UNIVERSITÀ DEGLI STUDI DI TRENTO

DEPARTMENT OF PHYSICS
BACHELOR DEGREE IN PHYSICS

An investigation of HURST EXPONENT

Candidate:
Alessandro Foradori

Supervisor:
Leonardo Ricci

17th November 2020

Hurst exponent

The Hurst exponent is a dimensionless estimator used to evaluate self-similarity and long-range dependence properties of time series.

$$\mathbb{E}\left[\frac{R(N)}{S(N)}\right] \propto N^H \quad \text{for } N \to \infty$$

Detrended fluctuation analysis

- C.-K. Peng et al., "Mosaic organization of dna nucleotides", Phys. Rev. E 49, 1685–1689 (1994)
- A. Foradori, *Hurst exponent*, https://github.com/f3fora/hurst-exponent, 2020

Time series

A discrete weakly stationary stochastic process $\{X_n\}_{n\in\mathbb{Z}}$ is called **time series**.

- V. Pipiras and M. S. Taqqu, *Long-range dependence and self-similarity*, Cambridge Series in Statistical and Probabilistic Mathematics (Cambridge University Press, 2017)
- B. Jan, *Statistics for long-memory processes*, Chapman and Hall/CRC monographs on statistics and applied probability (CRC Press, Boca Raton, FL, 1994)

Long-range dependence

- $\gamma_X(k) \propto k^{2d-1}$
- $\sum_k |\gamma_X(k)| = \infty$
- $\mathrm{Var}(X_1 + \cdots + X_N) \propto N^{2d+1}$

with $d \in (0, 0.5)$

Short-range dependence

$\sum_k |\gamma_X(k)| < \infty$

Antipersistance

$\sum_k |\gamma_X(k)| = 0$ and $d < 0$

Self-similarity

Intuitively, self-similarity means that a stochastic process scaled in time (that is plotted with a different time scale) looks statistically the same as the original process when properly rescaled in space.

$$\{X(ct)\}_{t\in\mathbb{R}} \stackrel{d}{=} \{c^H X(t)\}_{t\in\mathbb{R}} \quad \text{with } H > 0 \text{ and } c > 0$$

$$H = \frac{1}{2} + d$$

- $H \in (0.5, 1]$ long-range dependence
- $H = 0.5$ no time lag
- $H \in (0, 0.5)$ anti-persistent short-range dependence

The purpose of DFA is to estimate the variance of partial sums of the series $X = \{X_n\}_{n \in \mathbb{Z}}$. In this way, $\mathrm{Var}(X_1 + \cdots + X_N) \propto N^{2H}$ allows to estimate the Hurst exponent

Define the "profile" $\{Y(t)\}$ of $\{X(t)\}$.

$$Y(t) = \int_0^t dt'(X(t') - \bar{X})$$

where $\bar{X}$ is the mean of $\{X(t)\}$ computed on the whole time series.

Split the full time T in time windows τ.

Compute the linear regression $g^{(k)}(t,\tau,t')$ of order k of $\{Y(t')\}$ in range $t' \in [t, t+\tau]$.
Determinate the variance $f_2^{(k)}(t,\tau)$ between the profile $\{Y(t)\}$ and the regression $g^{(k)}(t,\tau,t')$.

$$f_2^{(k)}(t,\tau) \equiv \frac{1}{\tau}\int_t^{t+\tau} dt' \left(Y(t') - g^{(k)}(t,\tau,t')\right)^2$$

$F_q^{(k)}(\tau)$ is the momentum of order q over all time windows of length τ.

$$F_q^{(k)}(\tau) = \left(\int_0^{T-\tau} dt \left(f_2^{(k)}(t,\tau) \right)^{q/2} p(t,\tau) \right)^{1/q}$$

Determine the scaling behaviour of the fluctuation functions by analysing log–log plots of $F_q^{(k)}(\tau)$ versus τ.

$$F_q^{(k)}(\tau) \propto \tau^{\alpha(q)}$$

- $\alpha(2) \in (0, 1]$ stationary process with $\alpha(2) = H$
- $\alpha(2) \in (1, 2]$ non-stationary process with $\alpha(2) = 1 + H$

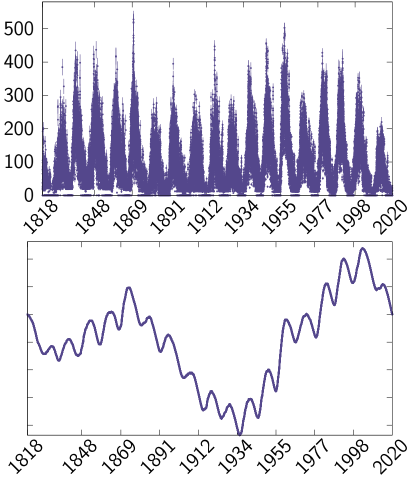
On the Sun's photosphere there is a strong magnetic field. However in some localized regions (called **sunspots**), the field is significant higher and the surface appear as spots darker than the surrounding areas. Their number varies according to the approximately 11-year solar cycle

- S. W. D. Center, "The international sunspot number", International Sunspot Number Monthly Bulletin and online catalogue (1749-2020)
- M. S. Movahed et al., "Multifractal detrended fluctuation analysis of sunspot time series", Journal of Statistical Mechanics: Theory and Experiment **2006**, P02003–P02003 (2006)

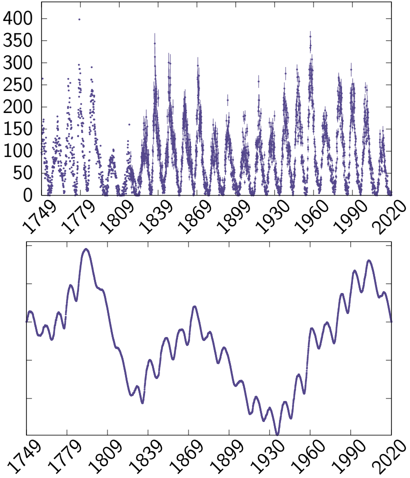
Sunspots

Raw data and profile

daily sunspot number

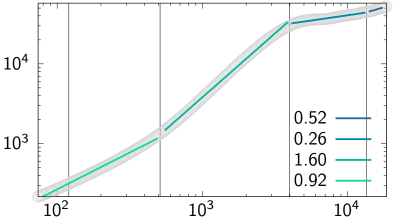


monthly mean sunspot number

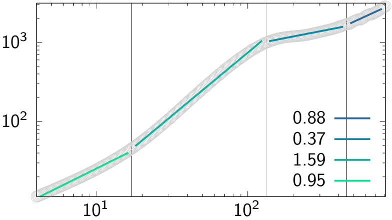


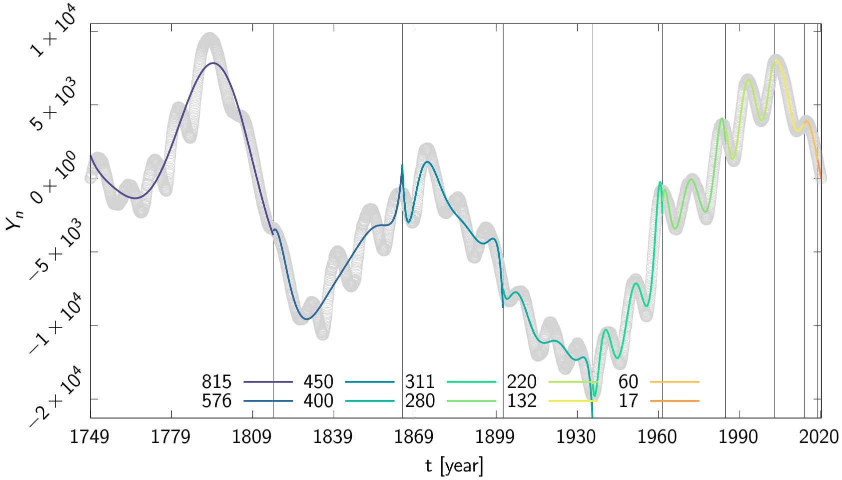
Vertical black lines at $\tau = 4, 17, 132, 450$ months.

daily sunspot number

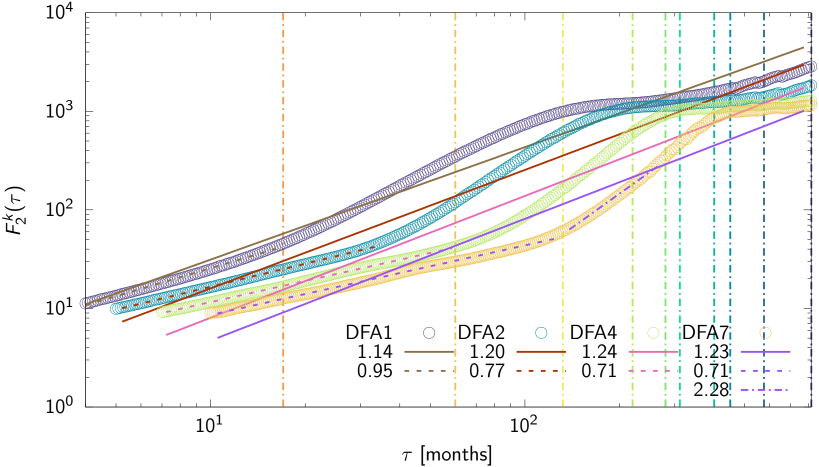


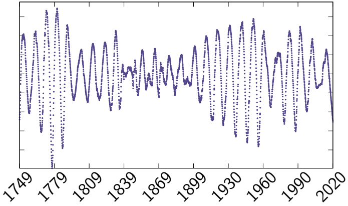
monthly mean sunspot number



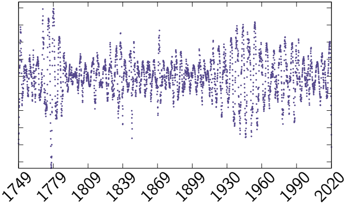


1×10^4
5×10^3
0×10^0
-5×10^3
-1×10^4
-2×10^4
Y_n
1749
1779
1809
1839
1869
1899
1930
1960
1990
2020
t [year]
815
576
450
400
311
280
220
132
60
17


10^4
10^3
10^2
10^1
10^0
$F_2^k(\tau)$
10^1
10^2
τ [months]
DFA1
DFA2
DFA4
DFA7
1.14
1.20
1.24
1.23
0.95
0.77
0.71
0.71
2.28



profile without 50 Fourier-terms



profile without 100 Fourier-terms