Power Coefficient – A Non-parametric Indicator for Measuring the Time Series Dynamics

Branko Pecar

University of Gloucestershire

Abstract

Time series resume various shapes and exhibit varying dynamics when presented as graphs. Partially, this dynamic is measured via the Hurst exponent and a related metric, the fractal dimension. Both metrics are not sufficiently sensitive to discriminate variations in time series dynamics. To address this need, a non parametric indicator, called the Power Coefficient (the PC coefficient) is suggested. The coefficient was calculated for a number of artificially generated and the real life time series to identify how selective it is. The results indicate that, although it is sensitive to the length of the time series, it is far superior to the existing metrics.

Keywords: Time series analysis, dynamics metrics, the Hurst exponent, the fractal dimension, the Power Coefficient, the PC coefficient

1 Introduction

A time series is a snapshot, contained in a time window, of the historical behavioural pattern of the observed variable. Various methods are used as descriptive and predictive tools related to time series, and we do not intend to focus on them (for further discussion, see Mills [1], Brockwell and Davis [2], Dunis [3], Enders [4], Kantz and Schreiber [5]). In order to apply a particular method, certain assumptions need to be made. These assumptions are explicitly, or implicitly, linked with the properties of the time series. The usual tests, or assumptions made, are mainly related to the linearity issues, stationarity, serial correlation, process memory and other similar assumptions. Nevertheless, it remains the fact that the general appearance of the time series will influence the choice of the method. The question associated with this is: how do we measure the appearance of the time series? This paper will focus on visual properties of time series and will attempt to classify them in accordance with an indicator that was specifically designed for such a purpose.

2 Background

Any conventional textbook in quantitative methods provides a variety of descriptive statistics that characterise a data set and which can, by default, be applied to time series, which are just one specific form of data sets. However, most of these statistics (the mean, variance, skewness, kurtosis, etc.) are fixed and applicable to stationary time series. The majority of the real-life time series fall into a non-stationary category, rendering effectively these descriptive statistics of little value. However, several other attempts were made to encapsulate complex time series behaviour into a single coefficient. One of the most successful attempts can be linked with the work H. E. Hurst conducted towards the middle of the last century in Egypt.

In 1951 Hurst [6] published a seminal paper in which he tried to establish the correct storage required in the Great Lakes of the Nile River Basin. His objective was to determine the capacity of a reservoir guaranteeing a minimum discharge when the water intake is low. To determine the optimum capacity, he examined the maxima and minima of a number of reservoirs over a number of periods and concluded that there is an optimum range required to maintain the average discharge. When this range was measured against different time horizons, he discovered strong scaling properties. This scaling property is estimated as the slope of the line on a log plot showing the changes in the rescaled range versus the number of cases used for the range (Range is a function of deviations from the mean. Rescaling the ranges means that they are divided by their appropriate standard deviation, i.e. R/σ.). Contrary to his expectation that this slope should be equivalent to 0.5 (which would be normal for random events), he discovered that the number is around 0.7 (to be precise it is 0.73). In other words, the scaling of certain natural phenomena (reservoirs included) does not follow chance, but shows an increasing amount of persistence and long process memory.

Benoit Mandelbrot [7] took Hurst's ideas as a foundation for making a further breakthrough in 1965. He renamed the rescale analysis exponent to the Hurst exponent and changed its original notation from K, as used by Hurst, to H. More importantly, he linked it with the fractal dimension and established that they are related as $d_F = 2 - H$, where d_f is the fractal dimension.

Mandelbrot established that (in particular for self-afine data sets) the smaller the value of the Hurst exponent, the rougher the surface of the curve describing the data set. As the Hurst exponent takes values $0<H<1$, this implies that for H close to zero, the series will closely resemble a white noise process, whilst for H close to unity, the series will be still jumping up and down, but with the element of persistence built into it. The Hurst exponent, therefore, indicates both how long the series memory is and how rough the series is. As Mandelbrot [8] put it "… fractional noises with a high value of H are the most violently fluctuating among fractional noises.". Until today the Hurst exponent and the fractal dimension remain the only two complementary methods of measuring the roughness of the data set and its capacity to fill in the plane.

Although Mandelbrot provided the relationship to estimate the Hurst exponent from the fractal dimension (usually it is other way round), it is also possible to infer the Hurst exponent from the spectral density exponent β. Spectral densities S_x have the approximate value of:

$$S_x \propto \frac{1}{f^{\beta}} \qquad (1)$$

Where, *f* are the frequencies and the exponent $\beta = 2H + 1$, provides a link with the Hurst exponent. However, there are several problems with calculating the Hurst exponent. First of all, the calculation algorithms are quite time-consuming and not so easy to implement without introducing some complex looping in the code. A very good freeware, called Fractan, for calculating several invariant measures, including the Hurst exponent, was created by Vyacheslav Sychyov and it is available from: http://impb.psn.ru/~sychyov/html/cv_e.shtml. The second point is that the Hurst exponent is not calculated explicitly, it is estimated. This fundamentally can cause some difficulties and misinterpretations.

As the Hurst exponent is one of the ways of estimating the fractal dimension d_F, we can say that essentially neither the Hurst exponent nor the fractal dimension discriminate enough the appearances of the time series. Essentially, as stated by Mandelbrot [7], for time series it is expected that the fractal dimension d_F will take values between one and two. The closer d_F is to one (i.e. the closer H to one is), the smoother the line. Conversely, the closer d_F is to two (i.e. the closer H to zero is), the more dynamic the line is. If the value of $d_F = 2$, we say that the time series is equivalent to the white noise process. For the value of $d_F=1.5$ we have a special case called ordinary Brownian motion. Any other value of d_F indicates fractional Brownian motion, i.e. for the values of $1.5<d_F<2$, the process is anti-persistent and for the values of $1<d_F<1.5$, the process is persistent, exhibiting long memory (Persistent process implies autocorrelation function following exponential decay, whilst anti-persistent process implies negative correlation between observations, i.e. every subsequent observation is more likely to go in the opposite direction from the previous one rather than follow the trend).

These concepts can also be translated into a visual appearance of a time series. The closer the value of H to 1 and the closer the value of d_F to 1, the less dynamic (erratic) the time series. The closer the value of H to 0 and the closer the value of d_F to 2, the more dynamic the time series and the more 'jumpy' it looks.

However, even some elementary experimentation can quickly demonstrate that the Hurst exponent and the fractal dimension are not sensitive enough and discriminating enough, unless we handle random walks. In addition to this, the value of the Hurst exponent has meaningful interpretation only for the random processes. For any other type of series it is inconsistent and ambiguous.

Although both coefficients (the Hurst exponent and the fractal dimension) can be used to describe some elementary properties of time series, this would constitute the most elementary feature descriptor if a time series was used as an element in some sort of repository, such as a Case Based Reasoning (CBR) system. Unfortunately, the emphasis should be on the word "elementary", as it is impossible to classify time series precisely in accordance with different values of the Hurst exponent. The case features have to be of higher resolution in order to secure better sensitivity and diagnostic/predictive properties. This paper will explore an alternative, a simple and non-parametric method for measuring the roughness of a time series.

3 The Power Coefficient (PC)

To suggest a measure of a time series appearance, we start with intuitive reasoning. In a time series the first and the last observation in a data set constitute the boundaries of the time window, which characterises certain behaviour of the data set. The shortest distance between these two points is a smooth, straight line. However, in reality a time series will exhibit some dynamics in this observed time window and will seldom follow the straight line. Very often, the first and the last observation have no effect on the overall behavioural pattern of the variable in this time window. It is more likely that, moving from period to period, the actual observations will deviate from the shortest path and will, in a way, wander or rumble through the space, creating a wiggly curve that characterises the behaviour of this variable in a given time window. The amount of curve wiggling will, in a way, show the 'energy' of the time series.

As the curve wanders, the implication is that the observations will travel certain distance which will inevitably be greater that the shortest path that connects the end points of the time series. The total distance that the variable will travel in a time window is the cumulative value of all the observation differences. Dividing it by the number of observations in a data set will give us the average travel distance between any two periods for this particular time window.

We can also observe that in a time series, at least one observation will inevitably be the maximum value for the given time window. Equally, at least one observation will be the minimum value, unless we deal with a horizontal line, in which case both the max and the min are the same. If a data set displays maximum level of dynamics for the given time window, then it would expect it to 'jump' from the minimum to maximum and back between any two periods in the series. In other words, by jumping up and down from minimum to maximum, the series would display maximum possible dynamics, or, maximum amount of energy embedded in its character. This means that the difference between the maximum and minimum value of the time series is a hypothetical average distance that this time series would travel if it had the most dynamic history.

These two values, the actual average distance and the hypothetical average distance, can be used to calculate a measure of energy the time series exhibits, and we will, therefore, name it tentatively as the Power Coefficient (PC). This coefficient is simply calculated as:

$$PC = \frac{\left| \frac{\sum_{i=1}^{k} m_i}{n} \right|}{\left| x_{max} - x_{min} \right| + \varepsilon} \times 100 \qquad (2)$$

Where,
x_i = actual observation
m_i = distance between two observations i.e. observation difference $m_i = x_i - x_{i-1}$
x_{min} = minimum value in the time series
x_{max} = maximum value in the time series
ε = arbitrarily small constant
k = number of observations in an interval, $2 < k < n$, (some reasonable number)
n = total number of observations

The Power Coefficient (PC) will take various values for different patterns and theoretically if a variable oscillated from minimum to maximum between any two periods, its maximum value is 100. On the other hand, zero is the minimum value that the PC will show, which applies to horizontal straight line only. Strictly speaking this is not true because in this case both max and min value are the same, yielding ∞. To avoid this inconvenience, present only in exceptionally small number of cases, an arbitrarily and infinitesimally small constant ε was introduced in the denominator.

In order to demonstrate how the value of the PC changes as the dynamics of the time series changes, we conducted experiments with 41 different time series. The list of all 41 time series and their PC values are given in Table 1. However, before we draw any conclusions, we need explain how some of the series were generated.

4 Data sets and results

The PC coefficient was experimentally applied to a number of time series. The experiments involved 41 different time series, such as: smooth curves, chaotic attractors, artificially generated random walks, periodic curves (sinusoids, see-saw curves and Wierestrasse curve), periodic curves with added random elements, stationary random processes (the white noise and ARCH) and the actual time series from the New York Stock Exchange (NYSE).

Four companies were arbitrarily picked from the NYSE, and they are BP, IBM, Coca Cola (KO) and Pfizer (PFE). Their daily and minute closing values were captured, varying in appearance from the white noise look to a random walk type of behaviour. Their returns for both daily and minute stock values were also calculated, generating time series that resemble either a white noise, or an ARCH process. The returns on closing values of stocks were calculated as:

$$R_t = \log X_{t+\Delta t} - \log X_t \qquad (3)$$

Where, X_t are stock values (either daily or minute values) and Δt is the time difference (a lag). This implies that for $\Delta t = 1$, i.e. for lag 1, we get short term returns (i.e. returns after one minute or one day, depending on the resolution of the data set) and for larger number of lags, some longer return value is obtained.

As a reference point, the Hurst exponents were also estimated for all 41 data sets and their fractal dimensions inferred. Essentially, the fractal dimensions d_F fell into three characteristic regimes:

1) d_F around 1 – All artificially generated random walks, all stock closing values (both daily and minute) and most of the smooth curves

2) d_F around 1.5 – All stock returns (both daily and minute), Lorenz attractor, one smooth curve, see-saw curve with added random component, sunspot numbers, the Wierestrass curve and the white noise process

3) d_F around 2 – All the sinusoids, MinMax series, ARCH process, Rossler and Henon attractor

Visual inspection of the graphs depicting the data sets that fall in the same d_F category indicates that the fractal dimension (or the Hurst exponent) does not discriminate these sufficiently and that it does not provide meaningful clues about the fundamental features of every time series, to say the least. The PC values, on the other hand, provide much better resolution for this problem.

The results in Table 1 indicate certain tendencies, and we have sorted the values in the ascending order of the PC coefficient. For all the smooth curves, the value of the PC is always 0.1 (for the horizontal line the value is zero and for all other smooth curves the part behind the decimal point depends on the value of the constant ε), and for the series number 41, which oscillates between its maximum and its minimum value every two periods, the PC is 100, as expected.

Table 1: Calculation of the PC values
for a number of time series

No	Series	Observations	PC	No	Series	Observations	PC
1	Horizontal Line	1000	0.000	21	Weierestrass curve	6001	2.038
2	Linear Nonstationary Trenc	1000	0.100	22	See-Saw nonstationary	114	2.564
3	Power curve	1000	0.100	23	IBM minute returns	5861	3.630
4	Expontial curve	1000	0.100	24	Coca Cola daily returns	8422	3.634
5	Logarithmic curve	1000	0.100	25	IBM daily returns	10404	4.279
6	Coca Cola daily stock valu	8423	0.257	26	Pfeizer minute returns	5472	4.568
7	IBM daily stock values	10405	0.306	27	Sunspots avg monthly nur	3042	4.840
8	BP daily stock values	6653	0.391	28	Low frequency (LF) sinus	873	5.561
9	Coca Cola minute stock va	5473	0.398	29	ARCH	5200	6.854
10	Pfeizer daily stock values	5391	0.429	30	Rossler attractor	1000	6.896
11	Random walk 3	8000	0.474	31	Pfeizer daily returns	5390	6.954
12	Random walk 5	8000	0.512	32	Lorenz attractor	1000	7.821
13	Random walk 1	8000	0.589	33	BP daily returns	6652	8.030
14	Random walk 2	8000	0.593	34	See-Saw stationary	1009	11.111
15	BP minute stock values	5458	0.615	35	LF sinusoid with white noi	873	11.520
16	Pfeizer minute stock value	5473	0.687	36	HF sinusoid with white noi	873	15.358
17	US Discount rates (52.8 ye	634	0.740	37	High frequency (HF) sinus	963	16.677
18	Random walk 4	8000	0.795	38	See-Saw stationary + ranc	1009	18.507
19	BP minute returns	5457	1.509	39	White noise	1500	33.531
20	Coca Cola minute returns	5472	1.745	40	Henon attractor	1000	38.229
				41	MinMax curve	1000	100

In general, the PC coefficients show seven characteristic regimes:

A) PC close to zero – All smooth curves

B) PC between 0 and 1 – All **closing** stock values and all artificially generated random walks

C) PC between 1 and 3 – The Weirestrasse curve, nonstationary see-saw curve and two stock minute return series (BP and Coca Cola)

D) PC between 3 and 5 – Majority of the **stock return** series (Coca Cola and IBM daily returns, and IBM and Pfeizer minute returns)

E) PC between 5 and 10 – The Lorenz and Rossler attractor, ARCH process, low frequency sinusoids and two daily return stock series (BP and Pfeizer)

F) PC between 10 to 20 – Majority of sinusoids and various see-saw curves

G) PC above 30 – The white noise process and the Henon attractor. MinMax series, as an exception, has the maximum value of 100.

Table 1 shows that, for practical reasons, we were not consistent with the sample sizes of the analysed data sets and that the lengths vary dramatically. In addition to this, some inconsistencies are visible, for example, the nonstationary see-saw series seems to be in a completely different category than the other see-saw series. It is very likely that the number of observations taken into account play a major role. To address this question we briefly explored how sensitive this coefficient is to the sample size, i.e. the length of the data set.

5 Sensitivity results

To explore this issue experimentally, we used only several data sets (see Table 2) and calculated their PC values for varying sample sizes. The sample size 1 in Table 2 implies that the PC value was calculated for the full length of the given time series. The sample size 0.5 means that the series was split into two half and the average PC value was calculated on the basis of these two half's PC values. The sample size 0.2 means that the series was split into five sub-samples and the average PC value was calculated on the basis of their PC values. And finally, the sample size 0.1 means that the series was split into ten sub-samples and the average PC value was calculated on the basis of their PC values.

The factors in the table imply that as the sample size goes down, the value of the PC goes up. To use an example of the Weirestrass curve, we can see that the value of the PC for the whole data set was 2.038. The samples that are ten times smaller (column under 0.1), show the value of PC 2.133 larger than the large data set, i.e. their estimated PC is 4.349, as opposed to 2.038.

As expected, for stationary random processes (ARCH and white noise) the sample size will not affect the value of PC. On the other end of the spectrum, for processes with long memory (Wiener process, daily and minute stock values), if the sample size is one tenth of what it actually should be, then the PC value could be overestimated between 4 to 5 times. For any other process in between these two extremes (in this experiment the daily and minute returns and the Weirestrasse curve), the overestimation of the PC value in case of a sample that is 10% of the actual series, is approximately 2 to 3 times.

We recall from the d_F regimes that Rossler and Henon attractors, as well as the ARCH process, fell into the same category. By observing the visual appearance of these time series we instinctively see that this does not make sense. On the other hand according to the PC values, Henon attractor is by far the most dynamic, even more dynamic than the white noise, whilst the Lorenz and Rossler attractors are clearly in a different league with much lower level of dynamics exhibited. This is much more in line with the visual appearance of these time series.

Table 2: Factor by which the PC value changes as the sample size reduces

Series	No of observations	PC value for all observat.	Sample size			
			1	0.5	0.2	0.1
White noise	1500	33.531	1	1.000	0.985	1.001
ARCH	5200	6.854	1	1.000	1.001	1.006
Weierestrass curve	6001	2.038	1	1.000	1.995	2.133
IBM min returns	5861	3.63	1	1.242	1.525	2.635
KO daily returns	8422	3.634	1	1.535	2.354	2.965
PFE min values	5473	0.687	1	1.246	2.253	3.428
Random walk	8000	0.474	1	1.747	2.792	4.084
BP daily values	6653	0.393	1	1.934	3.365	4.842

It is interesting to see how the actual stock values series vs. stock returns series were classified. Fractal dimension d_F clearly differentiates between the return values as opposed to actual stock values. The PC does the same. In fact, the PC discriminates between three regimes of returns. Some minute returns fall into the same category as the Weirestrasse curve and nonstationary see-saw type of curves. Some daily returns are close to chaotic attractors (the PC values close to the Lorenz and Rossler attractor). A majority of both minute and daily returns fall in the category between these two regimes.

6 Conclusions

The coefficient we suggested in this paper, the PC coefficient, is a nonparametric measure of the dynamics of the time series. It discriminates different time series in accordance with the amount of distance the time series travels from first to the last observation. For nonstationary time series resembling random walks, the coefficient is very sensitive to the length of the series, i.e. number of observations. Nevertheless, the coefficient provides a much better classification properties than the Hurst exponent and/or the fractal dimension.

References

[1] Mills, T.C., *The Econometric Modelling of Financial Time Series*. 1994, Cambridge: Cambridge University Press. 247

[2] Brockwell, P.J. and R.A. Davis, *Introduction to Time Series and Forecasting*. 2002, New York: Springer-Verlag

[3] Dunis, C., *Forecasting Financial Markets*. 1996, Chichester: John Wiley. 292
[4] Enders, W., *Applied Econometric Time Series*. 2004, Honoken: John Wiley.
[5] Kantz, H. and T. Schreiber, *Nonlinear Time Series Analysis*. 2000, Cambridge: Cambridge University Press. 304
[6] Hurst, H.E., *Long-term storage capacity of reservoirs.* Transactions of the American Societ of Civil Engineers, 1951. **116**: p. 770-808
[7] Mandelbrot, B., *The Fractal Geometry of Nature*. 1983, New York: W H Freeman & Co. 468
[8] Mandelbrot, B.B., *Gaussian Self-Affinity and Fractals*. 2002, New York: Springer Verlag. 654

APPENDIX

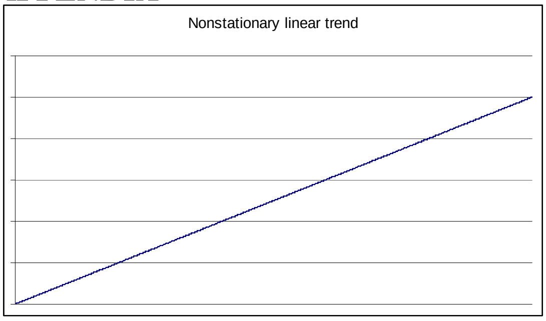

Fig 2. Nonstationary liner trend, PC=0.1

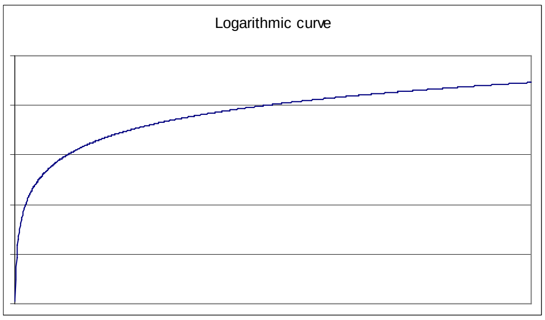

Fig 3. Log curve, PC =0.1

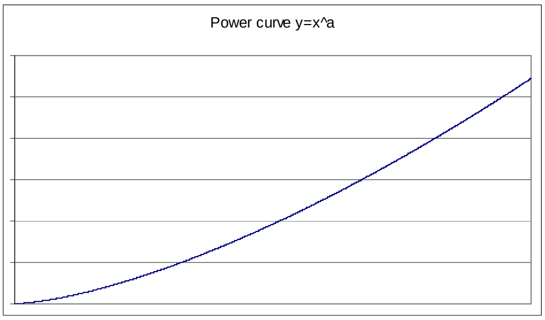

Fig 4. Parabola, PC =0.1

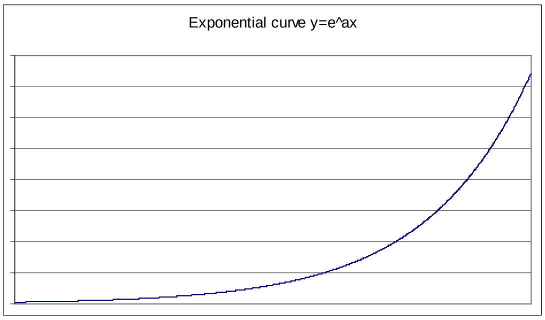

Fig 5. Exponential curve, PC =0.1

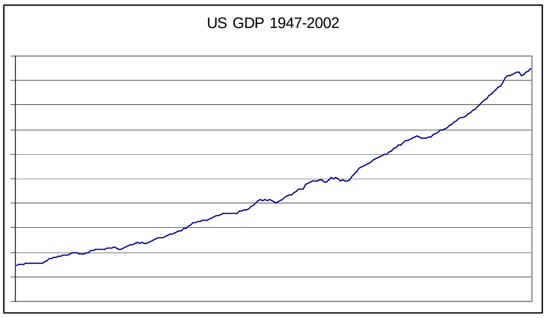

Fig 6. US GDP 1947-2002, PC =0.56

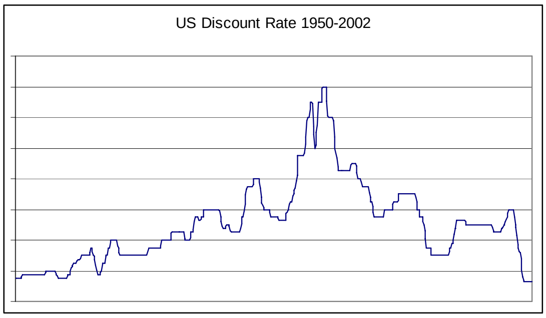

Fig 7. US Discount Rate 1950-2002, PC =0. 74

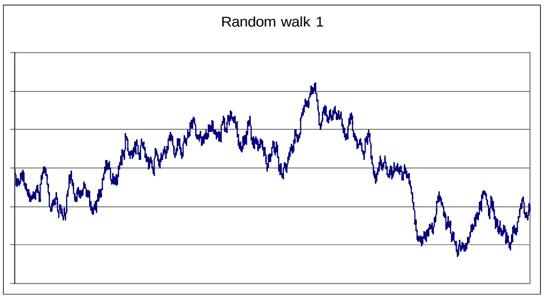

Fig 8. Random curve 1, PC =1.04

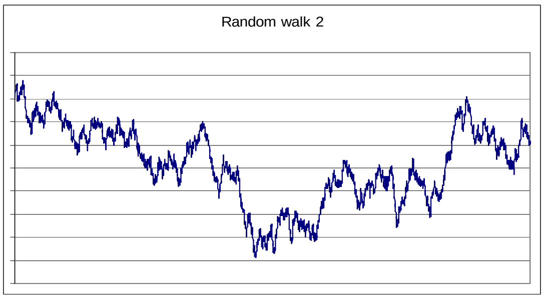

Fig 9. Random curve 2, PC =1.13

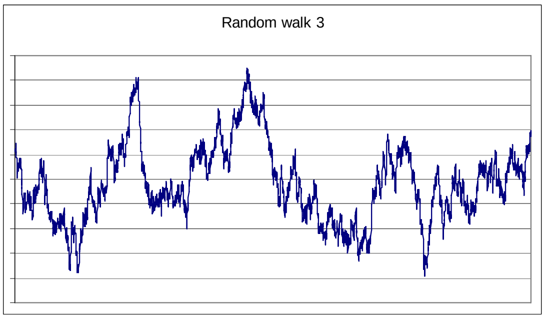

Fig 10. Random curve 3, PC =1.62

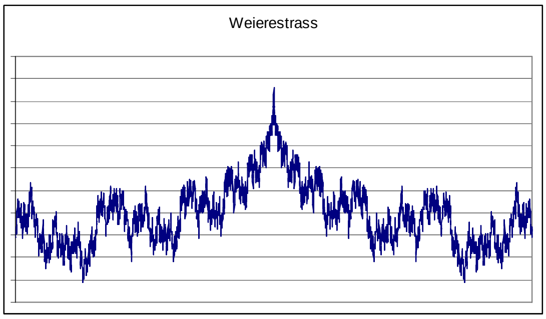

Fig 11. Weierestrasse curve, PC =2.03

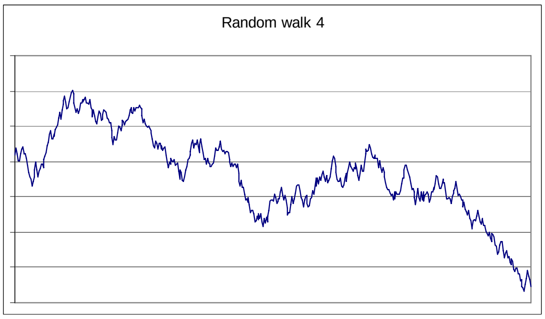

Fig 12. Random walk 4, PC =2.19

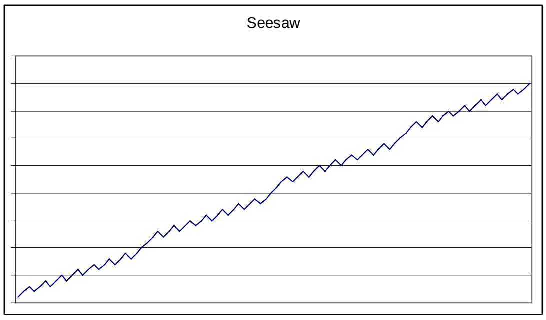

Fig 13. Nonstationary seesaw curve, PC =2.56

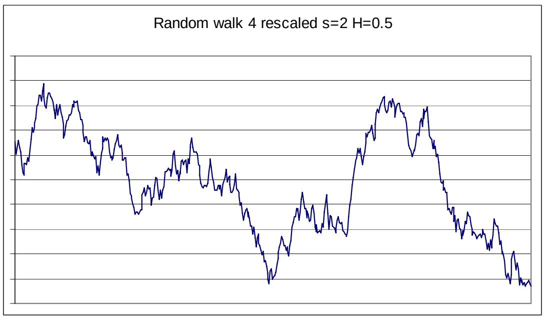

Fig 14. Rescaled random walk s=2 H=0.5, PC =6.86

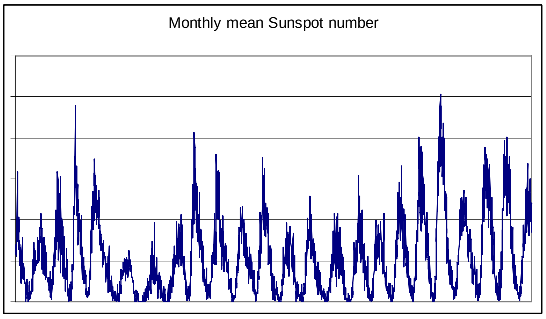

Fig 15. Monthly average sunspot number, PC =4.83

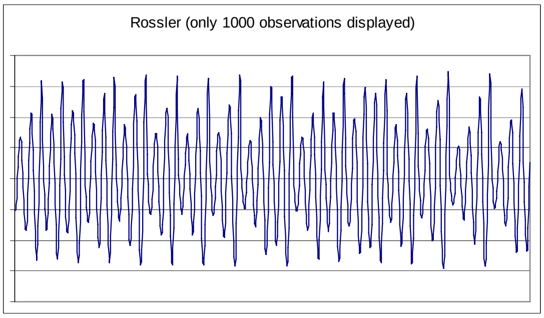

Fig 16. Rossler attractor, PC =6.86

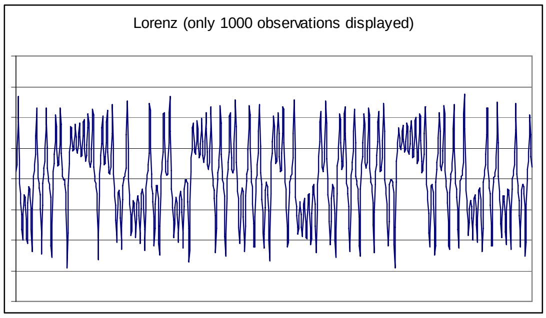

Fig 17. Lorenz attractor, PC =7.21

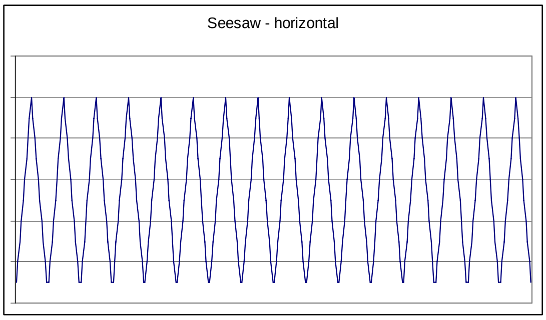

Fig 18. Stationary seesaw, PC =10.56

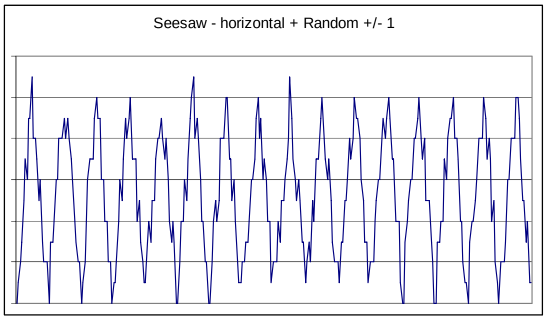

Fig 19. Stationary seesaw + random, PC =10.89

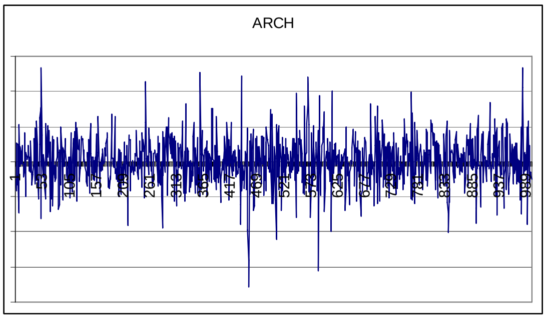

Fig 20. ARCH process, PC=11.97

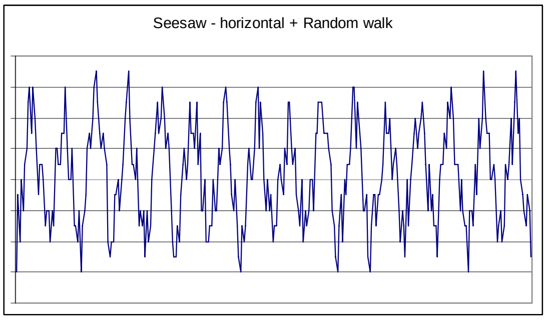

Fig 21. Stationary seesaw + random, PC =14.34

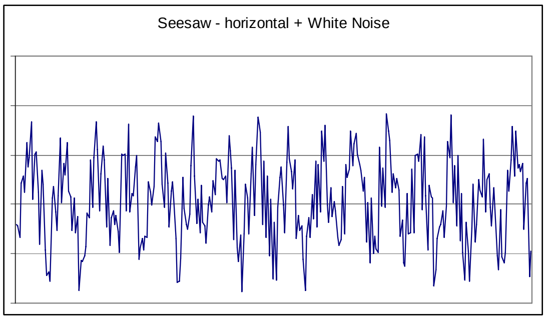

Fig 22. Stationary seesaw + white noise, PC =19.9

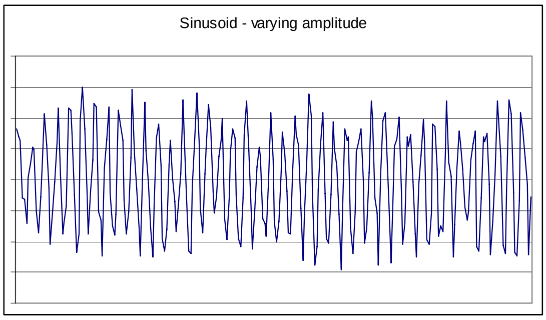

Fig 23. Sinusoid varying amplitude, PC =22.73

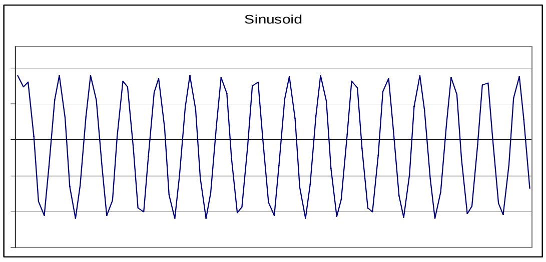

Fig 24. Sinusoid, PC =30.35

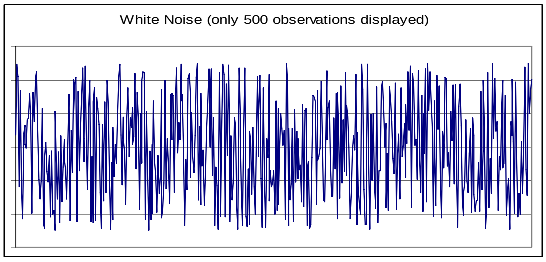

Fig 25. White noise, PC =33.26

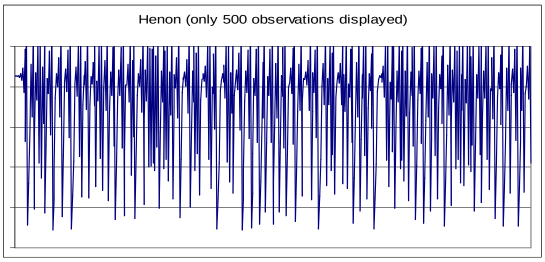

Fig 26. Henon attractor, PC =38.44

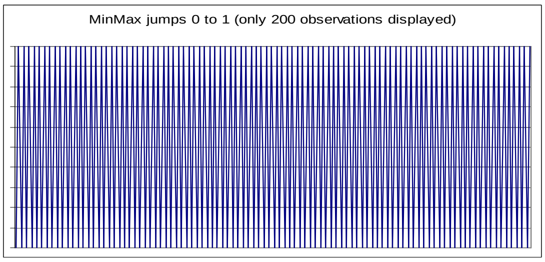

Fig 27. MinMax movements 0 to 1, PC =100.00

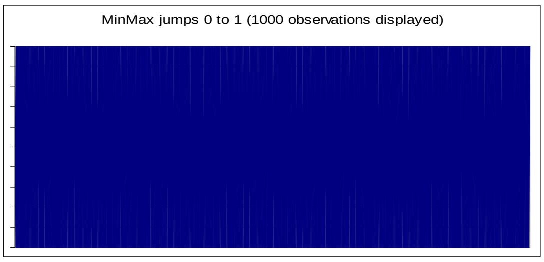

Fig 28. MinMax movements 0 to 1(1000 observations displayed)