q)}{dq} &= h(q), \\ D(h(q)) &= q h(q) - \tau(q). \end{aligned} \quad (3.2)$$

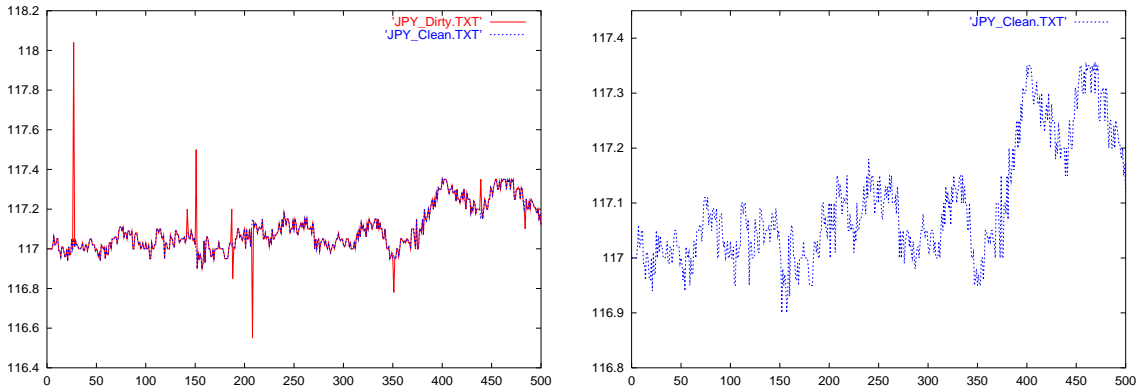


Figure 3: Left: ‘First_Dirty’. Right: ‘First_Clean’. The only difference between the two time series is a number of erroneous spikes which in ‘First_Clean’ were localised using external information and were removed by hand. The plot of ‘First_Clean’ still shows spikes which belong to the process investigated. The task of the methodology to be described is to detect the presence of erroneous spikes in time series and provide the means of localising them. The renormalised plot of ‘First_Clean’ reveals its complex form. Spiky events are still present in the time series, but these belong to the process.

From this transformation, we can directly obtain expressions for the average $h(q)$ and $D(h(q))$ in terms of the partition function over maxima values. From 3.2 and 3.1 we have:

$$h(q) = \frac{d\tau(q)}{dq} = \lim_{s \rightarrow 0} \frac{d}{dq} \frac{\log(\mathcal{Z}(s, q))}{\log(s)} = \lim_{s \rightarrow 0} \frac{1}{\log(s)} \sum_{\Omega(s)} P(s, q, \omega_i(s)) \log(W f \omega_i(s)) \quad (3.3)$$

where

$$P(s, q, \omega_i(s)) = \frac{(Wf\omega_i(s))^q}{\sum_{\Omega(s)} (Wf\omega_i(s))^q} \quad (3.4)$$

is the weighting measure for the statistical ensemble $\Omega(s)$ [10]. Similarly we obtain the expression for $D(h(q))$:

$$D(h(q)) = \lim_{s \rightarrow 0} \frac{1}{\log(s)} \sum_{\Omega(s)} p(s, q, \omega_i(s)) \log(p(s, q, \omega_i(s))) . \quad (3.5)$$

Usually this pair is used to obtain $D(h(q))$ spectrum in a parametric form (q is a parameter here). We will, however, also show $D(q)$ and $h(q)$ separately, since the explicit dependence on q is crucial for our purpose.

In figure 4, we show $D(h(q))$ evaluated for the ‘Dirty’ time series from figure 1 and for its cleaned version where the outlier spikes were removed using external information, see figure 3.

The difference in the spectral information is striking and well reflects the high sensitivity of the partition function method to outliers. The spectrum for the clean version is narrow and focused around the mean value of singularity strength $h_{mean} = 0.4$. For the dirty version, we have a very broad spectrum which gradually falls off to zero dimension values for decreasing $h < h_{mean}$. This fall off regime corresponds with positive q values which have the ability to select exponents of a relatively lower value than the h_{mean} value. Since (in this example) we are operating in the $h = 0.4$ range, thus capturing spikes for which $h = -1$ (in an isolated situation) is relatively easy.

Let us immediately remark that processes exist which have a wide range of h among their characteristic. Such processes are generically called *multifractal*, as a fractal dimension $D(h)$ is associated with each h . Hence, if there is a multitude of meaningful values $D(h_*)$ (associated with some h_*) constituting the spectrum of the process, the process can be of multifractal type. The difference between such processes and the processes with outliers is in the relative values of $D(q)$ and the spacing between the successive q ’s. For the process to have a meaningful multifractal spectrum, we will require dense coverage of $h(q)$ values. Also the $D(q)$ values should be relatively large for meaningful spectra. Dense support on a line corresponds with dimension 1. Single points, that is separated point-wise events, on the other hand have support 0. Therefore, if $D(q)$ is near 0, this indicates very weakly supported events and therefore a high probability of an outlier.

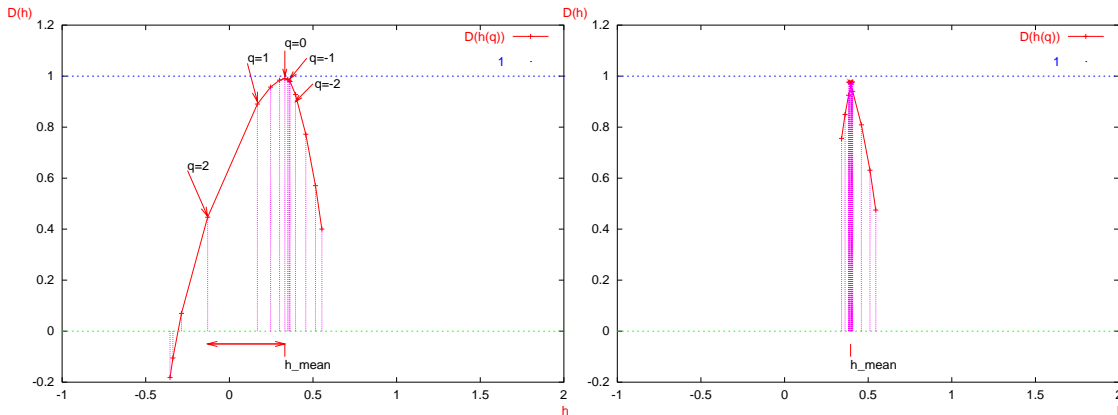


Figure 4: Left: $D(h(q))$ evaluated for the ‘First_Dirty’ time series from figure 1. Right: the same for the ‘Clean’ version of the same time series. (The outlier spikes were removed using external information.) A clear difference in D spectrum is visible.

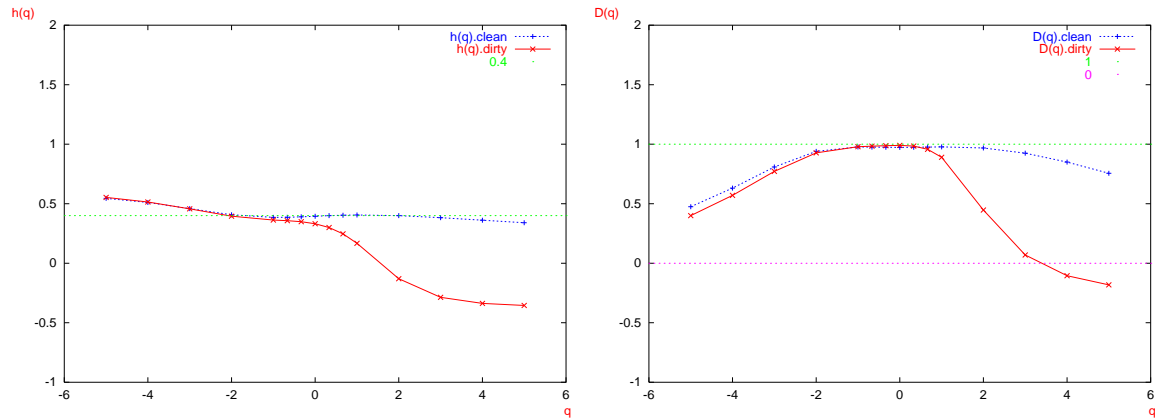


Figure 5: Left: $h(q)$ evaluated from 3.3 for both ‘Clean’ and ‘Dirty’ time series from fig 1. h shows a strong crossover, and for $q > 1$ quickly falls off the average 0.4 value for this process. Right: $D(q)$ evaluated from 3.5 for the same time series. Again a clear difference in behaviour is visible. D quickly approaches 0 for q larger than 1.

In figure 5 we check $D(q)$ and $h(q)$ separately in order to analyse the dependence on q . The test data in figure 5 is contaminated with Dirac delta type outliers. For comparison we also analyse outlier-free time series. $h(q)$ evaluated from Eq. 3.3, shows a strong crossover, and for $q > 1$ it quickly falls off the average h_{mean} value. Also in the case of $D(q)$ evaluated from Eq. 3.5 for the contaminated time series, a clear difference in behaviour is visible. $D(q)$ quickly approaches 0 for q larger than 1.

The conclusion that we can derive from such test results is that comparing the value of both the $h(q)$ and $D(q)$ for positive q 's may be useful for detecting the presence of spikes in the time series. In particular the second moment $q = 2$ seems to be suitable for use as a criterium in comparison with the reference $q = 0$ moment. The rule of thumb which we suggest is that if the value of $h(q = 2)$ differs from $h(q = 0)$ by some 0.5 and if $D(q = 2)$ is about 0.5 or less, the probability of spike presence is relatively high.

For actual localisation of spikes, we need the local value of $h(x)$ instead of the global average as defined in Eq. 3.3. Such a local $h(x)$ will make possible separating outliers from the residue, using some threshold value for h . This will be discussed further in the next section.

4. ESTIMATION OF THE LOCAL, EFFECTIVE HÖLDER EXPONENT USING THE MULTIPLICATIVE CASCADE MODEL

Note that even though the partition function method (discussed thus far) uses the maxima tree containing full local information about the singularities, this is lost at the very moment the partition function is computed. Therefore, there is no explicit local information present in the scaling estimates; τ , h or D , and all these are global statistical estimates. This is also where the strength of the partition function method lies - global averages are much more stable than local information and in some cases global information is all that it is possible to obtain.

We have shown in the previous section that the wavelet transform, and in particular its maxima lines, can be used in evaluating the Hölder exponent in isolated singularities. In most real life situations, however, the singularities in the time series are not isolated but densely packed. The logarithmic rate of increase or decay of the corresponding wavelet transform maximum line is usually not stable but fluctuates, following the action of the (hypothetical, multiplicative) process involved.

Indeed, it is generally not possible to obtain local estimates of the scaling behaviour other than in the case of isolated singular structures from the WT. This is why we introduced [23] an approach circumventing this problem while retaining local information - a local effective Hölder exponent, in which we model the singularities as created in some kind of a collective process of a very generic class.

To capture the fluctuations and estimate the related exponents (to which we will refer as an *effective* Hölder exponent of the singularity), we will model the singularities as created in some kind of a collective process of a very generic class - the multiplicative cascade model. Each point in this cascade is uniquely characterised by the sequence of weights $(s_1 \dots s_n)$, taking values from the (binary) set $\{1, 2\}$, and acting successively along a unique process branch leading to this point. Suppose that we denote the density of the cascade at the generation level F_i (i running from 0 to max) by $\kappa(F_i)$, we then have

$$\kappa(F_{max}) = p_{s_1} \dots p_{s_n} \kappa(F_0) = P_{F_0}^{F_{max}} \kappa(F_0)$$

and the local exponent is related to the rate of increase of the product $P_{F_0}^{F_{max}}$ over the gained scale difference. In any experimental situation, the weights p_i are not known and h has to be estimated. This can be simply done using the fact that for the multiplicative cascade process, the effective product of the weighting factors is reflected in the difference of logarithmic values of the densities at F_0 and F_{max} along the process branch:

$$h_{F_0}^{F_{max}} = \frac{\log(\kappa(F_{max})) - \log(\kappa(F_0))}{\log((1/2)^{max}) - \log((1/2)^0)}.$$

The densities along the process branch can be estimated with the wavelet transform, using its remarkable ability to reveal the entire process tree of a multiplicative process [25]. It can be shown that the densities $\kappa(F_i)$ corresponds with the value of the wavelet transform along the maxima lines belonging to the given process branch. The estimate of the effective Hölder exponent becomes:

$$\hat{h}_{s_{l_0}}^{s_{hi}} = \frac{\log(Wf\omega_{pb}(s_{l_0})) - \log(Wf\omega_{pb}(s_{hi}))}{\log(s_{l_0}) - \log(s_{hi})},$$

where $Wf\omega_{pb}(s)$ is the value of the wavelet transform at the scale s , along the maximum line ω_{pb} corresponding to the given process branch. Scale s_{l_0} corresponds with generation F_{max} , while s_{hi} corresponds with generation F_0 , (simply the largest available scale in our case).

4.1 Estimation of the Mean Hölder Exponent

For a multiplicative cascade process, a mean value of the cascade at the scale s can be defined as:

$$\mathcal{M}(s) = \frac{\sum_{\Omega(s)} \log(Wf\omega_i(s))}{\mathcal{Z}(s, 0)}, \quad (4.1)$$

where the $\mathcal{Z}(s, 0)$ is the partition function Eq. 3.1 for $q = 0$ and corresponds with the number of maxima at the scale s considered.

This mean is compatible with the canonical formalism based spectrum,² see Eq. 3.3, and gives the direct possibility of estimating the mean value of the local Hölder exponent as a linear fit to \mathcal{M} :

$$\log(\mathcal{M}(s)) = \bar{h} \log s + C. \quad (4.2)$$

Therefore, we estimate our mean Hölder exponent \bar{h} from 4.2. The estimate of the local Hölder exponent, from now on to be denoted as $\hat{h}(x_0, s)$ or just \hat{h} , now becomes:

$$\hat{h}_{s_{l_0}}^{s_{SL}} \cong \frac{\log(Wf(s_{l_0})) - (\bar{h} \log(s_{SL}) + C)}{\log(s_{l_0}) - \log(s_{SL})}.$$

²In Ref. [23], we used a different, 'micro-canonical' mean, which in general differs slightly from the canonical one. For reasons of the compatibility of the spectra, we take the canonical mean here.

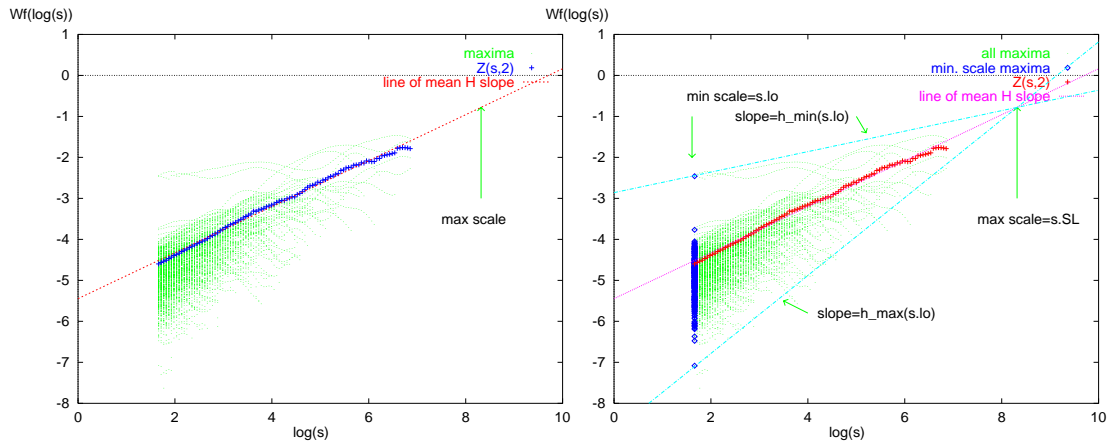


Figure 6: Left: the projection of the maxima lines of the WT along time. The mean value of the Hölder exponent can be estimated from the log-log slope of the line shown. Also, the beginning of the cascade at the maximum scale s_{hi} is indicated. Right: the maxima at the smallest scale considered are shown in the projection along time. The effective Hölder exponent can be evaluated for each point of the maximum line at s_{lo} scale. Two extremal exponent values are indicated, for minimum and maximum slope.

4.2 Employing the Effective Hölder Exponent in Local and Global Spectra Estimation Examples

Such an estimated local $\hat{h}(x, s)$ is a function of the same x parameter (time, position) as the analysed function $f(x)$, and can be analysed in a local fashion or histogrammed in order to study its distribution properties.

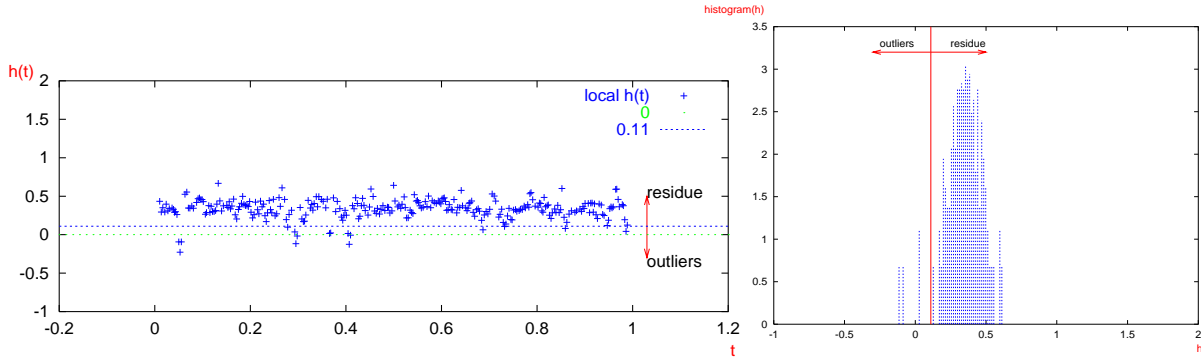


Figure 7: Left: the local Hölder exponent for the time series from figure 3 left. Right: the corresponding log-histograms of the local Hölder exponent. Thresholding on h separates outliers from residue.

In figure 7, we depict it in a temporal fashion for the record contaminated with Dirac type events, see figure 4. The local $\hat{h}(x, a = \log(10))$ plot is rather varied, but still clustered around the $h_{mean} = 0.4$. There are several 'drop down' events present, indicating the presence of strong singular events. These can be selected using an appropriate threshold on h .

Generic criteria for threshold choice can be obtained from the statistical distribution of the h exponent. For this purpose, the log-histogram³ of h can be analysed, see figure 7 right. For noisy

³It is obtained by taking the logarithm of the measure in each histogram bin. This conserves the monotonicity of the original histogram, but allows us to compare the log-histograms with the spectrum of singularities $D(h)$. By following the evolution of the log-histograms along scale, we will be able to extract the spectrum of the singularities $D(h)$.

time series, the corresponding histograms show considerable widening, often accompanied by visible fragmentation (discontinuity).

Determining the ‘noiseless’ width of the histogram and the point above/below which ‘noise’ starts is a crucial task in this procedure. The histogram can provide good insight, but it is somehow hard to automatize the choice of the threshold level and make it arbitrary. We have, therefore, resorted to a simple heuristic.

A good threshold value which worked well in our tests (for an account of the experimental work we refer the reader to Ref [24]) was determined using the scaling exponent of (the square root of) the second moment of the measure $\mathcal{Z}(s, 2)$:

$$\mathcal{M}'(s) = \sqrt{\frac{\mathcal{Z}(s, 2)}{\mathcal{Z}(s, 0)}}, \quad (4.3)$$

and the thresholding \hat{h} exponent is then determined from the linear fit:

$$\log(\mathcal{M}'(s)) = \hat{h} \log(s) + C' . \quad (4.4)$$

In cases where there are outliers in the time series, this quantity, the ‘micro-canonical’ geometric mean, \hat{h} does not coincide with the h_{mean} mode value of the $D(h)$ distribution. In such cases it can be used for the threshold between the outliers and the residue, as in figure 7. In fact, the \hat{h} value can also be used as an additional criterium (besides that suggested in the previous section) for testing for the presence of outliers. In cases where the h_{mean} differs significantly from \hat{h} , the probability of outliers is high.

Finally, the histogram of the h exponent is shown in figure 7, with the tail to the lower h values clearly visible. The same threshold \hat{h} as in the local plot separates the residual bulk and the tail outlier events.

5. CONCLUSIONS

We have presented a method of detecting and localising outliers in stochastic processes. The method checks the internal consistency of the scaling spectrum of the process within the paradigm of the multifractal spectrum. Deviation from the expected spectrum is interpreted as the potential presence of outliers. We suggested a rule of thumb to detect the presence of outliers based on the global, partition function based multifractal formalism. The detection part is then supplemented by the localisation analysis part, using the local scaling properties of the time series. Each data point has an (effective) Hölder exponent associated with it, and based on some threshold level, the outliers can be separated from the residue of the exponent distribution. In particular, localised outliers can then be removed one by one, with the possibility of the dynamic verification of spectral properties (in a dynamic programming approach).

A possible extension to the methodology is in estimating the ‘size’ of outliers from the scale at which the associated maxima lines appear. Properties of the size distribution of all singular events could then be used with the related Hölder exponent distribution for outlier discrimination and localisation.

Another direction is the real time, fast computation of the time series local effective Hölder exponent and scaling properties with the use of the Haar wavelet [26]. Testing in real time for the likelihood of the current sample (that last acquired) being an outlier is crucial in some financial applications.

References

1. A. Johansen, D. Sornette, Stock Market Crashes are Outliers, *Eur.Phys.J. B* **1**, pp. 141-143 (1998).
2. A. Johansen, D. Sornette, Large Stock Market Price Drawdowns Are Outliers arXiv:cond-mat/0010050, 3 Oct 2000, rev. 25 Jul 2001.
3. V.S. L'vov, A. Pomyalov, I. Procaccia, Outliers, Extreme Events and Multiscaling, arXiv:nlin.CD/0009049, 27 Sept 2000.
4. A. Arneodo, J.F. Muzy, D. Sornette, *Eur. Phys J. B*, **2**, 277 (1998).
5. A. Fisher, L. Calvet, B.B. Mandelbrot, Multifractality of the deutschmark/us dollar exchange rate, Cowles Foundation Discussion Paper, (1997).
6. M.E. Brachet, E. Taflin, J.M. Tch  ou, Scaling transformations and probability distributions for financial time series, preprint cond-mat/9905169, (1999).
7. F. Schmitt, D. Schwertzer, S. Levejoy, Multifractal analysis of Foreign Exchange data, preprint, McGill University Montreal (1999).
8. A. Arneodo, E. Bacry, J.F. Muzy, Wavelets and Multifractal Formalism for Singular Signals: Application to Turbulence Data, *PRL*, **67**, No 25, 3515-3518, (1991).
9. A. Arneodo, E. Bacry and J.F. Muzy, The Thermodynamics of Fractals Revisited with Wavelets. *Physica A*, **213**, 232 (1995).
J.F. Muzy, E. Bacry and A. Arneodo, The Multifractal Formalism Revisited with Wavelets. *Int. J. of Bifurcation and Chaos* **4**, No 2, 245 (1994).
10. A. Arneodo, E. Bacry and J.F. Muzy, Wavelet Analysis of Fractal Signals: Direct Determination of the Singularity Spectrum of Fully Developed Turbulence Data
11. A. Arneodo, A. Argoul, J.F. Muzy, M. Tabard, E. Bacry, Beyond Classical Multifractal Analysis using Wavelets: Uncovering a Multiplicative Process Hidden in the Geometrical Complexity of Diffusion Limited Aggregates, *Fractals* **1**, 629-646, (1995).
12. A. Arneodo, E. Bacry, P.V. Graves, J.F. Muzy, Characterizing Long-Range Correlations in DNA Sequences from Wavelet Analysis, *PRL*, **74**, No 16, 3293-3296, (1995).
13. P.Ch. Ivanov, M.G. Rosenblum, C.-K. Peng, J. Mietus S. Havlin, H.E. Stanley, A.L. Goldberger, Scaling Behaviour of Heartbeat Intervals Obtained by Wavelet-based Time-series Analysis, *Nature*, **383**, 323-327, (1996).

14. I. Daubechies, *Ten Lectures on Wavelets*, (S.I.A.M., 1992).
15. M. Holschneider, *Wavelets - An Analysis Tool*, (Oxford Science Publications, 1995).
16. S.G. Mallat and W.L. Hwang, Singularity Detection and Processing with Wavelets. *IEEE Trans. on Information Theory* **38**, 617 (1992).
17. S.G. Mallat and S. Zhong, Complete Signal Representation with Multiscale Edges. *IEEE Trans. PAMI* **14**, 710 (1992).
18. S. Jaffard and Y. Meyer, Wavelet Methods for Pointwise Regularity and Local Oscillations of Functions. *Memoirs of AMS*, 123 (1996).
19. B. Guiheneuf and J. Lévy Véhel, 2-Microlocal Analysis and Application in Signal Processing, in *Proc. of Int. Wavelets Conference*, Tangier (1998).
20. A. Arneodo, E. Bacry and J.F. Muzy, Oscillating Singularities in Locally Self-Similar Functions. *PRL*, **74**, No 24, 4823 (1995).
21. A. Arneodo, E. Bacry and J.F. Muzy(1994): Solving the Inverse Fractal Problem from Wavelet Analysis. *Europhysics Letters*, **25**, No 7, 479–484.
22. Z.R. Struzik, Removing Divergences in the Negative Moments of the Multi-Fractal Partition Function with the Wavelet Transformation. *CWI Report*, **INS-R9803**. Also see ‘Fractals and Beyond - Complexities in the Sciences’, M.M. Novak, Ed., World Scientific, 351 (1998).
23. Z. R. Struzik, Local Effective Hölder Exponent Estimation on the Wavelet Transform Maxima Tree, in *Fractals: Theory and Applications in Engineering*, Eds: M. Dekking, J. Lévy Véhel, E. Lutton, C. Tricot, Springer Verlag, (1999).
24. Z.R. Struzik, A. Siebes, Outlier Detection and Localisation with Wavelet Based Multifractal Formalism, *CWI Report*, **INS-R0008** (2000).
25. Z.R. Struzik, The Wavelet Transform in the Solution to the Inverse Fractal Problem. *Fractals* **3** No.2, 329 (1995).
Z.R. Struzik, From Coastline Length to Inverse Fractal Problem: The Concept of Fractal Metrology. *Thesis*, University of Amsterdam. (1996).
26. Z.R. Struzik, Oversampling the Haar Wavelet Transform, *CWI Report*, **INS-R0102** (2001).