

Ecological diagnosis from biotic data by Hurst exponent and the R/S analysis adaptation to short time series

Máurea N. Flynn¹,

Professor credenciado no curso pós graduação em Tecnologia Ambiental,
FT, Limeira – Unicamp, 13484-332, Limeira/SP.

William R. L. S. Pereira²

Bio.Sensu Consultoria Ecológica, 04419-030, São Paulo/SP.

Abstract. In recent years, there has been a concerted effort to develop methods for estimating the scaling exponents of time series data, thus permitting a characterization of their underlying dynamical behavior. This task becomes rather inaccurate with data of limited length (less than 100 points), as is the case in many ecological studies where observation time is constrained. In this paper, an arithmetical growth in time window $\tau = 2n$ is considered for ecological time series. This slight alteration allows the adaptation of *R/S* analysis to any ecological series, opening Hurst-Mandelbrot ideas to ecological diagnosis.

Key words: *R/S analysis, Hurst exponent, time series length, time windows length, biotic data.*

1. Introduction

Abiotic (environmental) and biotic (life) parameters are constantly interacting and influencing each other. This simple premise permeates the development of Ecology and it has been the aim of innumerable ecological studies to describe and explain how these interactions take place at individual, population, community and ecosystem levels. Statistical tests, however, frequently

¹maureaflynn@gmail.com

²william.roberto.luiz@hotmail.com

show no correlation whatsoever among abiotic and biotic parameters. The explanation for the apparent lack of relation can be derived from the complexity of mechanisms that defines biological interactions (see Flynn and Pereira, 2011). The interactive parts inside an ecological system is numerous (high degree of freedom) making it difficult to understand its dynamics. In the mathematical field of dynamic and nonlinear systems such condition of multiple interactions is present and relations between the parts concur throughout time and if the result of a variable upon another is not directly proportional to the action strength, the system is said to show nonlinear dynamics.

Hurst (1951) created the Rescaled Range Analysis (R/S analysis) studying time series of floods and droughts determined by rain and volume of the Nile tributaries. The R/S analysis have been applied to ecological time series: Arino and Pimm (1995) to 115 different species (population counted annually per a minimum of 25 years); Keitt and Stanley (1998), to American bird species data from a period of 31 years, covering 3000 observation routes; Halley and Stergiou (2005) to 344 fishes landings, comparing the results with 431 terrestrial populations abundances. Those authors found a distinct exponent value (Hurst exponent, H), revealing a variety of dynamics for ecological time series. A serious problem for R/S analysis application and consequent attainment of H value is the generally short length of ecological series data (Pilgram and Kaplan, 1999). Different methodologies, such as H quantification based on V/S analysis, have been proposals to solve this difficulty but the validity for short or irregular time series is questioned (Katsev and L'Heureux, 2003).

The aim of this article is to propose an adaption to H calculation in order to adequate its application to ecological time series.

2. Objectives

- To describe the methodology to implement Rescaled Range Analysis Algorithm and obtain Hurst Exponent (H), some developments and applications
- To discuss the dependence of H on time series and time windows length
- To address the problem of extremely short time series (< 20 points) and suggest alternative manner to apply R/S ratio in time windows that grow arithmetically

- To suggest a new interpretation to biotic data (population density) based on ecology theory, complex systems and Hurst Exponent

3. Metodology: R/S analysis and Hurst exponent

The phenomenon of persistence, the occurrence of extended periods with systematic deviations from the long-term mean, was first considered in the context of Nile floodings (Hurst, 1951). Often present in hydrological data sets, it is quantified by calculating the Hurst exponent H , using the rescaled range or R/S statistics (Mandelbrot and Ness, 1968). This is a prototypical example of a long-term nonlinear measure. The application of R/S analysis requires the implementation of an algorithm. Alluding to Hurst (1951) and his hydrological studies in the Nile River, we have:

In a certain year t a reservoir receives a $w(t)$ water amount and must liberate part of it to maintain an acceptable water level. The average water intake in r years is given by:

$$\bar{w}_\tau = \frac{1}{\tau} \sum_{t=1}^{\tau} w(t)$$

The water accumulated in the reservoir as a function of the average water amount liberated and received is:

$$X(t, \tau) = \sum_{m=1}^t (w(t) - \bar{w}_\tau)$$

The difference between the water maximum and minimum amount accumulated, denoted by R , representing the required capacity to keep an average outflow \bar{w}_τ as a function of an annual inflow $w(t)$. For a reservoir that never overflows or is empty, R represents the difference between the maximum and minimum amount of water presents in the reservoir in a given t years time period:

$$R = \left[\max_{1 \leq t \leq \tau} \sum_{t=1}^{\tau} (w(t) - \bar{w}(\tau)) - \min_{1 \leq t \leq \tau} \sum_{t=1}^{\tau} (w(t) - \bar{w}(\tau)) \right]$$

Hurst divided R by the standard error to find a dimensionless test quantity:

$$s_\tau = \sqrt{\frac{1}{\tau} \sum_{\tau} (w(t) - \bar{w}(\tau))^2}$$

Empirically R/S is proportional to the observed time τ raised to a characteristic exponent, called later, by Benoit Mandelbrot, Hurst exponent (H). H is obtained by plotting log-log of $R/S(\tau)$ against the time window τ for the scale relation:

$$\left(\frac{R}{S}\right)_\tau \propto \tau^H$$

In the same year Feller (1951) demonstrated that a composed series of variables identically distributed and with finite variance presents R/S value that equals:

$$\left(\frac{R}{S}\right)_\tau = \left(\frac{\tau}{2}\right)^{0.5}$$

If Hurst scaling is found, the expected H values are $0 \leq H \leq 1$; $H < 0.5$ means anti-persistent behavior, which is rarely observed in experimental data, $H = 0.5$ is ordinary Brownian motion and $H > 0.5$ is fractional Brownian motion with increasing persistence strength as H approaches 1. Persistence as quantified by H is also strongly related to scaling in the autocorrelation function and to the occurrence of extreme events.

Analyzing the Nile data, Hurst noticed that flood years and drought years followed no apparent pattern. The series, apparently not organized, showed not random periods and H value could be related to a memory effect meaning that occurrence of drought years and flood years depends on what had happened the moment before. This phenomenon, called the 'Joseph Effect' in allusion to the biblical passage where Joseph warned followers of an oncoming seven years of abundance followed by seven years of scarcity, refers to the memory defined by Hurst. Nile floods deposit a large amount of sediment in crop areas along the river margins, renewing soil fertility each year. Hurst found $H \approx 0.7$ for the Nile River flooding system.

Therefore, R/S analysis scans events and denotes the existence of memory through H exponent value. H provides an exact measure of randomness in any particular time series considered. $H = 0$ features an anti-persistent time series, in other words, an event that occurs subsequently after will have a 100% tendency of an opposite behavior; $H = 1$ indicates that the series is persistent,

meaning that an event that occurs subsequently after will have a 100% tendency of maintaining the same tendency; when $H = \frac{1}{2}$ the event is random, as the Brownian motion, meaning it is not possible to predict the behavior of the next event based on what has occurred in the past (Figure 1).

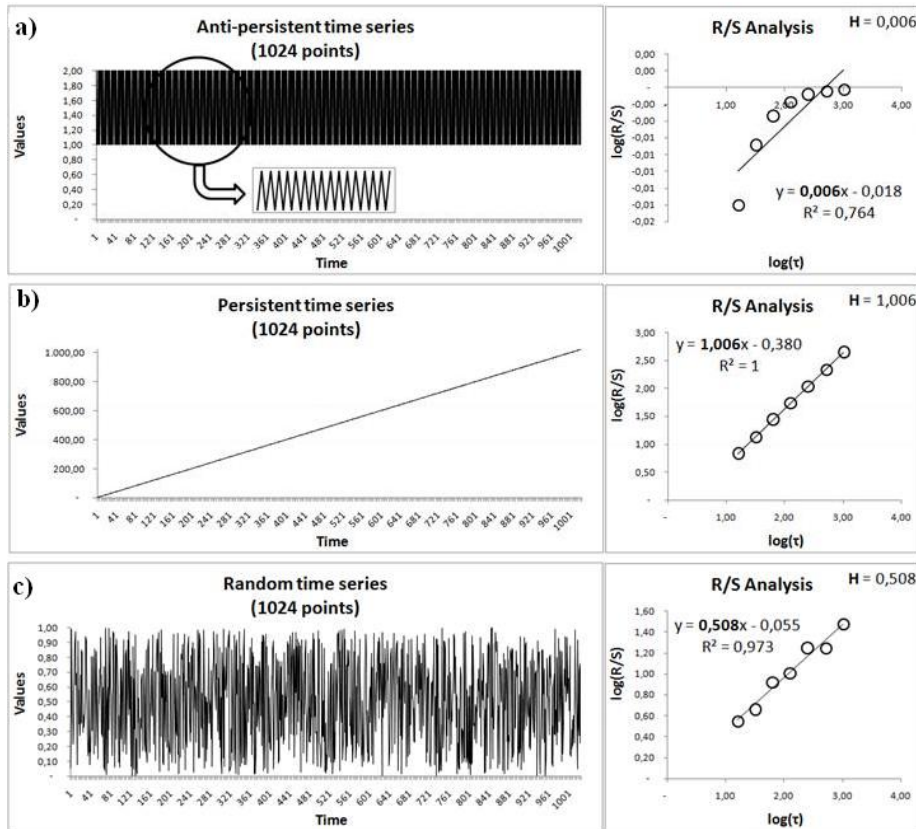


Figure 1: Time series with 1024 data points and respective values for H . a) Anti-persistent time series, $H = 0.006$. b) Persistent time series, $H = 1.006$. c) Random time series $H = 0.508$. In log-log plot, the first three points were out because we don't have an expressive time windows length to calculate R/S . Note in a) the time series consists in a true up-and-down behavior.

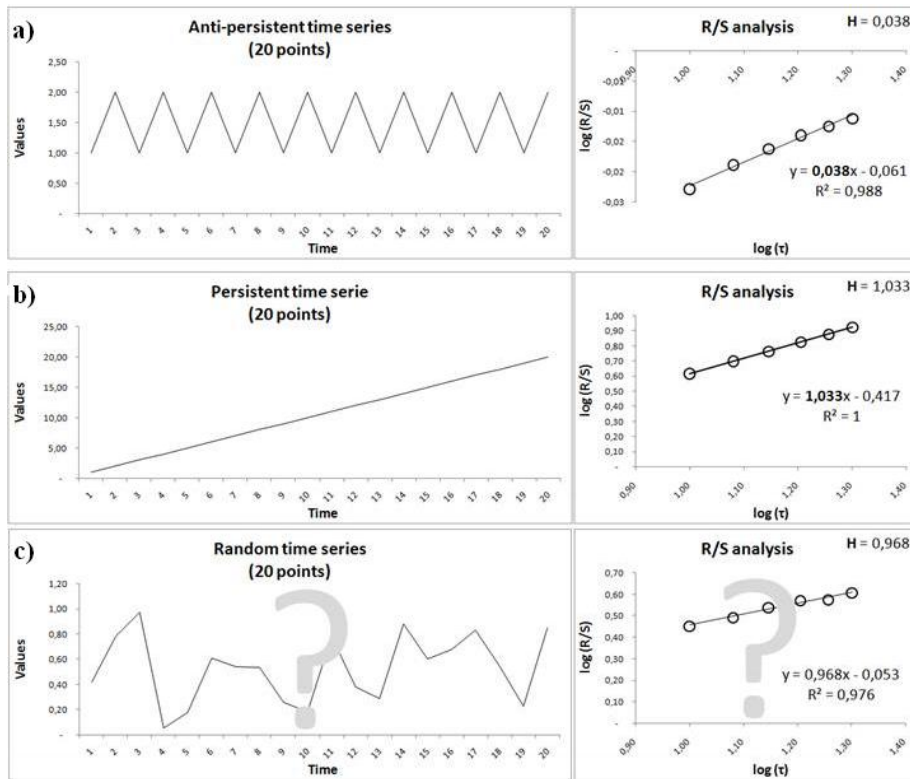


Figure 2: Time series with 20 data points and respective values for H . a) Anti-persistent time series, $H = 0.038$. b) Persistent time series, $H = 1.033$. c) Random time series $H = 0.968$. In log-log plot, the first four points were out because we don't have an expressive time windows length to calculate R/S . Note that there is an inconsistency for H in a random time series, because $H = 0.968$, different from the expected value $H = 0.50$. This will be discussed below.

Any given time series may exhibit a variety of autocorrelation structures. For example, successive terms may show strong, moderate or no positive correlation with previous terms. The strength of these correlations provides useful information about the inherent 'memory' of the system.

4. Results and Discussion

Hurst exponent is generally applied to series containing more than 1000 points and the uncertainty of the H measurement can decrease from 0.1 to 0.05 when good quality data are generated. However, in some field conceived data are relatively small and time series may present certain 'anomalies' (such as leaps, pulses or peaks), which may produce incorrect results (Katsev and L'Heureux, 2003). For the standard R/S methodology, the rescaling is done in τ time windows with geometric growth, with an $\tau = 2n$ format where $n = 1, 2, \dots, 10$ generating $\tau = 2, 4, 8, 16, 32, 64, 128, 256, 512, 1024$ (figures 1, 3a). The series must necessarily contain 1024 points for R/S analysis implementation. The log-log plot that reveals H value shows at least 10 points. Here it is propose an arithmetical growth for ecological time window, $\tau = 2n$ inform, where $n = 1, 2, \dots, 512$ generating $\tau = 2, 4, 6, 8, \dots, 512$ (figures 2, 3b). This slight alteration allows R/S analysis adaptation to any ecological series, even the shortest ones, since the 10 points generated in the log-log plot is attainable for 20 points size series.

R/S analysis with a $\tau = 2n$ time window can be easily applied for financial time series where data is continuously generated. Cajueiro and Tabak (2007) demonstrated that stock market indices of developed countries presents $H \approx 0.5$, while in emerging countries H value is slightly higher (see Lim, 2008). Investors seek certain patterns in stock prices for potential gains. The gain can be quantified through the return rate $\left(\frac{N_{t+1}}{N_t} - 1\right) * 100\%$.

If R/S analysis is applied upon the daily quotation of a specific asset (for example, NATU3) and expected return rate, it is possible to obtain Hurst exponent (Figure 4). The asset assumes strong persistence, while the return is very close to Brownian (random). Such results justify the fact that investments in stock market are profitable in medium and long-term. Immediate returns are hard to obtain because the asset assumes near Brownian movement.

Ecological time series are not monitored as intensely as economics variables. In Brazil long-term ecological projects such as LTER (Long-Term Ecological Research) and LTREB (Long-Term Research in Environmental Biology) are being implemented. Fact is, few long-term time series are available and this condition restrains the interpretation of R/S analysis for short-term memory with $\tau = 2n$ time windows.

R/S analysis was here adapted for *Clethrionomys rufocans* data ob-

tained by Saitoh et al. (1997) considering 10 populations from 10 distinct areas sampled annually in autumn, from 1963 to 1992, which generated exactly 30 points. H was obtained eliminating in an arbitrary manner the first four windows in the log-log plot. A great variability in H was detected (Figure 5).

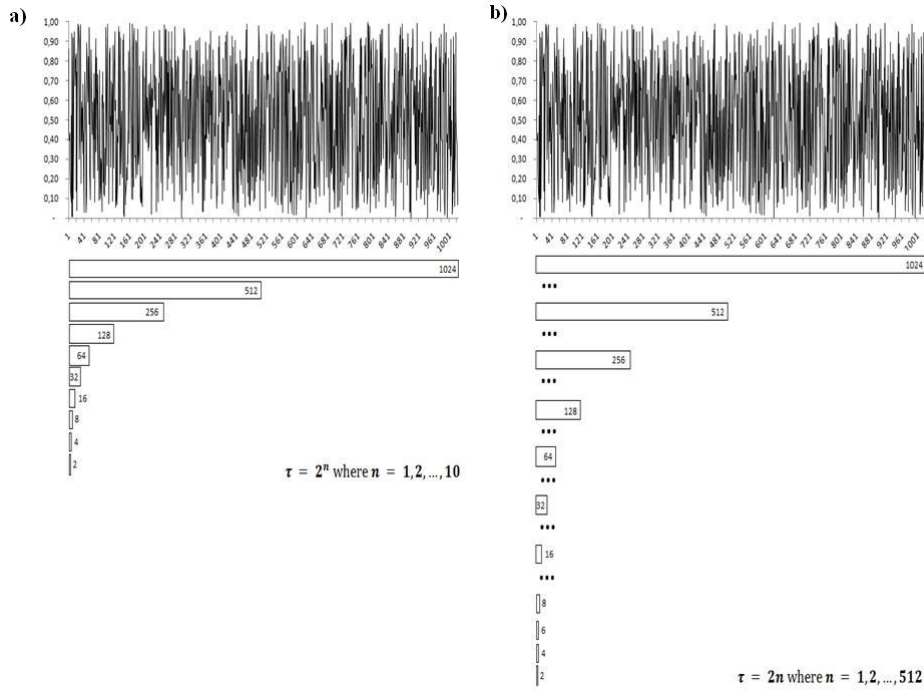


Figure 3: Simulated time series from normally distributed random values and their respective time window length $T = 1024$. a) Time window with geometrical growth, $\tau = 2^n$ where $n = 1, 2, 3, \dots, 10$ generating $\tau = 2, 4, \dots, 1024$. b) Time window with arithmetical growth, $\tau = 2n$ where $n = 1, 2, \dots, 512$ generating $\tau = 2, 4, 6 \dots, 1024$. Note that with $\tau = 2n$, R/S analysis is run with small number of windows (10), while in $\tau = 2^n$ with a much larger number of windows (512).

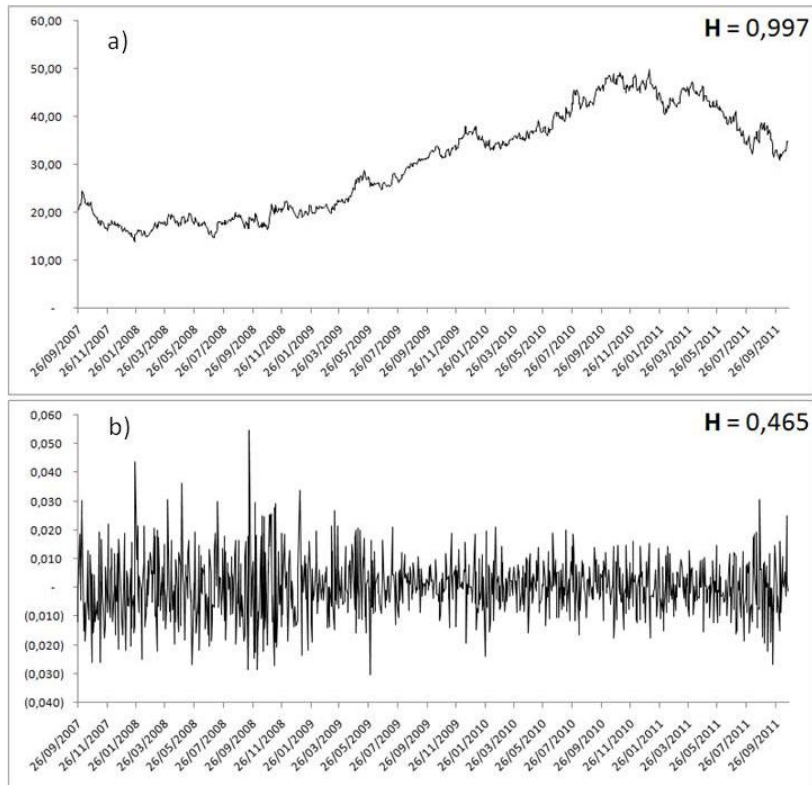


Figure 4: a) NATU3 daily quotation and b) daily rate of return, with respective H considering 1024 points time series. Raw data extracted from www.natura.com.br.

The species considered and, consequently, the population intra-specific regulation mechanisms are the same. However, each population can be influenced by different abiotic variables and/or inter-specific mechanisms unknown initially, being possible to relate H variability to the different ways in which system parts may be concurring in each populations. If $H = \frac{1}{2}$, events are independent and the series presents Brownian movement, denoting that the population dynamics hardly could be modelled. If $H = 1$ or near, the population dynamics could be driven by some biotic or abiotic variables leading possibly to short-term prediction.

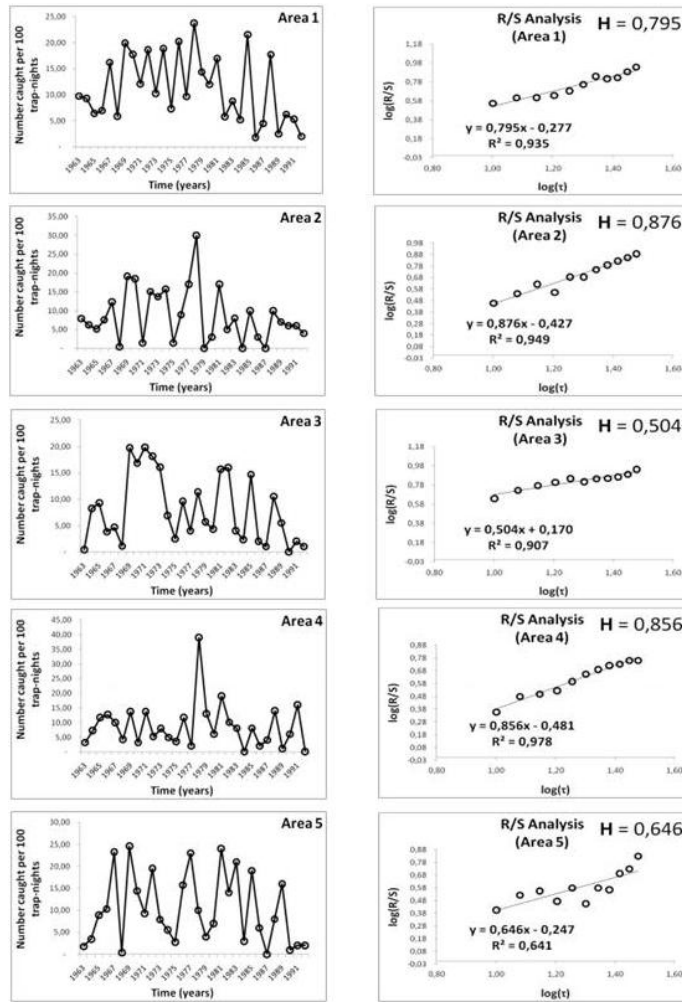


Figure 5: *C. rufocans* population density time series for 10 different locations (Saitoh et al., 1997) over a period of 30 consecutive years and Hurst exponents. Each series did not reveal any apparent pattern, although H values varied from $0.5 \leq H \leq 0.1$.

Two other interpretations are possible. In a rich ecosystem, as interacting parts are very diverse and events are independent due to high degree of freedom, $H = \frac{1}{2}$ is expected. Otherwise, a weak or nonexistent interactions between variables and target variables dynamics related to few determinants,

lead to the assumption that the considered ecosystems is simpler, and $H \neq \frac{1}{2}$ is expected. Considering *C. rufocans* data (Saitoh et al., 1997), the persistency could be related to an ecological law, for examples, Calder's Law, (Calder, 1983) which states that large body size mammal species have longer population cycles. Since *C. rufocans* have small body size, population cycles would be shorter and more intense within a time interval (high frequency cycles) and revealing $H \approx 1$.

When adjusting R/S analysis to ecological series some difficulties are present. If the series is too short, memories are ephemeral, and events that occurred in the near past may have no correlation with events that may occur far away. However, analysis will present the characteristic H value for the series length, which can reveal the existence of short-term memory. For short life cycle organisms (bacteria or other small organisms), such as the rodent *C. rufocans* mentioned above, the seemingly chaotic population fluctuation ups and downs can reveal a certain degree of correlation, even in trend apparently undetected. H values could also be associated to fleeting memories with transitory states of a system. In dynamic systems the interacting parts compete with each other and the different values of H can reveal different states of the same system.

The implementation of R/S analysis for short time window could be dislocated through time in relatively long time series so that H would act as continuous variable, and its value would carry the system different transient states, characterizing its dynamics.

It can be concluded that R/S analysis can be adjusted to any size time series, without losing its fractals characteristics (figures 1, 2). It's known a relation between fractal dimension and Hurst exponent by the equation (Mandelbrot, 1982):

$$D = 2 - H$$

If $H = 0$, $D = 2$. In geometric sense, $D = 2$ means that an object 'fills' all the two-dimensional space (a given area) (figures 1a, 2a); if $H = 1$, $D = 1$, or in other words, the object 'fills' all the one-dimensional space (a line) (figures 1b, 2b); other H values shown fractated values to fractal dimension, or in other words, the object 'slightly fills' the two-dimensional space (H higher than $H = 1$) or 'heavily fills' the two-dimensional space (H smaller than $H = 2$) (figure 1c).

But we have statistical problems, because of the length time series and

consequently the time window (figure 2c). By sorting true random values we could expect $H = 0.5$ to any time series length, but computational tests show that H converges to $H = 0.5$ when the time windows grows (Mesa and Poveda, 1993; Sánchez Granero et al., 2008). In mathematical notation:

$$E\left[\frac{R(\tau)}{S(\tau)}\right] = C\tau^{0.5} \text{ when } \tau \rightarrow \infty$$

This can be explained by the nature of random number. In a sequence of true random numbers, sometimes we could see an apparently order (Mlodinow, 2009) inside a long time series by drawing random numbers the and R/S methodology certify that $H = 0.50$ only for longer and longer time windows. So, by considering that time window grows arithmetically in our proposition, the equality above should be taken account.

And due to short-term memory (ephemeral memory) others implications have to be taken into account when interpreting results. R/S analysis may reveal states of a system that changes over time (e.g. in financial assets) or allows the characterization of a variable inserted in different systems (e.g. rodents? densities), suffering different degrees of interactions. Its concepts from statistical-physics were highly successful in revealing important aspects of economics and other sciences, and can, therefore, encourage the development of an index that synthesizes the operation of a same variable within a system and provides clues of the system towards this same variable.

5. Conclusion

Although Hurst exponent is an indicator highly applied in economy, it has been almost forgotten by ecology, mostly because of its dependence on long time series, which for biotic ecological data are, in most cases, short. Trough the assessment of particularities of the Rescaled Range Analysis algorithm, some reasons arises. Hurst exponent can be an efficient tool in revealing some aspects of population density data. It can measure the role played by certain variable in an ecosystem, since highly complex systems in which a great source of influences (biotic and abiotic) are expected, leads us to a rich ecosystem ($H = \frac{1}{2}$), while if some trend in time series ($H \neq \frac{1}{2}$) appears, a reduce number of interactions is perceived leading to a simpler and predictable system (until $H = 1$ or $H = 0$). It is suggest ways to test short time windows that grows arithmetically to $H = \frac{1}{2}$ from true random numbers.

Agradecimentos

To Adriana Cristiane Ruy for the careful adaptation of the manuscript into Latex language.

References

- Arino, A. and Pimm, S. L. (1995). On the nature of population extremes. *Evolutionary Ecology*, 9:429–443.
- Cajueiro, D. O. and Tabak, B. M. (2007). Testing for long-range dependence in world stock markets. *Chaos, Solitons & Fractals*, 37(3):918–927.
- Calder, W. A. (1983). An allometric approach to population cycles of mammals. *Journal of Theoretical Biology*, 100:275–282.
- Feller, W. (1951). The asymptotic distribution of the range of sums of independent random variables. *Annals of Mathematical Statistics*, 22:427.
- Flynn, M. N. and Pereira, W. R. L. S. (2011). Abordagem populacional na ecotoxicologia. *RevInter Revista Intertox de Toxicologia, Risco Ambiental e Sociedade*, 4(3):79–91.
- Halley, J. M. and Stergiou, K. I. (2005). The implications of increasing variability of fish landings. *Fish and fisheries*, 6:266–276.
- Hurst, H. (1951). Long term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 116:770–799.
- Katsev, S. and L’Heureux, I. (2003). Are hurst exponent estimated from short or irregular time series meaningful? *Computers & Geosciences*, 29:1085–1089.
- Keitt, T. H. and Stanley, H. E. (1998). Dynamics of north american breeding bird population. *Nature*, 393:257–260.
- Lim, K. P. (2008). Long-term dependence in world stock markets: Does legal origin matter? *ICFAI Journal of Applied Finance*, 14(5):57–68.
- Mandelbrot, B. B. (1982). *The fractal geometry of Nature*. W. H. Freeman and Company, San Francisco.

- Mandelbrot, B. B. and Ness, J. W. V. (1968). Noah, joseph and operational hydrology. *Water Resources Research*, 4:909–918.
- Mesa, O. J. and Poveda, G. (1993). The hurst effect: the scale of fluctuation approach. *Water Resources Research*, 9(12):3995–4002.
- Mlodinow, L. (2009). *O andar do bêbado: como o acaso determina nossas vidas*. Jorge Zahar, Rio de Janeiro.
- Pilgram, B. and Kaplan, D. T. (1999). A comparison of estimators for 1/f noise. *Physica D*, 114:108–122.
- Saitoh, T., Stenseth, N. C., and Bjornstad, O. N. (1997). Density dependence in fluctuating grey-sided vole populations. *Journal of Animal Ecology*, 66:14–24.
- Sánchez Granero, M. A., Trinidad Segovia, J. E., and García Pérez, J. (2008). Some comments on hurst exponent and the long memory processes on capital markets. *Physica A*, 387:5543–5551.