Wavelet methods in (financial) time-series processing

Zbigniew R. Struzik*

Centre for Mathematics and Computer Science (CWI), Amsterdam, The Netherlands

Abstract

We briefly describe the major advantages of using the wavelet transform for the processing of financial time series on the example of the S&P index. In particular, we show how to uncover the local scaling (correlation) characteristics of the S&P index with the wavelet based effective Hölder exponent (Struzik, in: Fractals: Theory and Applications in Engineering, Dekking, Lévy Véhel, Lutton, Tricot, Springer, Berlin, 1999; Fractals 8 (2) (2000) 163). We use it to display the local spectral (multifractal) contents of the S&P index. In addition to this, we analyse the collective properties of the local correlation exponent as perceived by the trader, exercising various time horizon analyses of the index. We observe an intriguing interplay between such (different) time horizons. Heavy oscillations at shorter time horizons, which seem to be accompanied by a steady decrease of correlation level for longer time horizons, seem to be characteristic patterns before the biggest crashes of the index. We find that this way of local presentation of scaling properties may be of economic importance. © 2001 Elsevier Science B.V. All rights reserved.

1. Introduction

Economics has been producing more and more complicated models, trying to capture deviations of the model situation from reality. But patching a model to increase its complexity may not be an optimal way of modelling. Any economic system is extremely complicated but, to a large degree, this is due to the enormous number of

* Tel.: +31-20-592-4123; fax: 31-20-592-4199.

E-mail address: zbigniew.struzik@cwi.nl (Z.R. Struzik).

degrees of freedom. The interactions between economic entities do not need to be very complicated. Approaches aiming at improving the description of a system by means of high order corrections may, therefore, have a rather slow convergence to the satisfactory model.

On the contrary, statistical physics uses very simple models of interactions between components of a huge ensemble. Concepts from statistical physics have proven quite successful in the work of *econophysicists*. The power of modern computers allows not only the testing of such models but also the analysis of data for the presence of multiscaling characteristics, in particular locally in the data. An analysis of correlations can be linked to the models of interaction between the system components, and so can the multiscale distribution analysis.

Physics has a long tradition of dealing with systems with an extreme number of degrees of freedom, where entities are coupled with very simple interaction rules. This is probably why economics has recently experienced great interest among physicists, following the pioneering work of Gene Stanley et al. [1,2], Rosario Mantegna [3,4], Alain Arneodo et al. [5,6], Marcel Ausloos et al. [7,8] and, not to forget, the early work of Benoit Mandelbrot [9,10].

Whereas the physicists derive the final model from the characteristics of the data analysed, using their learned physics knowledge and scientific intuition, the latest computer science trend is to assist the data analyser in model discovery. *Data mining* is the name for this recent direction in machine learning, see for example Refs. [11–13]. Market/shopping basket analysis and insurance or loan scoring are already widely done using data mining techniques. We expect that the number of contributions of data mining/model discovery approaches in the economical sciences will grow rapidly in the coming decade.

Therefore, working without an a priori assumed model is characteristic of the modern approach to economics. Instead, the model is to be inferred from the data. The data is analysed in terms of very generic analysis methods like, for example, wavelet decomposition. The wavelet transform components are then analysed and, of course, simple or complex models can be fitted to such decomposition components. The scaling of moments or distributions can be tested and a hypothesis drawn, but the first step is data analysis in the most generic terms.

The possibility of doing analysis locally is another very attractive option. The stationarity of almost any statistical characteristic fails when applied to the financial data. A similar situation pertains in other complex phenomena (take, for example, the human heartbeat [14–16]). But where the non-stationarities occur, interestingness begins, and with tools like the wavelet transform, capable of taming the non-stationarities (trends), interesting (local) patterns can be discovered in the data.[1]

In Section 2, we briefly introduce the wavelet transformation in its continuous form, we describe the requirements for the wavelet in Section 3 and discuss the advantages for

---

[1] Of course, ultimately we would like to be able to feed such local patterns back into the global theory, but at the moment we will remain more modest and simply local.

time series processing. In particular, we focus in Section 4 on the ability of the wavelet transformation to characterise scale-free behaviour through the Hölder exponent. We describe in brief a technical model enabling us to estimate the scale-free characteristic (the effective Hölder exponent) for the branches of a multiplicative process. A more extensive coverage of this method is available in Refs. [17,18]. In Section 5, we use the derived effective Hölder exponent for the local temporal description of the S&P index. Section 6 provides an extension to the fluctuation analysis of the effective Hölder exponent of the S&P index. Section 7 closes the paper with conclusions.

2. Why wavelets?

The wavelet transform is a convolution product of the signal with the scaled and translated kernel—the wavelet $\psi(x)$ [19,20]. The scaling and translation actions are performed by two parameters; the scale parameter s 'adapts' the width of the wavelet kernel to the resolution required and the location of the analysing wavelet is determined by the parameter b:

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

(1)

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

The 3D plot in Fig. 1 shows how the wavelet transform reveals more and more detail while going towards smaller scales, i.e. towards smaller $\log(s)$ values. The wavelet transform is sometimes referred to as the 'mathematical microscope' [5,6], 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.

Whether we want to use a continuous or discrete WT, see Fig. 2, is largely a matter of application. For coding purposes, one wants to use the smallest number of coefficients which can be compressed by thresholding low values or using correlation properties. For this purpose a discrete (for example a dyadic) scheme of sampling the scale s, position b space is convenient. Such sampling often spans an orthogonal wavelet base.

For analysis purposes, one is not so concerned with numerical or transmission efficiency or representation compactness, but rather with accuracy and adaptive properties of the analysing tool. Therefore, in analysis tasks, continuous wavelet decomposition is mostly used. The space of scale s and position b is then sampled semi-continuously, using the finest data resolution available.

For decomposition, a simple base function is used. The *wavelet* ψ, see Eq. (1) took its name from its wave-like shape. It has to cross the zero value line at least once since its mean value must be zero. The criterion of zero mean is referred to as the *admissibility* of the wavelet, and is related to the fact that one wants to have the possibility of reconstructing the original function from its wavelet decomposition.

This condition can be proven formally, but let us give a quick, intuitive argument. We have seen that wavelets work at smaller and smaller scales, covering higher and
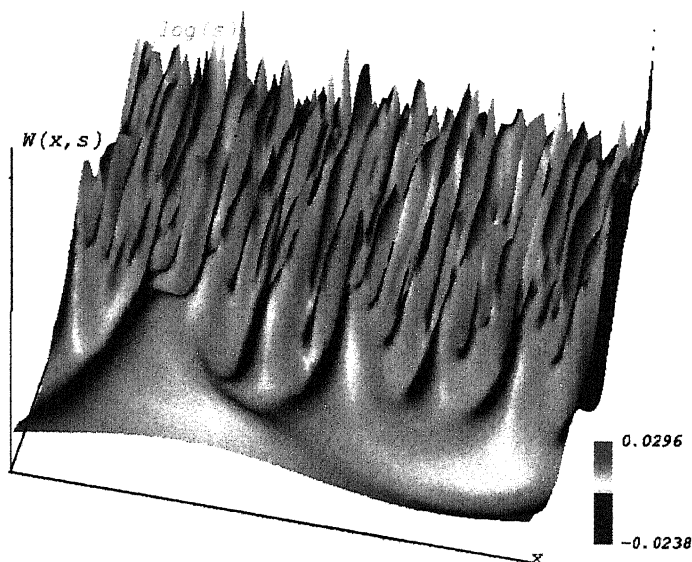
Fig. 1. Continuous wavelet transform representation of the random walk (Brownian process) time series. The wavelet used is the Mexican hat—the second derivative of the Gaussian kernel. The coordinate axes are: position x, scale in logarithm $\log(s)$, and the value of the transform $W(s,x)$.
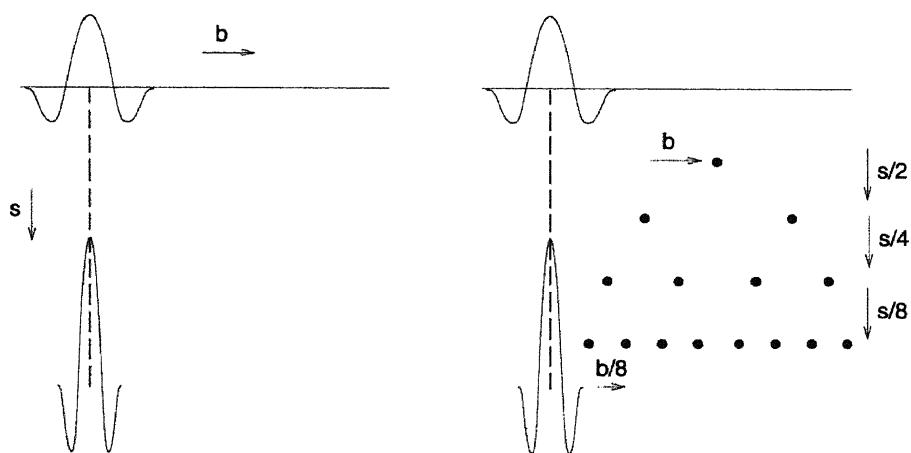

Fig. 2. Continuous sampling of the parameter space (left) versus discrete (dyadic) sampling (right).

higher frequency bands of the signal being decomposed. This is the so-called band pass filtering of the signal. Only a certain band of frequencies (level of detail) is captured by the wavelets working at one scale. Of course, at another scale a different set of details (band of frequencies) is captured. But other frequencies (in particular zero frequency) are not taken into the coefficients. This is the idea of decomposition.

Kernels like the Gaussian smoothing kernel are low-pass filters, which means they evaluate the entire set of frequencies up to the current resolution. This is the idea of approximation at various resolutions.

The reader may rightly guess here that it is possible to get band pass information (wavelet-coefficients) from subtracting two low-pass approximations at various levels of resolution. This is, in fact, the so-called multi-resolution scheme of decomposition into WT components. But back to the admissibility—reconstruction from multiple resolution approximations would not be possible since the same low frequency detail would be described in several coefficients of the low pass decomposition. This, of course, is not the case for wavelets; they select only a narrow band of detail with very little overlap (in the orthogonal case no overlap at all!). In particular, if one requires that the wavelet is zero for frequency zero, i.e., it fully blocks zero frequency components, this corresponds with the zero mean admissibility criterion.

3. The wavelet

The only *admissibility* requirement for the wavelet ψ is that it has zero mean—it is a wave function, hence the name *wavelet*.

$$\int_{-\infty}^{\infty} \psi(x)\,\mathrm{d}x = 0 \,. \tag{2}$$

However, in practice, wavelets are often constructed with orthogonality to a polynomial of some degree n.

$$\int_{-\infty}^{\infty} x^n \psi(x)\,\mathrm{d}x = 0 \,. \tag{3}$$

This property of the wavelets—orthogonality to polynomials of degree n—has a very fine application in signal analysis. It is referred to by the name of the number of vanishing moments. If the wavelet is orthogonal to polynomials of a degree up to n and including n, we say that it has $m = n + 1$ vanishing moments. So one vanishing moment is good enough to filter away constants—polynomials of zero degree P_0. This can be done, for example, by the first derivative of the Gaussian kernel plotted in Fig. 3. Similarly the second derivative of the same Gaussian kernel, which is often used upside down and then appropriately called the Mexican hat wavelet, has two vanishing moments and in addition to constants can also filter linear trends P_1. Of course, if the wavelet has m vanishing moments, it can filter polynomials of degree $m - 1, m - 2, \ldots, 0$.

4. The Hölder exponent

The use of vanishing moments becomes apparent when we consider local approximations to the function describing our time series. Suppose we can *locally* approximate
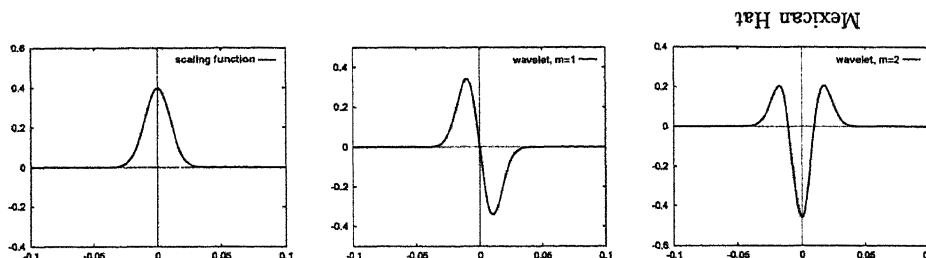
Fig. 3. Left: the smoothing function: the Gaussian. Centre: wavelet with one vanishing moment, the first derivative of the Gaussian. Right: wavelet with two vanishing moments, the second derivative of the Gaussian.

the function with some polynomial P_n, but the approximation fails for P_{n+1}. One can think of this kind of approximation as the Taylor series decomposition. In fact the arguments to be given are true even if such Taylor series decomposition does not exist, but it can serve as an illustration.

For the sake of illustration, let us assume that the function f can be characterised by the Hölder exponent $h(x_0)$ in x_0, and f can be locally described as:

$$f(x)_{x_0} = c_0 + c_1(x - x_0) + \cdots + c_n(x - x_0)^n + C|x - x_0|^{h(x_0)}$$

$$= P_n(x - x_0) + C|x - x_0|^{h(x_0)}.$$

The exponent $h(x_0)$ is what 'remains' after approximating with P_n and what does not yet 'fit' into an approximation with P_{n+1}. More formally, our function or time series $f(x)$ is locally described by the polynomial component P_n and the so-called Hölder exponent $h(x_0)$.

$$|f(x) - P_n(x - x_0)| \leqslant C|x - x_0|^{h(x_0)}.\tag{4}$$

It is traditionally considered to be important in economics to capture trend behaviour P_n. It is, however, widely recognised in other fields that it is not necessarily the regular polynomial background but quite often the transient singular behaviour which can carry important information about the phenomena/underlying system 'producing' the time series.

One of the main reasons for the focus on the regular component was that until the advent of multi-scale techniques (like WT) capable of locally assessing the singular behaviour, it was practically impossible to analyse singular behaviour. This is because the weak transient exponents h are usually completely masked by the much stronger P_n.

However, wavelets provide a remedy in this case! The reader has perhaps already noted the link with the vanishing moments of the wavelets. Indeed, if the number of the vanishing moments is at least as high as the degree of P_n, the wavelet coefficients will capture the local scaling behaviour of the time series as described by $h(x_0)$.

In fact, the phrase 'filtering' with reference to the polynomial bias is not entirely correct. The actual filtering happens only for wavelets the support of which is fully incorporated in the biased interval. If the wavelet is at the edge of such an interval
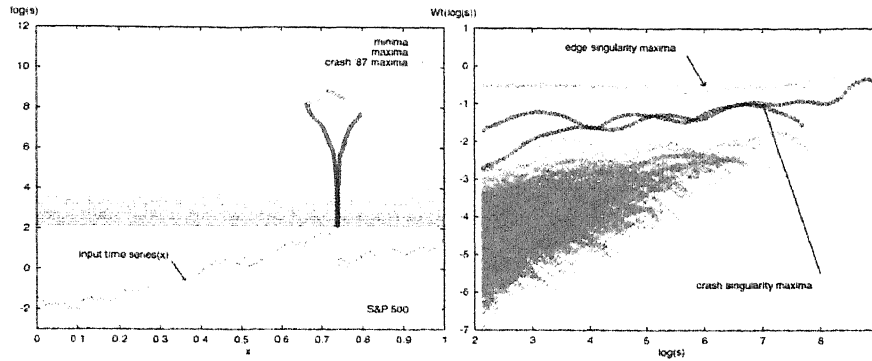
Fig. 4. Left: the input time series with the WT maxima above it in the same figure. The strongest maxima correspond to the crash of '87. The input time series is debiased and L1 normalised. Right: we show the same crash related maxima highlighted in the projection showing the logarithmic scaling of all the maxima.

(where bias begins) or if the current resolution of the wavelet is simply too large with respect to the biased interval, the wavelet coefficients will capture the information pertinent to the bias. This is understandable since the information does not get 'lost' or 'gained' in the process of WT decomposition. The entire decomposed function can be reconstructed from the wavelet coefficients, including the trends within the function. What wavelets provide in a unique way is the possibility to tame and manage trends in a local fashion, through localised wavelets components.

Above, we have suggested that the function can locally be described with Eq. (4). Its wavelet transform $W^{(n)}f$ with the wavelet with at least n vanishing moments now becomes:

$$W^{(n)}f(s,x_0) = \frac{1}{s} \int C|x - x_0|^{h(x_0)} \psi\left(\frac{x - x_0}{s}\right) dx = C|s|^{h(x_0)} \int |x'|^{h(x_0)} \psi(x') dx' .$$

Therefore, we have the following power law proportionality for the wavelet transform of the (Hölder) singularity of $f(x_0)$:

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

From the functional form of the equation, one can immediately attempt to extract the value of the local Hölder exponent from the scaling of the wavelet transform coefficients in the vicinity of the singular point x_0. This is indeed possible for singularities which are isolated or effectively isolated, that is that can be seen as isolated from the current resolution of the analysing wavelet. A common approach to trace such singularities and to reveal the scaling of the corresponding wavelet coefficients is to follow the so-called maxima lines of the CWT converging towards the analysed singularity. This approach was first suggested by Mallat et al. [21,22] and later used and further developed among others in Refs. [5,6,23,24].

In Fig. 4, we plot the input time series which is a part of the S&P index containing the crash of '87. In the same figure, we plot corresponding maxima derived from the

CWT decomposition with the Mexican hat wavelet. The maxima corresponding to the crash stand out both in the top view (they are the longest ones) and in the side log–log projection of all maxima (they have a value and slope different from the remaining bulk of maxima). The only maxima higher in value are the end of the sample finite size effect maxima. These observations indicate that the crash of '87 can be viewed as an isolated singularity in the analysed record of the S&P index for practically the entire wavelet range used.

This is, however, (luckily) an unusual event and in general in time series we have densely packed singularities which cannot be seen as isolated cases for a wider range of wavelet scales. The related Hölder exponent can then be measured either by selecting smaller scales or by using some other approach. A possibility we would like to suggest is using the multifractal paradigm in order to estimate what we call the effective Hölder exponent. The detailed discussion of this approach can be found in Refs. [17,18], but let us quickly point out that the effective Hölder exponent captures local deviations from the mean scaling exponent of the decomposition coefficients related to the singularity in question. This approach has been quite successful in evaluating histograms of the scaling exponents, singularity spectra and collective properties of the local Hölder exponent.

Dense singularities can be seen as evolving from a multiplicative cascading process which takes place across scales. The CWT has been successfully used in revealing such a process and in recovering its characteristics. In short, all that is then needed to evaluate the local effective Hölder exponent for a singularity at a particular scale is the gain in process density across scales with respect to the scale gain:

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

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 bp. Scale s_{lo} corresponds with generation F_{max}, while s_{hi} corresponds with generation F_0, (simply the largest available scale in our case).

5. Employing the local effective Hölder exponent in the characterisation of time series

Such an estimated local $\hat{h}(x_0, s)$ can be depicted in a temporal fashion, for example with colour stripes, as we have done in Fig. 5. The colour of the stripes is determined by the value of the exponent $\hat{h}(x_0, s)$ and its location is simply the x_0 location of the analysed singularity (in practice this amounts to the location of the corresponding maximum line). Colour coding is done with respect to the mean value, which is set to the green colour central to our rainbow range. All exponent values lower than the mean value are given colours from the 'warmer' side of the rainbow, all the way towards dark red. All higher than average exponents get 'colder' colours, down to dark blue.
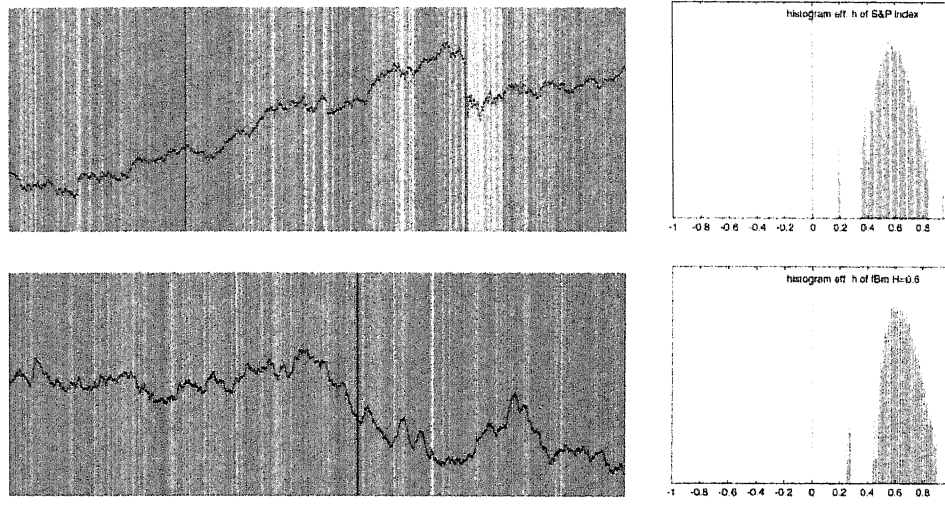
Fig. 5. Left: example time series with local Hurst exponent indicated in colour: the record of healthy heartbeat intervals and white noise. The background colour indicates the Hölder exponent locally, centred at the Hurst exponent at green; the colour goes towards blue for higher $\hat{h}$ and towards red for lower $\hat{h}$. Right: the corresponding log–histograms of the local Hölder exponent. Full colour version can be downloaded from www.cwi.nl/~zbyszek

The first example time series is a record of the S&P500 index from time period 1984– 1988. There are significant fluctuations in colour in this picture, with the green colour centred at $h = 0.55$, indicating both smoother and rougher components. In particular, one can observe an extremal red value at the crash of '87 coordinate, followed by very rough behaviour (a rather obvious fact, but to the best of our knowledge not reported to date in the rapidly growing coverage of this time series record), see e.g. Ref. [25].

The second example time series is a computer generated sample of fractional Brownian motion with $H = 0.6$. It shows almost monochromatic behaviour, centred at $H = 0.6$; the colour green is dominant. There are, however, several instances of darker green and light blue, indicating locally smooth components.

It is important to notice that $h = H_{Brownian\_Walk}$, the Hölder exponent value equal to the Hurst exponent of an uncorrelated Brownian walk, corresponds with no correlation in time series.[2] An ideal random walk would have only monochromatic components of this value. Of course, an ideal, infinitely long record of fractional Brownian motion of $H = 0.6$ is correlated, but this correlation would be stationary in such an ideal case and no fluctuations in correlation level (in colour) would be observed. By the same argument, we can interpret the variations in h as the local fluctuations of correlation in the S&P index. The more red the colour, the more unstable, the more anti-correlated the index. And the more blue, the more stable and correlated.

---

[2] Theoretically this is $h = 0.5$, but finite size sample effects usually add some degree of correlation, slightly increasing this value.

To the right of Fig. 5, the log–histograms of the Hölder exponent displayed in the colour panels are shown. They are made 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 $D(h)$. The log–histograms are actually closely related to the (multifractal) spectra of the Hölder exponent [18]. The multifractal spectrum of the Hölder exponent is the 'limit histogram' $D_{s\to 0}(h)$ of the Hölder exponent in the limit of infinite resolution. Of course we cannot speak of such a limit other than theoretically and, therefore, a limit histogram (multi-fractal spectrum) has to be estimated from the evolution of the log–histograms along scale. For details see Ref. [18].

Let us point out that the width of the spectra alone is a relatively weak argument in favour of the hypothesis of the multifractality of the S&P index. The log–histogram of the S&P is only slightly wider than the log–histogram of a record of fractional Brownian motion of comparable length, see Fig. 5. An interesting observation is, however, that the crash of '87 is clearly an outlier in the sense of the log–histogram of Hölder exponent (and therefore in the sense of the MF spectrum). The issue of crashes as outliers has been extensively discussed by Johansen [26] and recently by L'vov et al. [27]. Here we support this observation from another point of view.[3]

6. Discovering structure through the analysis of collective properties of non-stationary behaviour

Non-stationarities are usually seen as the curse of the exact sciences, economics not excluded. Let us here present a different opinion: where the non-stationarities occur, interestingness begins! Non-stationarities can be seen as a departure from some (usually) simple 'model'. For example, this can be the failure of the stationarity of the effective Hölder exponent (see Fig. 5).

In some sense, they indicate that the simple model used is not adequate, but this does not necessarily mean that one needs to patch or replace this low level model. On the contrary, the information revealed by such a low level model may be used to detect higher order structures. In particular, correlations in the non-stationarities may indicate the existence of interesting structures. An intriguing example of such an approach in the financial domain is the work by Arneodo et al. where a correlation structure in S&P index has been revealed [28].

The simplest way of detecting structure, we suggest, is detecting fluctuations or the collective behaviour of the local effective h. This has already been successfully applied in human heartbeat analysis [16]. Here we will present some preliminary results for the S&P index.

---

[3] The careful reader will notice that similar outliers can be seen in the log–histogram for the fBm in Fig. 5. These outliers are the end of the sample singularities and are caused by finite sample size. As such they are clearly outliers.
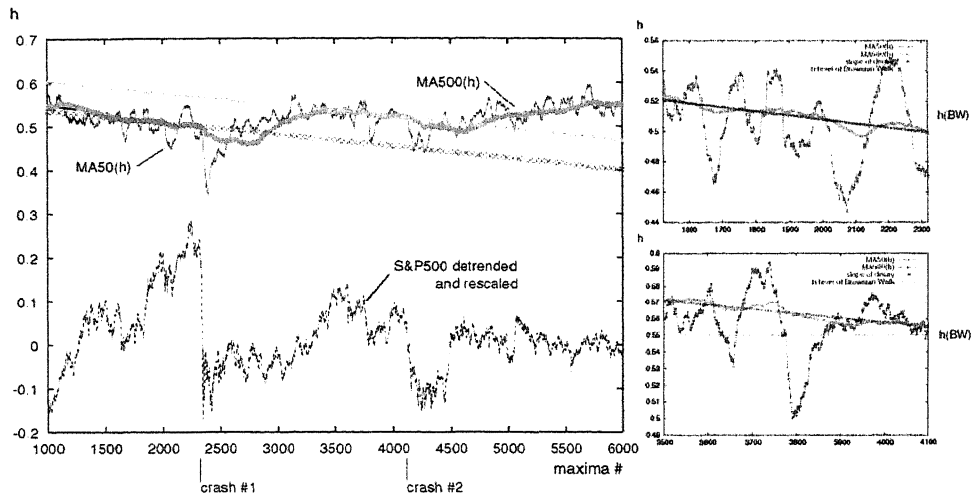
Fig. 6. The effective Hölder exponent smoothed with two windows (MA50 and MA500) is shown. The two plots to the right show windows on the smoothed effective Hölder exponent just before crash #1 and crash #2. (The crashes are visible in the left figure but not in the windows.) Visible oscillations of MA50 and decay of MA500 characterise precursors of both crashes. The average level of the effective Hölder exponent for the uncorrelated Brownian walk is also indicated.

The non-stationary behaviour in h can be quantified, and for this purpose we use a low pass moving average filter (MA) to detect/enhance trends. This processing is, of course, done on the Hölder exponent value set $\{h_i(f(x))\}$, *not* on the input signal $f(x)$. A n-MA filtering of n base is defined as follows:

$$h_{MA_n}(i) = \frac{1}{n}\sum_{i=1}^{i=n} h_i(f(x)) \,, \tag{5}$$

where $h_i(f)$ are the subsequent values of the effective Hölder exponent of the time series f.

Let us now go back to the S&P index and its effective Hölder exponent description. Different window lengths in our MA filter represent different horizons for the trader. If the index is all that is available, in order to evaluate the risk associated with the trading (or in other words, to predict the risk of an index crash), the trader might want to know how 'stable' the market/index is on a daily or monthly time scale. In fact a comparison between the two indicators of stability might be even more indicative.

This is exactly what we have done using two different time scales (two trading horizons) for the MA smoothing, see Fig. 6. The smoothed input is the effective Hölder exponent of the S&P index. It corresponds closely with the logarithm of the local volatility and as such it reflects the stability of the market.

We made the following observations from this experiment: the short time horizon MA shows a strong oscillatory pattern in collective behaviour of the h. These oscillations have already been observed by Liu et al. [2] and by Vandewalle et al.

[7,8]. This is, however, not log-periodic behaviour in our results and it does not converge to a moment of crash. What can perhaps be used in order to help the trader in evaluating the growing risk is the interplay of the various time horizons. The second MA filter has a time horizon 10 times longer and it shows practically no oscillations. However, its value decays almost monotonically, in the moment just before the crash, reaching the level of correlations characteristic for the random walk (see Fig. 6 right inserts). Note that the crashes themselves are not visible in the insert plots. Let us recall that the main advantage of the effective Hölder exponent above some traditional measures of volatility is that it describes the local level of correlation in the time series. If the value of h is below $h = H_{Brownian\_Walk}$, this means we have an anti-correlated time series which intuitively corresponds with a rather unstable process. The h above $h = H_{Brownian\_Walk}$ indicates the presence of correlations and generally can be associated with 'stability'. Please note that the oscillations in MA50 before the crashes bring the collective h up and down between the correlated and the anti-correlated regimes. Similarly MA500 steadily decays towards the anti-correlated regime just before the crashes.

7. Conclusions

The local effective Hölder exponent has been applied to evaluate the correlation level of the S&P index locally at an arbitrary position (time) and resolution (scale). In addition to this, we analysed collective properties of the local correlation exponent as perceived by the trader exercising various time horizon analyses of the index. A moving average filtering of Hölder exponent based variability estimates was used to mimic the various time horizon analysis of the index. We observed intriguing interplay between different time horizons before the biggest crashes of the index. We find that this way of local presentation of scaling properties may be of economic importance.

References

[1] M.H.R. Stanley, L.A.N. Amaral, S.V. Buldyrev, S. Havlin, H. Leschhorn, P. Maass, M.A. Salinger, H.E. Stanley, Can statistical physics contribute to the science of economics? Fractals 4 (3) (1996) 415–425.
[2] H.E. Stanley, L.A.N. Amaral, D. Canning, P. Gopikrishnan, Y. Lee, Y. Liu, Econophysics: can physicists contribute to the science of economics? Physica A 269 (1999) 156–169.
[3] R.N. Mantegna, H.E. Stanley, Scaling behaviour in the dynamics of an economic index, Nature 376 (1995) 46–49.
[4] R. Mantegna, H.E. Stanley, An Introduction to Econophysics, Cambridge University Press, Cambridge, 2000.
[5] A. Arneodo, E. Bacry, J.F. Muzy, The thermodynamics of fractals revisited with wavelets, Physica A 213 (1995) 232.
[6] J.F. Muzy, E. Bacry, A. Arneodo, The multifractal formalism revisited with wavelets, Int. J. Bifurcation Chaos 4 (2) (1994) 245.
[7] N. Vandewalle, M. Ausloos, Coherent and random sequences in financial fluctuations, Physica A 246 (1997) 454–459.

[8] N. Vandewalle, Ph. Boveroux, A. Minguet, M. Ausloos, The crash of October 1987 seen as a phase transition amplitude and universality, Physica A 255 (1998) 201–210.

[9] B.B. Mandelbrot, H. Taylor, On the distribution of stock price differences, Operations Research 15 (1962) 1057–1062.

[10] B.B. Mandelbrot, The variation of certain speculative prices, J. Business 36 (1963) 394–419.

[11] J.M. Żytkow, J. Rauch (Eds.), Principles of Data Mining and Knowledge Discovery, Lecture Notes in Artificial Intelligence, Vol. 1704, Springer, Berlin, 1999.

[12] D.J. Hand, Data mining: statistics and more? Am. Statist. 52 (1998) 112–118.

[13] D.J. Hand, H. Mannila, P. Smyth, Principles of Data Mining, MIT Press, Cambridge, MA, 2000.

[14] 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 (1996) 323.

[15] P.Ch. Ivanov, M.G. Rosenblum, L.A. Nunes Amaral, Z.R. Struzik, S. Havlin, A.L. Goldberger, H.E. Stanley, Multifractality in human heartbeat dynamics, Nature 399 (1999) 461–465.

[16] Z.R. Struzik, Revealing local variability properties of human heartbeat intervals with the local effective Hölder exponent, Fractals 9 (1) (2001) 77–93. See also CWI Report, INS-R0015, June 2000, available from www.cwi.nl/~zbyszek

[17] Z.R. Struzik, Local Effective Hölder Exponent Estimation on the Wavelet Transform Maxima Tree, in: M. Dekking, J. Lévy Véhel, E. Lutton, C. Tricot (Eds.), Fractals: Theory and Applications in Engineering, Springer, Berlin, 1999.

[18] Z.R. Struzik, Determining Local Singularity Strengths and their Spectra with the Wavelet Transform, Fractals 8 (2) (2000) 163–179.

[19] I. Daubechies, Ten Lectures on Wavelets, SIAM, Philadelphia, PA, 1992.

[20] M. Holschneider, Wavelets—An Analysis Tool, Clarendon, Oxford, 1995.

[21] S.G. Mallat, W.L. Hwang, Singularity detection and processing with wavelets, IEEE Trans. Inform. Theory 38 (1992) 617.

[22] S.G. Mallat, S. Zhong, Complete signal representation with multiscale edges, IEEE Trans. PAMI 14 (1992) 710.

[23] S. Jaffard, Multifractal formalism for functions: I. results valid for all functions, II. Self-similar functions, SIAM J. Math. Anal. 28 (4) (1997) 944–998.

[24] R. Carmona, W.H. Hwang, B. Torrésani, Characterisation of signals by the ridges of their wavelet transform, IEEE Trans. Signal Process. 45 (10) (1997) 480–492.

[25] Y. Liu, P. Gopikrishnan, P. Cizeau, M. Meyer, C.-K. Peng, H.E. Stanley, The statistical properties of the volatility of price fluctuations, arXiv:cond-mat/9903369 v2, 25 March 1999.

[26] A. Johansen, D. Sornette, Stock market crashes are outliers, arXiv:cond-mat/9712005 v3, 16 December 1997.

[27] V.S. L'vov, A. Pomyalov, I. Procaccia, Outliers, extreme events and multiscaling, arXiv:nlin.CD/0009049, 27 September 2000.

[28] A. Arnéodo, J.-F. Muzy, D. Sornette, 'Direct' casual cascade in the stock market, Eur. Phys. J. B 2 (1998) 277–282.