Multifractal Detrended fluctuation Analysis (MFDFA)

Fahimeh Najafi

95203626

Introduction

we can investigate non linearity, fractal and multifractal properties of a time series using MFDFA (multifractal detrended fluctuation analysis). We also apply these methods to surrogated data (rank-wised surrogate and random-phase surrogate) as well as the original time series.

Non-detrending methods can not be used for data which involve trends. In those types of data non-detrending methods like using correlation function yield wrong results. In real life experimental data we usually have some kind of trend (or in general non-stationarities). therefore to be able to work with these type of data and investigate scaling behavior of fluctuations, a good detrending tool is needed. Multifractal Detrended Fluctuation Analysis (MFDFA) is a well established method for analyzing the scaling behavior and fractal properties of a time series when an unknown trend is present.

Multifractal Detrended fluctuation Analysis method

MFDFA method for multifractal characterization and long-range correlation detection of non-stationary time series is explained in steps here.

1. integrating over the time series $x(i)$ with size N

$$Y(i)=\sum_{k=1}^{i} x(k)-\langle x\rangle \qquad i = 1, \dots ,N$$

$Y(i)$ is called the profile.

2. we then divide the integrated time series into segments of equal size. (non-overlapping segments) for a segment size s , number of segments is $N_s=\lfloor N/s \rfloor$ (integer part of N/s).

If N is not a multiple of s , the whole length of data will not be scanned. To include the time series completely in our calculations, we repeat this process of segmentation, this time starting from the end of time series, so there will be $2N_s$ segments in total.

3. the next step is detrending , which involves for any ν^{th} segment , fitting a polynomial P_ν of desired degree (fitting linear, quadratic or cubic polynomials will correspond to MFDFA-1, MFDFA-2 or MFDFA-3 respectively) . These polynomial are then taken away from each segment. This process eliminates the local trend from our data set.

Least-square variance is defined as follows :

$$F_s^2(\nu) = \frac{1}{s}\sum_{i=1}^{s} [Y((\nu-1)s+i) - P_\nu(i)]^2 \qquad \nu = 1, \ldots ,N_s$$

this is a measure for fluctuation of $Y_s(i)$

The generalized q^{th} order fluctuation function is defined by Average over all segments :

$$F_q(s) = \{ \frac{1}{2N_s}\sum_{\nu=1}^{2N_s} [F_s^2(\nu)]^{\frac{q}{2}} \}^{\frac{1}{q}}$$

$F_q(s) \sim s^{h(q)}$ where $h(q)$ is the generalized Hurst exponent.

for $q=2$, $F_2(s)$ is the standard fluctuation function used in standard DFA method. And in this case, in $F_2(s) \sim s^{h(2)}$, for a stationary process $h(2)=H$ is the Hurst exponent. It can be shown that for a non-stationary process $h(2) > 1$ and the Hurst exponent is $H=h(2)-1$.

q can be any real number. But for $q=0$ the fluctuation function is computed using logarithmic averaging process :

$$F_0(s) = \exp\{ \frac{1}{4N_s}\sum_{\nu=1}^{2N_s} \ln(F_s(\nu)) \}$$

In the case of a monofractal with long-range correlation, we have a single Hurst exponent and $h(q)$ is independent of q . But for a multifractal two or more Hurst exponents are present and $h(q)$ is dependent on q .

If we consider positive values of q , the segments ν with large variance $F^2(s,\nu)$ (i.e. large deviations from the corresponding fit) will dominate the average $F_q(s)$. Thus, for positive values of q ; $h(q)$ describes the scaling behavior of the segments with large fluctuations. for negative values of q , the segments ν with small variance $F^2(s,\nu)$ will dominate the average

$F_q(s)$. Hence, for negative values of q , $h(q)$ describes the scaling behavior of the segments with small fluctuations.

$\tau(q)$ is the classical multifractal scaling exponent or Renyi exponent.

$$\tau(q) = qh(q) - 1$$

$D(q)$, multifractal dimensions is defined as

$$D(q) = \frac{\tau(q)}{q-1}$$

$D(0)$ gives us the standard fractal dimension.

Another useful parameter to characterize multifractal properties is $f(\alpha)$. $f(\alpha)$, the singularity spectrum, is the Legendre transformation of $\tau(q)$.

$$f(\alpha) = q\alpha - \tau(q)$$

or in terms of $h(q)$: $f(\alpha) = q(\ \alpha - h(q)\ ) + 1$

where : $\alpha = \tau'(q)$

α is the singularity strength or Holder exponent.

because in the multifractal case, the different parts of the structure give different values of α , spectrum $f(\alpha)$ exists for multifractal time series. While a Holder exponent denotes monofractality.

$$\alpha = \frac{d\tau(q)}{dq}$$

Which in numerical computation translates as : $\alpha(i) = [\tau(i+1) - \tau(i)]\ /\ [q(i+1) - q(i)]$

Surrogated Data

As well as the original data, MFDFA is also applied to rank-wised and random-phase surrogated data. The aim for these analysis is to determine the origin of multifractality. There are two signs of multifractality for a time series. The first being the existence of fat-tailed probability density function, and second one , existence of linear and non-linear correlation in the time series.

Rank-wised surrogate changes the PDF of a time series into a Gaussian distribution, in other words it removes non-Gaussian distribution effects, but keeps the linear and non-linear correlation of the

original time series. Considering these properties of rank-wised surrogate, this transformation of data will eliminate the effect of a fat-tailed distribution on multifractality.
The method is that we take a random Gaussian data set with same size as our original data and sort it so that the order of numbers is the same as original time series.

Random-phase surrogate also eliminates the effect of fat-tailed PDF (changing it to a Gaussian PDF), while keeping linear correlation of the original time series. But it also eliminates the effect of non-linear correlation. In this method we apply Fourier transform to our time series. Keeping the absolute value of it, we give random phase to the data (phases coming from a uniform distribution). Finally an inverse Fourier transform is applied.

Results

we calculated $F_q(s)$ for the given time series. The scaling behavior and non-stationary properties of this time series can be investigated base on $F_q(s)$. the log-log plot of $F_q(s)$ versus s for some different q values is shown. (figure 1)

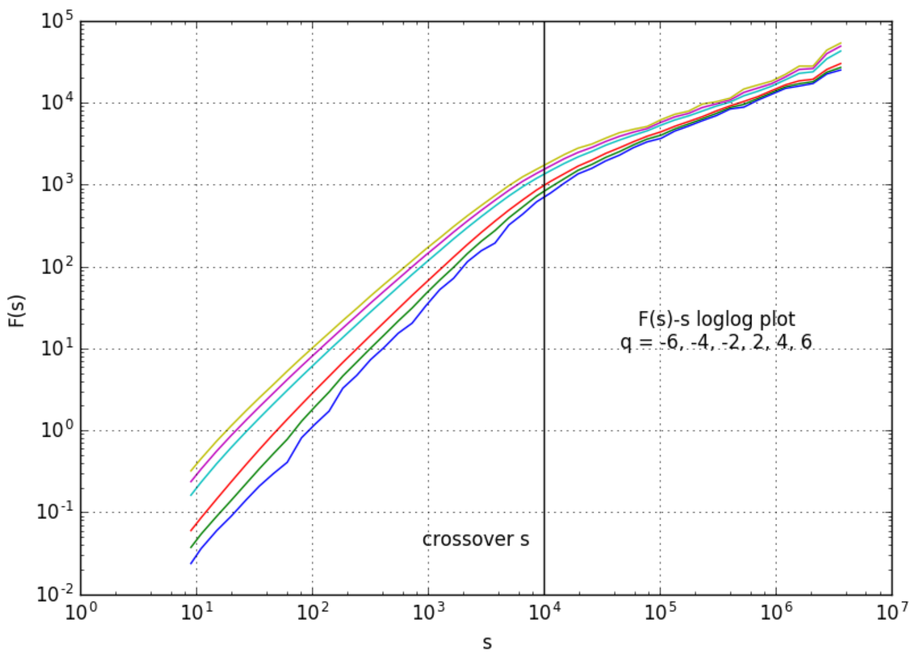

figure 1: log-log graph for $F(s)$ versus s for different q s. crossover time scale at $s=10^4$

In our calculations s is in the range of $9 - 6.3*10^6$ and changes logarithmically. q changes between -6 to $+6$ with steps equal to 0.2 .

it is clear from the graphs that there is a crossover scale (s_c) at about 10^4 . it means that there are two regions that we must discuss multifractality and correlation properties. The dominant fluctuations behave differently for different scales. From this graph we can also predict that multifractality properties are mostly for the first region ($s<s_c$).

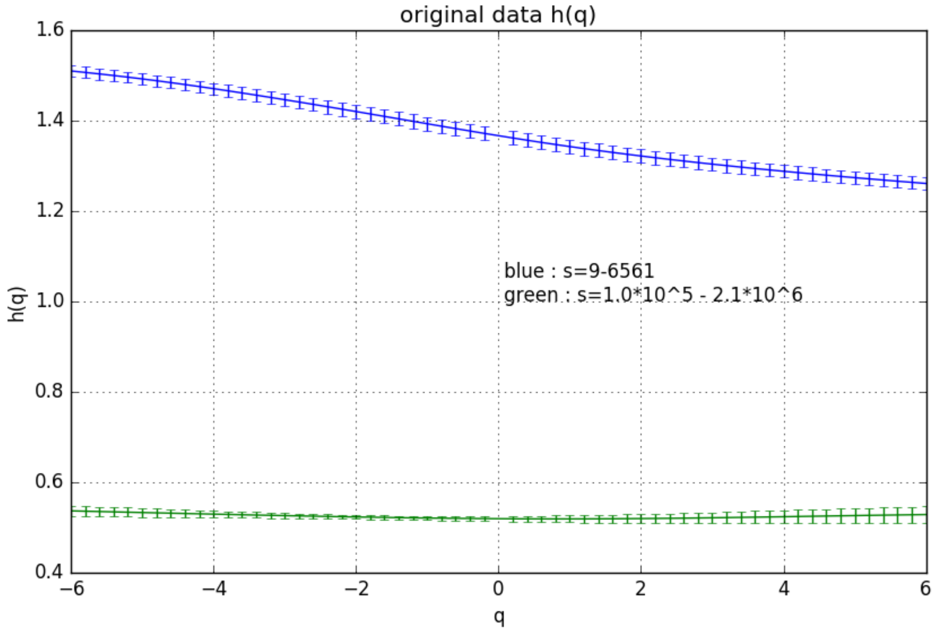

figure 2: comparison of $h(q)$ *for two regions (* $s<s_c$ *and* $s>s_c$ *)*

For s smaller than crossover time scale, $h(q)$ is clearly more q dependent and also takes values larger than 1 which is a sign of non-stationarity. But for s larger than crossover time scale $h(q)$ is close to being q independent and takes values close to 0.52 . (keeping in mind that Hurst exponent for a completely uncorrelated series i.e. random walk, is 0.5).

$F_q(s)$ Versus s for $q=2$ with fitted line for two regions is plotted (figure 3). The slope of these lines is actually the generalized Hurst exponent $h(q)$ for the this value of q . we can see that for the first region $s<s_c$, $h_1(2) = 1.322 \pm 0.15$ which is greater than 1 and therefore shows non-stationary characteristics for this region. For $s>s_c$, $h_2(2) = 0.520 \pm 0.10$ which is smaller than 1. for time scales larger than s_c the time series almost shows monofractal properties. in the first region we can obviously see the difference between slopes for different q values. That is exactly stating the dependence of $h(q)$ to q . we mostly focus on this region which reveal the multifractal behavior of the time series.

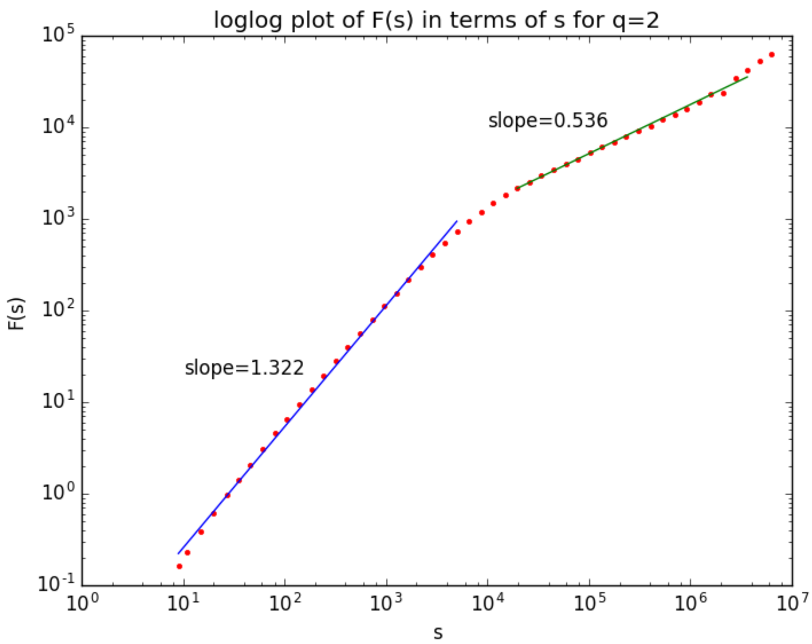

figure 3: log-log graph of F(s) versus s for q=2. Slope for each region is the Hurst exponent of that region.

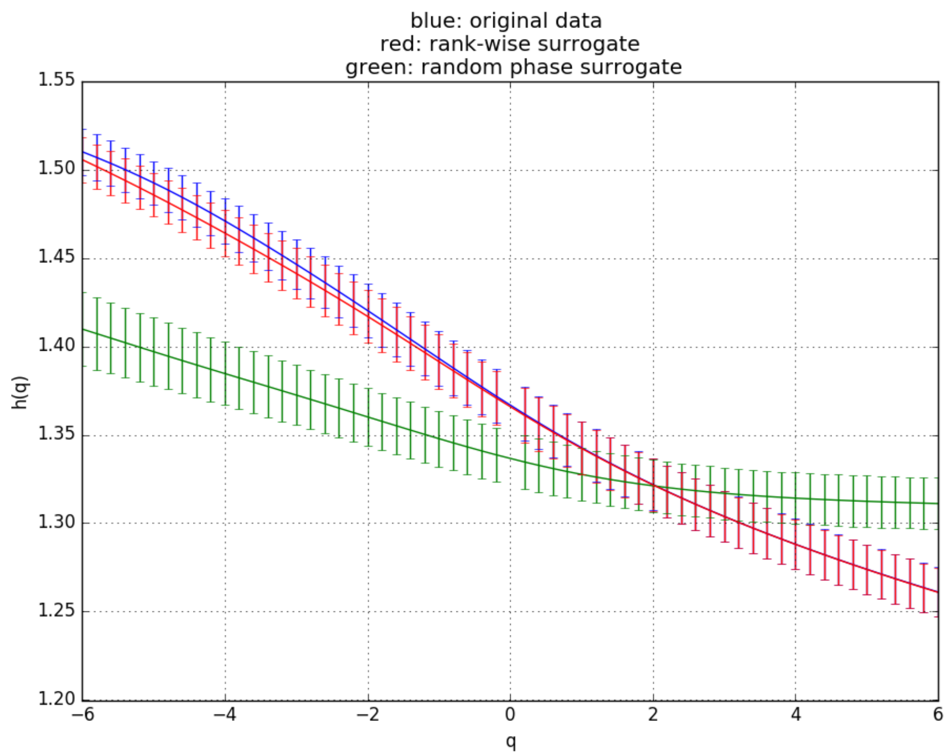

figure 4: $h(q)$ *versus* q *comparison of original data, Rank-wised data and Random-phase data*

The width of singularity spectrum , $\Delta\alpha = \alpha_{max} - \alpha_{min}$, and the range of $h(q)$ are indicators of the degree of multifractality as well. In figure 5 $f(\alpha)$ is compared for original data, Rank-wise surrogated data and Random-phase surrogated data. Original data and Rank-wise surrogated data almost fit which means fatness of PDF is not the reason behind multifractality here. But Random-phase surrogated data clearly has smaller $\Delta\alpha$, which corresponds to weaker multifractality. The parameter that this surrogate does not keep, comparing to Rank-wise surrogate, Is non-linear correlation.

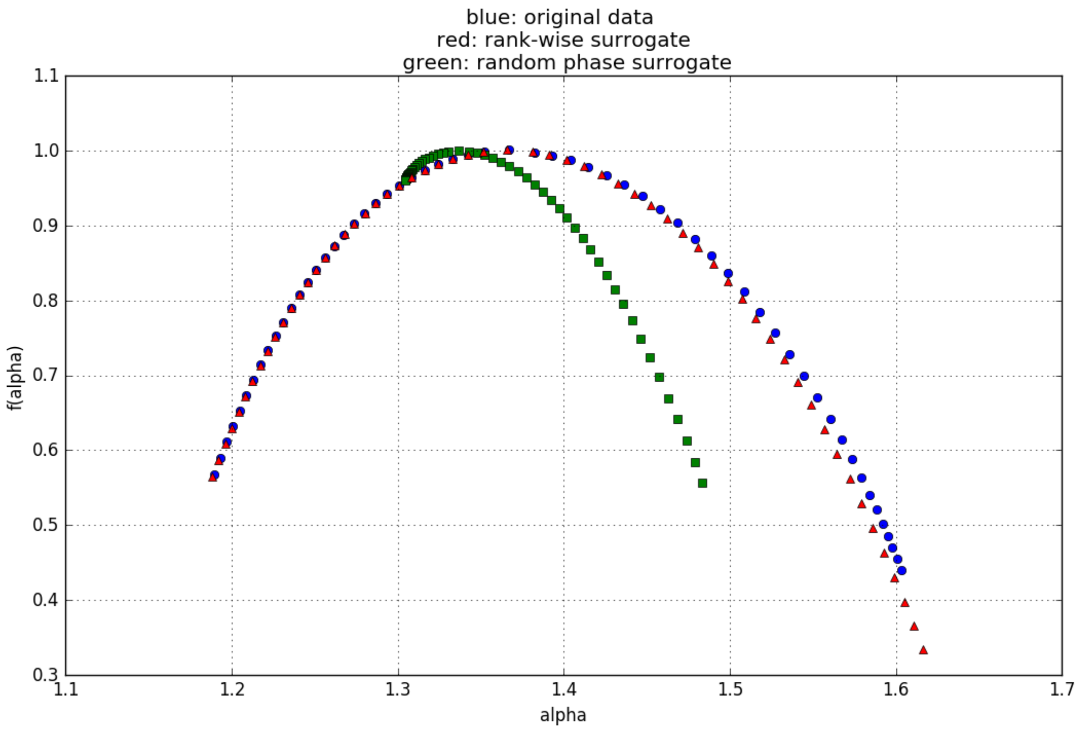

figure 5: singularity spectrum $f(\alpha)$ versus α comparison of original data, Rank-wised data and Random-phase data

data	$h_1(2)\quad(s<s_c)$	$h_2(2)\quad(s>s_c)$	$\tau(q=2)\quad(s<s_c)$	$\Delta\alpha$
Original data	1.322 ± 0.015	0.520 ± 0.010	1.644 ± 0.029	0.4141
Rank-wised	1.321 ± 0.015	0.517 ± 0.013	1.642 ± 0.029	0.1791
Random-phase	1.322 ± 0.015	0.520 ± 0.010	1.643 ± 0.029	0.4281

Hurst exponent for $s<s_c$ is $h(q)-1$ .for example for $q=2$ in original data $H=0.322$.
in $q=2$ original data and surrogates agree due to having the same linear correlation.

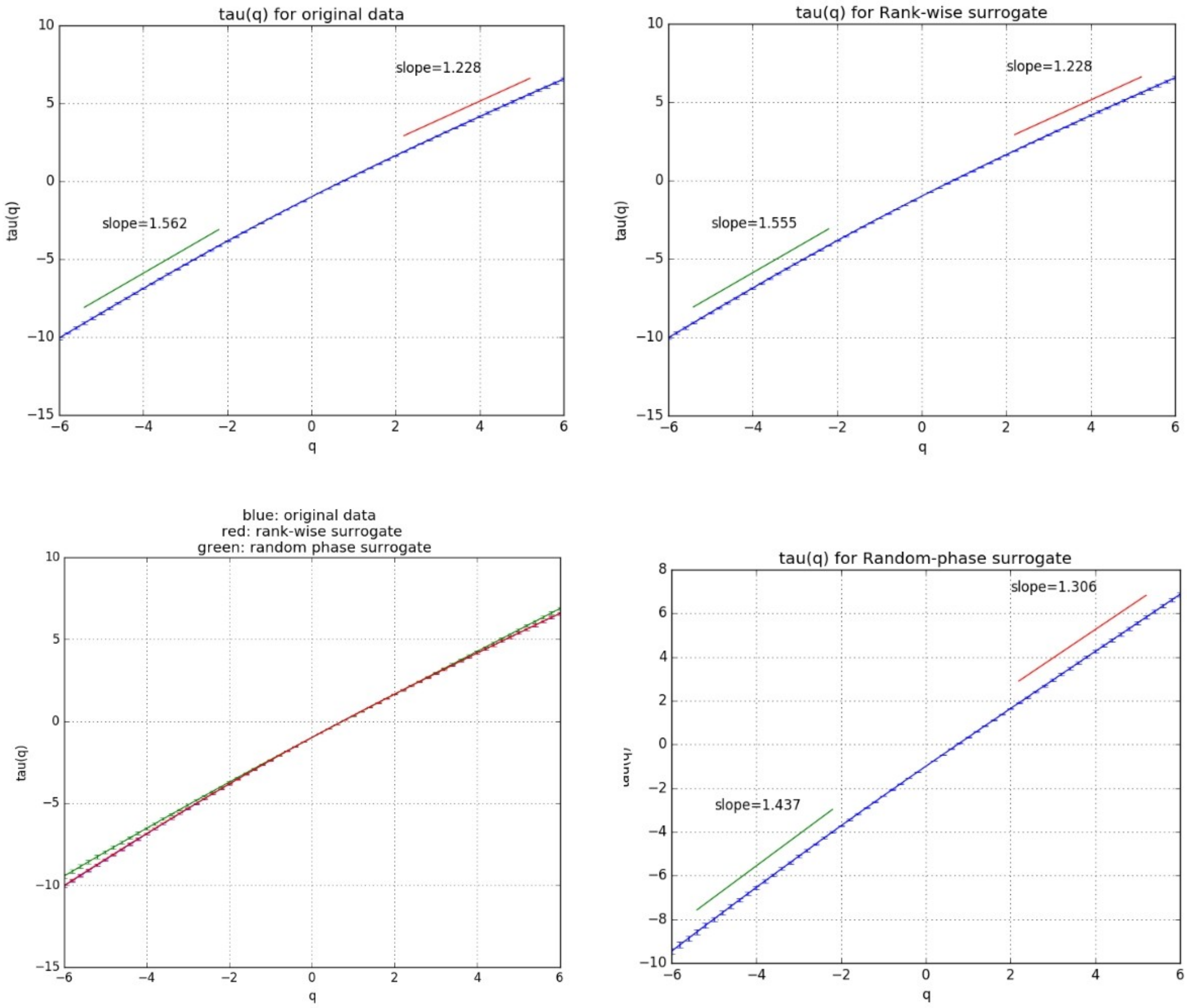

figure 6: $\tau(q)$ *versus* q *comparison of original data, Rank-wised data and Random-phase data*

The difference of slope for two regions of $\tau(q)$, for positive and negative q values show the degree of non-linearity. From the following table it is clear than only Random-phase surrogate decreased the degree of non-linearity considerably.

data	$\tau(q)$ Slope for $q<0$	$\tau(q)$ Slope for $q>0$	Slope difference
Original data	1.562	1.228	0.334
Random-phase	1.437	1.306	0.131
Rank-wised	1.555	1.228	0.327

Original time series and Rank-wised surrogated data behave very similarly in $h(q)$, $\tau(q)$ and $D(q)$. from this it can be concluded that the origin of multifractality in this time series is mostly due to the non-linear correlation. Rank-wise surrogate which only changed the PDF of our data didn’t change the multifractality parameters, but Random-phase surrogate which only kept linear correlation did make change in those parameters. q Dependence of $h(q)$ is clearly less for Random-phase surrogate as well as $\tau(q)$ being closer to a straight line (less difference between slope of two fitted line for negative and positive q values)

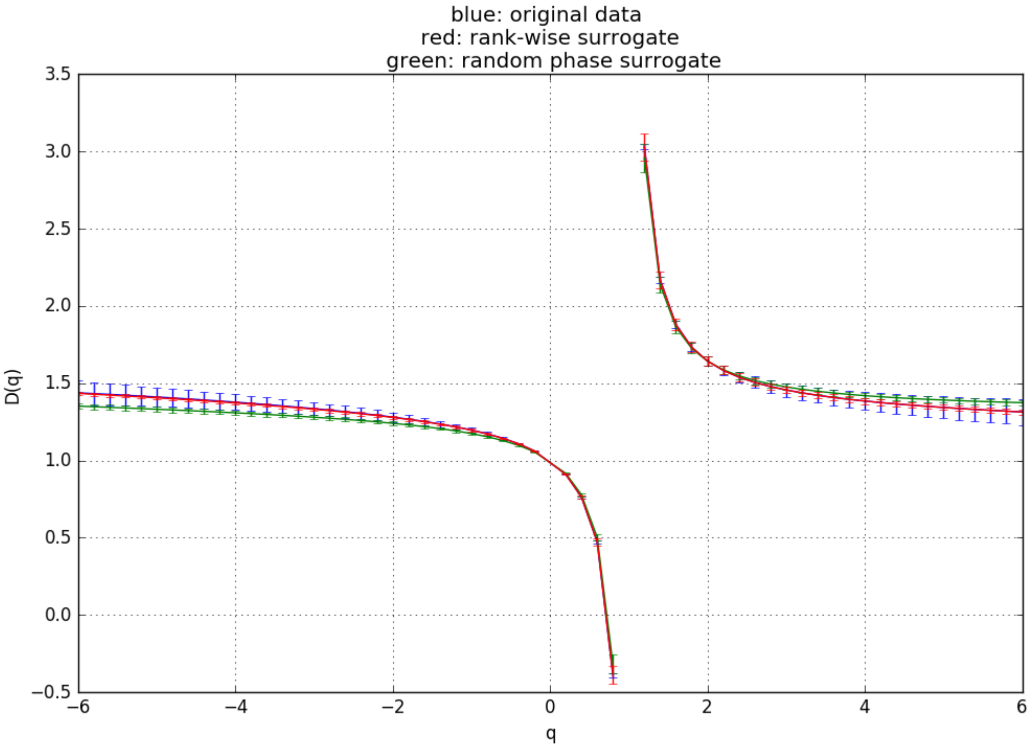

figure 7: $D(q)$ *versus* q *comparison of original data, Rank-wised data and Random-phase data.* $D(0)=1$ *, the standard fractal dimension, as expected*

conclusion

Multifractal Detrended Fluctuation Analysis (MFDFA) is is proven to be useful in determining multifractal properties of a time series. Using this method we found out that the analyzed time series has one crossover time scale. The q dependency of $h(q)$ and $\tau(q)$ proved the multifractality of this time series. Comparing the generalized Hurst exponent of the original time series with Rank-wised and Random-phase surrogated data, we found that multifractality here is mostly due to the non-linear correlation, linear correlation and broadness of the probability density function have less contribution. Investigation of $f(\alpha)$ and $\Delta\alpha$ supported this claim as well.

Error calculations

Main error estimation in our computation arises when we fit a line to two sections ($s<s_c$ or $s>s_c$) of $F(s)$ versus s log-log plot.

$y=mx+c$

having x_i y_i as two sets of points.

covariance matrix : $cov = \begin{bmatrix} cov(x,x) & cov(x,y) \\ cov(y,x) & cov(y,y) \end{bmatrix}$

it can be shown that $m = \dfrac{cov(x,y)}{cov(x,x)}$

s is defined : $s = \sqrt{\dfrac{cov(y,y) - \dfrac{cov(x,y)^2}{cov(x,x)}}{n-2}}$

Em is the error for slope of the fitted line

$$Em = \frac{s}{\sqrt{cov(x,x)}}$$

for MFDFA method this would be the error for $h(q)$ and error for other parameters can be calculated from that.

$$\Delta h(q) = Em$$

$$\Delta\tau(q) = q \ \Delta h(q)$$

$$\Delta D(q) = \frac{\Delta\tau(q)}{q-1}$$

References :

- J. W. Kantelhardt, S. A. Zschiegner, Eva Koscielny-Bunde, S. Havlin, A. Bunde and H. E. Stanley, Physica A316, 87-114 (2002).

- M. Sadegh Movahed, F. Ghasemi, Sohrab Rahvar, M.Reza Rahimi Tabar, arXiv:astro-ph/0602461.

- A. Madanchi, M. Absalan, G. Lohmann, M. Anvari and M.R. Rahimi Tabar – PACS number: 05.45.Tp, 05.45.Df, 96.60.Ub

- http://mathworld.wolfram.com/LeastSquaresFitting.html