Delignières D, & Marmelat V. (2013). *Advances in Experimental Medicine and Biology, 782*, 127-148.

Theoretical and methodological issues in serial correlation analysis

Didier Delignières & Vivien Marmelat
EA 2991 Movement to Health – University Montpellier 1, France

Abstract

We present in this paper some theoretical and methodological problems related to the analysis of serial correlations in experimental data. A very common observation in behavioral and physiological experiments is the presence of long-range correlations in time series. In this case the current observation seems to keep the memory of a large set of previous observations. This kind of process has been referred to as long-range dependence, long-term memory, fractal correlation, or 1/*f* noise. There is now a general agreement for considering long-range correlations as reflecting the complexity of the system, defined as the flexible and adaptable coordination between its multiple components and sub-systems. Long-range correlations are supposed to sign an optimal compromise, between order and disorder, order reflecting a too strict and rigid coordination, and disorder the absence of coordination. Long-range correlations are considered the signature of health and adaptability, and deviations towards order and disorder have been described in elderly or pathological populations. As such, the detection of long-range correlations and the assessment of their alteration in specific populations or situations appear as an important scientific goal.

However, the detection of long-range correlation in empirical series is not straightforward. A number of recent experiments have showed that the pattern of correlation observed in empirical series combines short-range and long-range correlation processes. This is important because apparent alterations in long-range correlations could in fact be due to the superimposition of short-term correlation processes.

An additional problem is that short-range correlation processes can sometimes mimic long-range correlations. Classical methods, such as Detrended Fluctuation Analysis or Power Spectrum Density analysis, seem unable to distinguish between short- and long-range correlated processes. Some specific methods have been developed for testing for the effective presence of long-range correlations in experimental series.

Key words: Long-range correlations, 1/*f* fluctuations, time series, Detrended Fluctuation Analysis, Power Spectral Density

Time series and auto-correlation function

Time series analysis is a relatively recent approach in experimental psychology and movement sciences. For a long time repeated measurements of performance were performed for obtaining more accurate estimates of average performance, or for assessing performance variability. From this point of view, successive performances were considered independent, and their order of occurrence was not considered. In other words, successive performances were considered white noise fluctuations around a mean stable value. However, Slifkin and Newell (1998) argued that this "white noise assumption" could be more the exception rather than the rule, and a number of experiments have evidenced the presence of correlations between successive measurements of performance. The aim of time series analyses is to reveal and characterize serial dependence in series of observations[1]..

The concept of serial dependence refers to systematic relationships between observations over time. The autocorrelation function represents a straightforward way for analyzing serial dependence. Autocorrelation is defined as the correlation of a time series with itself, considering a given lag. The auto-correlation function represents the series of auto-correlation values obtained for increasing lags. By definition, the auto-correlation function of a white noise process presents

[1] Note that we just define a time series as a series of data ordered in time. A strict definition supposes that successive data are spaced by regular time intervals, but in most experiments series are simply composed of ordered data.

non-significant, close-to-zero values, revealing the absence of dependence in the series (see Figure 1, left column).

Serial dependence can occur on the short term: In that case, the current value is dependent on the previous value, or on a few set of preceding values. A good example is provided by auto-regressive series, obeying the following equation:

$$y_i = \varphi y_{i-1} + \varepsilon_i \qquad (1)$$

where φ is the autoregressive parameter and ε_i is a white noise process with zero mean and unit variance. In this kind of process, the current value (y_i) is defined as a fraction of the preceding value (y_{i-1}), plus a random perturbation (ε_i). The strength of the relation between successive values is entirely defined by the value of the autoregressive parameter. In this very simple process, the current value is directly dependent on the immediately preceding value. Serial correlations can appear between y_i and more distant observations (y_{i-2}, y_{i-3}, etc.), but these relations are not direct and are explained by the presence of intermediate observations (e.g., the observed relationship between y_i and y_{i-2} is explained by the effective relationships between y_i and y_{i-1}, and between y_{i-1i} and y_{i-2}).

The autocorrelation function of this kind of process typically presents a significant value at lag one (depending on the value of φ), and then a quick decay over time (see Figure 1, central column).

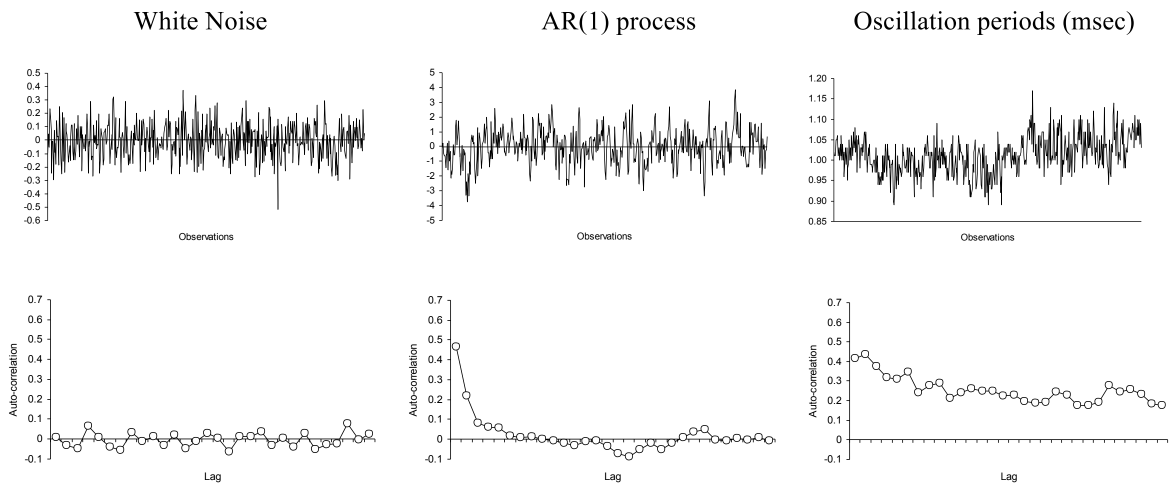

Figure 1: Time series (top raw) and auto-correlation functions (bottom raw). Left column: white (uncorrelated) noise; middle column: one-order auto-regressive process; right column: experimental series of periods produced during repetitive oscillations of the forearm (data from Delignières, Lemoine & Torre, 2004).

Long-range correlations

This exponential decay of the autocorrelation function, however, is rarely observed in empirical series. Often the autocorrelation function presents a very slow decay over time, suggesting that the current value presents dependencies with a large set of preceding values. For example, we present in Figure 1 (right column) a series of periods produced during the repetitive oscillations of the forearm (Delignières, Lemoine & Torre, 2004). The auto-correlation function presents a very slow decay over time, and remains significant over a number of lags. This persistence of auto-correlation signs the presence of long-range correlations (LRC) in the series.

LRC are characterized by the presence of dependence which tends to persist over dozens or even hundreds of data. In this kind of process the current observation seems to keep the memory of a large set of previous observations. LRC can be understood through the fact that over multiple, interpenetrated time scales, an increasing trend in the past is likely to be followed by an increasing trend in the future, and conversely a decrease in the past is likely to be followed by a decrease in the future. This kind of process has been referred to as long-range dependence, long-term memory, fractal process, or 1/*f* noise.

LRC have been discovered in the dynamics of a number of natural and physical systems, including for example the series of discharges of the Nile River (Hurst, 1951), the series of magnitudes of earthquakes (Matsuzaki, 1994), the evolution of traffic in Ethernet networks (Leland, Taqqu, Willinger, & Wilson, 1994), or the dynamics of self-esteem over time (Delignières, Fortes, & Ninot, 2004). In the domain of human movement, LRC have been evidenced in serial reaction time (Gilden, 1997; van Orden, Holden, & Turvey, 2003), in finger tapping (Gilden, Thornton, & Mallon, 1995; Lemoine, Delignières & Torre, 2006), in stride duration during walking (Hausdorff, Peng, Ladin, Wei, & Goldberger, 1995), or in relative phase in a bimanual coordination task (Torre, Delignières & Lemoine, 2007a).

Fractional Gaussian noise and fractional Brownian motion

Most often, LRC has been modeled through a family of processes introduced by Mandelbrot and van Ness (1968), namely *fractional Gaussian noises* and *fractional Brownian motions*. A formal introduction to this framework is necessary for a complete understanding of this approach. The authors first defined a family of processes they called fractional Brownian motions (fBm). The starting point of this definition is ordinary Brownian motion, a well-known stochastic process that could represent, for example, the movement of a single particle along a straight line, due to a succession of random jumps. Then, an important property of Brownian motion is that its successive increments in position are uncorrelated: each displacement is independent of the former, in direction as well as in amplitude. Einstein (1905) showed that, on average, this kind of motion moves a particle from its origin by a distance that is proportional to the square root of the time.

fBm series differ from ordinary Brownian motion by the fact that in a fBm successive increments are correlated. A positive correlation signifies that an increasing trend in the past is likely to be followed by an increasing trend in the future; in that case the series is said to be persistent. Conversely, a negative correlation signifies that an increasing trend in the past is likely to be followed by a decreasing trend, and the series is said to be anti-persistent.

Mathematically, fBm is characterized by the following scaling law:

$$SD(x) \propto \Delta t^{H} \qquad (2)$$

which signifies that the standard deviation of the process is a power function of the time interval (Δt) over which it was computed. H is the so-called Hurst exponent and can be any real number in the range $0 < H < 1$. Eq. 2 describes the so-called *diffusion property* of fBm series: the higher H, the higher the diffusion of the process over time. For this point of view, ordinary Brownian motion corresponds to the special case $H = 0.5$ and constitutes the frontier between anti-persistent ($H < 0.5$) and persistent fBm series ($H > 0.5$). We present in Figure 2 three example fBm series, for three contrasted H exponents.

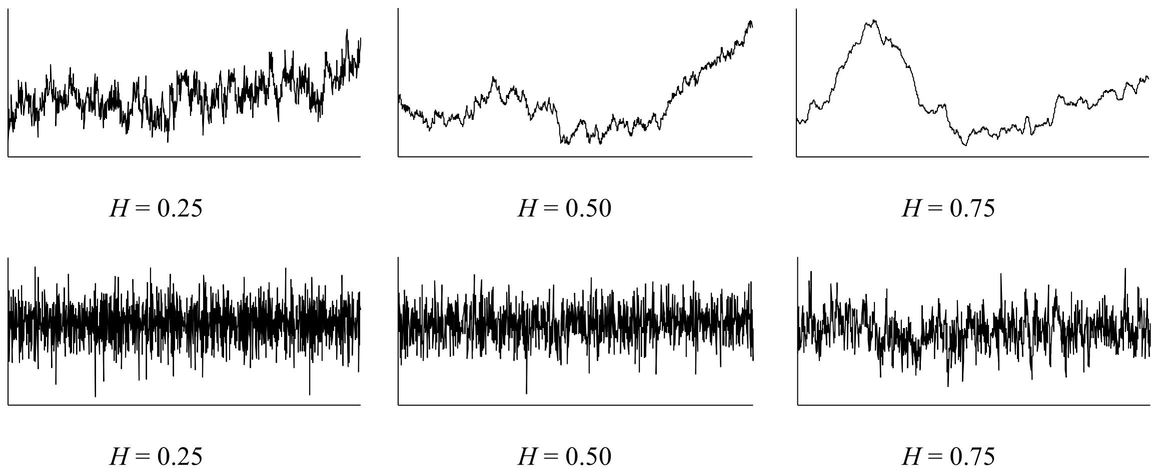

Figure 2: Top raw: example series of fractional Brownian motions (fBm) for three typical values of the scaling exponent. The central graph represents an ordinary Brownian motion (H=0.5). The left graph shows an anti-persistent fBm (H=0.25) and the right graph a persistent fBm (H=0.75). The corresponding fractional Gaussian noises series (fGn) are displayed in the bottom raw.

Fractional Gaussian noise (fGn) represents another family, defined as the series of successive increments in a fBm. In other words a fGn is the differentiation of a fBm, and conversely the integration of a fGn gives a fBm. Each fBm series is then related to a specific fGn series, and both are characterized by the same H exponent. We present in the bottom row of Figure 2 the fGn series corresponding to the just above fBm series. The fGn family is centered around white noise ($H = 0.5$), which represents the frontier between anti-persistent ($H < 0.5$) and persistent fGn ($H > 0.5$).

These two families of processes possess fundamentally different properties: fBm series are non-stationary with time-dependent variance (diffusion property), while fGn are stationary with a constant expected mean value and constant variance over time. As previously explained, fGn and fBm can be conceived as two superimposed families, invertible in terms of differentiation and integration.

Another useful conception is to conceive these two families as representing a continuum, ranging from the most anti-persistent fGn to the most persistent fBm. This fGn/fBm continuum is characterized by the presence of scaling laws that could be expressed in the frequency or in the time domain. In the frequency domain, a scaling law relates power (i.e. squared amplitude) to frequency according to an inverse power function, with an exponent β:

$$S(f) \propto 1/f^{\beta} \qquad (3)$$

This scaling law is exploited by the Power Spectral Density (PSD) method that reveals β as the negative of the slope of the log-log representation of the power spectrum (Figure 3). The fGn/fBm continuum is then characterized by exponents β ranging from -1 to 3 (see Figure 5).

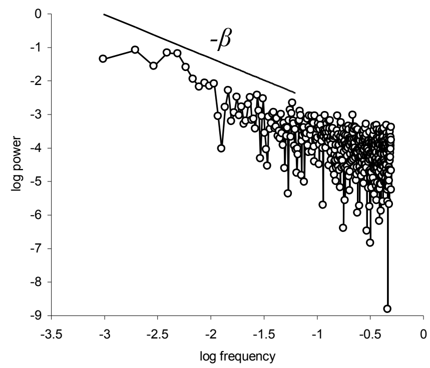

Figure 3: Power Spectral Density analysis. The exponent β is given by the negative of the slope of the log-log representation of the power spectrum.

In the time domain, the typical scaling law states that the standard deviation of the integrated series is a power function of the time over which it is computed, with an exponent α. Considering a time series $x(i)$:

$$\begin{cases} y(i) = \sum_{k=0}^{i} x(k) \\ SD(y) \propto n^{\alpha} \end{cases} \qquad (4)$$

This scaling law is exploited by the Detrended Fluctuation Analysis (DFA) that reveals α as the slope of the log-log diffusion plot (Figure 4). The fGn/fBm continuum is characterized by exponents α ranging from 0 to 2 (see Figure 5). Note that the scaling law expressed in Eq. 4 just derives from the original definition of fBm (Eq. 2). If the series $x(i)$ is a fGn, y(i) is the corresponding fBm and α is the Hurst exponent. If $x(i)$ is a fBm, $y(i)$ belongs to another family of over-diffusive processes, characterized by exponents α ranging from 1 to 3, and in that case $\alpha = H + 1$.

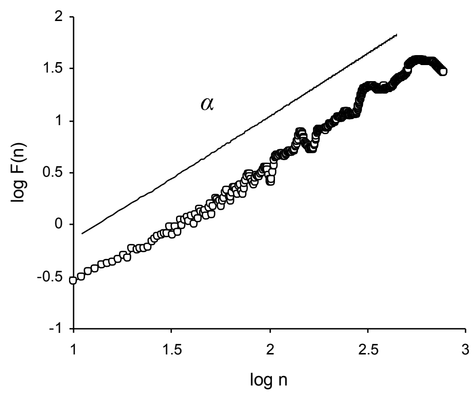

Figure 4: Detrended Fluctuation Analysis. The exponent α is determined as the slope of the log-log diffusion plot.

Note that the different exponents characterizing these scaling laws are mutually linked by very simple equations:

For fGn series:

$$\beta = 2H - 1 \quad \text{and} \quad \alpha = H \qquad (5)$$

For fBm series:

$$\beta = 2H + 1 \quad \text{and} \quad \alpha = H + 1 \qquad (6)$$

For fGn and fBm series:

$$\beta = 2\,\alpha + 1 \quad \text{or} \quad \alpha = (\beta - 1)/2 \qquad (7)$$

The exponents provided by PSD and DFA (β and α, respectively), are useful because they allow to unambiguously distinguish between fGn and fBm series, which could be characterized by the same H exponents. As an example, the two series presented in the right column of Figure 2 are characterized by the same H exponent (H = 0.25). However, for the fGn series (bottom raw), β = -0.5 and α = 0.25, and for the fBm series (top raw), β = 1.5 and α = 1.25.

In this fGn/fBm continuum, LRC are generally considered to appear in a narrow range, between β = 0.5 and β = 1.5 (i.e. between α = 0.75 and α = 1.25, see Wagenmakers, Farrell and Ratcliff, 2004). This range is centered on β = α = 1, corresponding to the ideal $1/f$ noise. Long-range correlated series present typical fluctuations, often referred to as $1/f$ fluctuations, characterized by multiple interpenetrated waves, a weak stationarity, and a marked level of roughness.

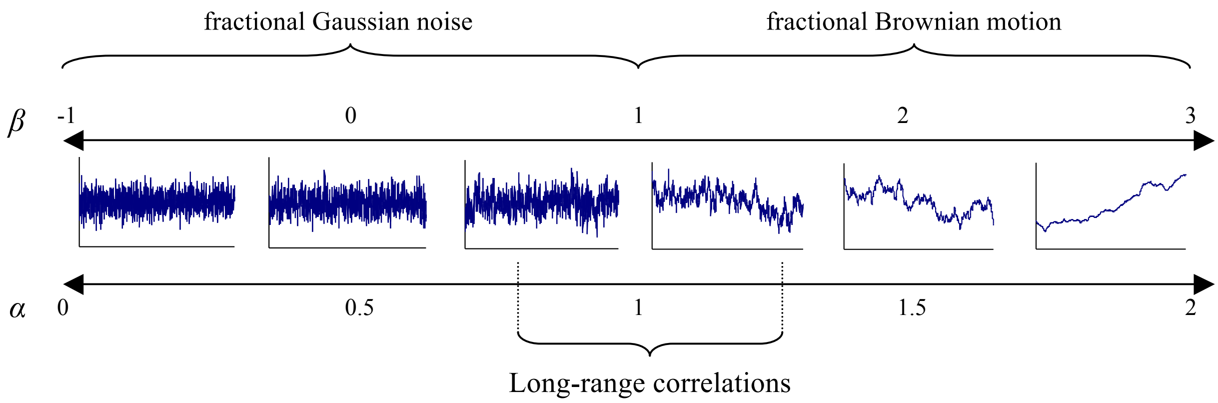

Figure 5: Representation of the fGn/fBm continuum. The continuum is characterized by exponents β ranging from -1 to 3, and by exponents α ranging from 0 to 2. White noise corresponds to β = 0 and α = 0.5, and Brownian motion to β = 2 and α = 1.5. Long range correlations are considered to appear between β = 0.5 and β = 1.5. $\beta = \alpha = 1$ corresponds to ideal 1/*f* noise.

Long-range correlations, complexity and health

As previously explained, LRC have been discovered in many physical and natural systems. Science being generally based on the detection of differences (between groups, or between conditions), one could question the scientific interest of studying such a ubiquitous feature. From this point of view, 1/*f* fluctuations are considered an omnipresent background noise, without discriminative power. As such, LRC could be thought just as an amazing phenomenon or a mathematics fantasy.

Another line of criticism is to consider 1/*f* fluctuations as unexpected perturbations, due to prolonged observation times. For example the presence of LRC in tapping experiments can be supposed to reflect basically the inability of participants to keep a steady tempo, especially in long sequences. From this point of view, 1/*f* fluctuations can be thought as just as undesirable as any other noise. In other words, researchers focusing on LRC just create the problem they whish to study, and 1/*f* fluctuations hide the deterministic or 'true' part of the collected signals. In sum, LRC are sometimes conceived at best a mathematical curiosity, at worst an experimental artifact (Pressing & Jolley-Rogers, 1997).

A number of authors, however, have adopted a more heuristic point of view, considering that 1/*f* fluctuations reveal the structural and functional properties of the system that generated them. Considering the ubiquity of 1/*f* fluctuations in the outcome series produced by natural, physical and biological systems, some authors have suggested that the occurrence of such fluctuation could correspond to some universal principles (Beltz & Kello, 2006; Ihlen & Vereijken, 2010; Kello, Beltz, Holden, & Van Orden, 2007; Kello, Brown, Ferrer-i-Cancho, Holden, Linkenkaer-Hansen, Rhodes, & Van Orden, 2010). During some years a debate opposed these authors to the proponents of an alternative, so-called *mechanistic* approach, seeking for specific explanations and models for the presence of 1/*f* fluctuations in specific systems or situations (Torre & Wagenmakers, 2009), but by now a general agreement seems to emerge favoring the *nomothetic* view considering LRC as related to universal properties (Diniz et al., 2011). Form this point of view, LRC are considered to reflect the complexity of the system, defined as the flexible and adaptable coordination between its multiple components and sub-systems (Diniz et al., 2011; Kello et al., 2007). 1/*f* fluctuations do not arise from some specific component within the system but, rather, from the complex, multiplicative interactions between the multiple components, acting at different time scales, that compose the system. Several characteristic features of complex systems have been advocated to play a central role in the emergence of LRC, such as self-organized criticality (van Orden, Holden & Turvey, 2003), multiscale dynamics (Hausdorff, 2005), metastability (Kello, Beltz, Holden, & van Orden, 2007), or degeneracy (Delignières, Marmelat & Torre, 2011).

As previously explained, LRC seem to appear in series produced by healthy, sustainable, adaptable and flexible systems (Bassingthwaighte, Liebovitch, & West, 1994; Goldberger, Amaral, Hausdorff, Ivanov, Peng, & Stanley, 2002; van Orden, 2007; West, 1990). For example Goldberger et al (2002) have showed that the heart inter-beat intervals series recorded in young and healthy subjects typically presented 1/*f* fluctuations. In contrast, heartbeat series appear more random in patients with a cardiac arrhythmia, and conversely more periodic and predictable in patients with severe congestive heart failure. Similar alterations of LRC have been evidenced by Hausdorff, Mitchell, Firtion, Peng, Cudkowicz, Wei, and Goldberger (1997), who showed that stride interval series during walking presented 1/*f* fluctuations in young and healthy participants, but tended to loose their correlation structure (i.e. became more random) in elderly and in patients suffering from neurodegenerative diseases. Note that this alteration of LRC, especially in elderly, is consistent with the *loss of complexity* hypothesis proposed by Goldberger, Peng, and Lipsitz (2002).

The observation of such alterations suggests that LRC represent an optimal compromise, between order and disorder, order reflecting a too strict and rigid coordination, yielding stereotyped behavior and low adaptability, and disorder an absence of coordination. LRC can be considered a macroscopic biomarker of health, and their detection and the assessment of their alteration in specific populations or situations appear as an important scientific goal (Eke et al., 2000).

Time series analysis methods

We previously evoked two of the most used methods, PSD and DFA. These methods are useful because they can be applied on both fGn and fBm series. PSD has been improved by Eke et al. (2000) with the introduction of some pre-processing operations. In a first step the mean of the series is subtracted from each value, and then a parabolic window is applied: each value in the series is multiplied by the following function:

$$W(j) = 1 - \left(\frac{2j}{N+1} - 1\right)^2 \quad \text{for } j = 1, 2, \ldots, N. \qquad (8)$$

Thirdly a bridge detrending is performed by subtracting from the data the line connecting the first and last point of the series. The Fast Fourier Transform algorithm is applied on the resulting series. Finally the estimation of β, on the basis of Eq. 3, excludes the high-frequency power estimates ($f > 1/8$ of maximal frequency). This method was proven by Eke et al. (2000) to provide more reliable estimates of the spectral exponent (see also Delignières et al., 2006).

DFA was initially proposed by Peng et al. (1993). This method exploits the scaling law expressed in Eq. 4, and includes a detrending procedure in order to cope with local non-stationarities in the series. The analyzed series $x(t)$ is first integrated, by computing for each t the accumulated departure from the mean of the whole series:

$$X(k) = \sum_{i=1}^{k} \left[x(i) - \bar{x} \right] \qquad (9)$$

This integrated series is then divided into non-overlapping intervals of length n. In each interval, a least squares line is fit to the data (representing the trend in the interval). The series $X(t)$ is then locally detrended by substracting the theoretical values $X_n(t)$ given by the regression. For a given interval length n, the characteristic size of fluctuation for this integrated and detrended series is calculated by:

$$F = \sqrt{\frac{1}{N} \sum_{k=1}^{N} \left[X(k) - X_n(k) \right]^2} \qquad (10)$$

This computation is repeated over all possible interval lengths (in practice, the shortest length is around 10, and the largest $N/2$, giving two adjacent intervals). Typically, F increases with interval length n. A power law is expected, as

$$F \propto n^{\alpha} \qquad (11)$$

α is expressed as the slope of a double logarithmic plot of F as a function of n.

Delignières et al. (2006) presented an evaluation of the efficiency of these methods for measuring the strength of LRC in simulated series. As can be seen in Figure 6, both methods gave satisfying results. DFA worked especially well with fGn signals ($0 < \alpha < 1$), but presented a systematic underestimation bias in fBm. PSD produced a slight underestimation bias in fGn series and in low diffusive fBm (especially for $\alpha = 1.1$), and conversely an overestimating bias in high diffusive fBm series ($\alpha > 1.7$). In all cases the estimates presented moderate variability levels.

Some other methods have been proposed, but they often present some limitations in their ranges of application. For example the Rescaled Range analysis (Hurst, 1965; Caccia, Percival, Cannon, Raymond, & Bassingthwaigthe, 1997) and the Dispersional analysis (Caccia et al., 1997) only work on fGn signals, and gave erroneous results when applied to fBm signals. Conversely, Windowed Variance Analysis methods work only on fBm signals (Cannon, Percival, Caccia, Raymond, & Bassingthwaighte, 1997; for a general presentation see Eke et al., 2000, 2002; Delignières et al., 2006).

It is important to note that the accuracy of the estimation of fractal exponents is directly related to the length of the analyzed series. However, Delignières et al. (2006) showed that fractal analysis methods provided acceptable results with series of 1024 or 512 points, except PSD that appeared severely affected by the shortening of series, despite the algorithmic improvements introduced by Eke et al. (2000). The results obtained with series of 256 data points were not so bad, especially with DFA. This observation is very important, because of the difficulty to obtain long time series in psychological and behavioral experiments. These results suggests that a better estimate of scaling exponents could be obtained, with a similar time on the experimental task, from the average of four exponents derived from distinct 256 data points series (with an appropriate period of rest between two successive sessions), than from a single session providing 1024 data points. This could open new perspectives of research in areas that were until now reticent for using this kind of analyses.

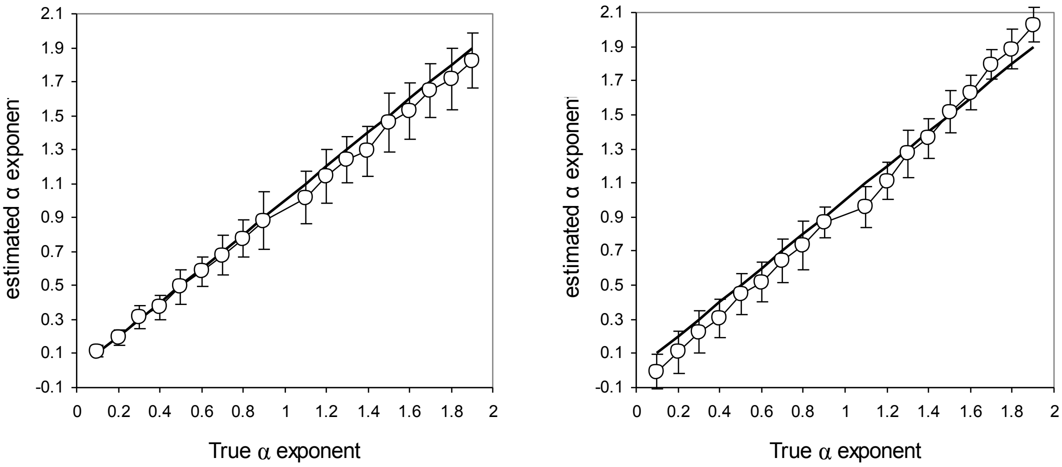

Figure 6: Accuracy and variability of DFA (left), and PSD (rigth) estimates of scaling exponents, based of the analysis of simulated series with known exponents. The graphs represent the relationship between the exact exponents (expressed in αDFA, from α = 0.1 to α = 0.9 for fGn, and from α = 1.1 to α = 1.9 for fBm), and the mean estimates provided by the two methods. Error bars represent standard deviation (data from Delignières et al., 2006).

Short-term and long-range correlations

The detection of long-range correlation is not so straightforward, however, in empirical series. A number of studies have been published in the last decade, which showed that the pattern of correlation in a given signal appeared as a complex combination of short-range and long-range correlation processes.

A very classical example of such combination has been described by Gilden, Thorton and Mallon (1995), in the series of time intervals produced during finger tapping. The application of PSD on these series produces a very specific power spectrum, with a negative slope in the low frequency region of the log-log spectrum, and a positive slope in the high frequency region (Figure 7).

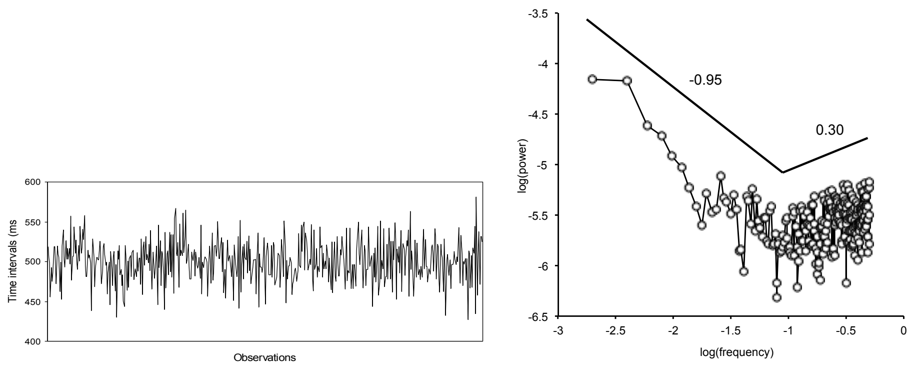

Figure 7: Right: An example time series of inter-tap intervals collected in a self-paced finger tapping task. Left: averaged log-log power spectrum (N = 11, from Delignières, Lemoine & Torre, 2004).

The authors interpreted these results on the basis on the well-known model by Wing and Kristofferson (1973), suggesting that inter-tap intervals are determined by the cognitive intervals produced by an internal clock, plus a term of differenced white noise due to the motor delay that characterizes each tap.

$$I_i = C_i + M_i - M_{i\text{-}1} \qquad (12)$$

In the initial formulation of the model, C_i and M_i were both considered as white noise sources. According to Gilden and collaborators, the positive slope in high frequencies is due to the negative short-term correlations induced by the differenced white noise term ($M_i - M_{i\text{-}1}$). They concluded that LRC in tapping series were induced by the timekeeper (C_i), which was then supposed to possess fractal properties.

The results are more complex when tapping is performed in synchronization with a metronome. In that case the application of PSD on time interval series yields a positive slope over the entire spectrum, suggesting the presence of negative correlations in the series (Figure 8).

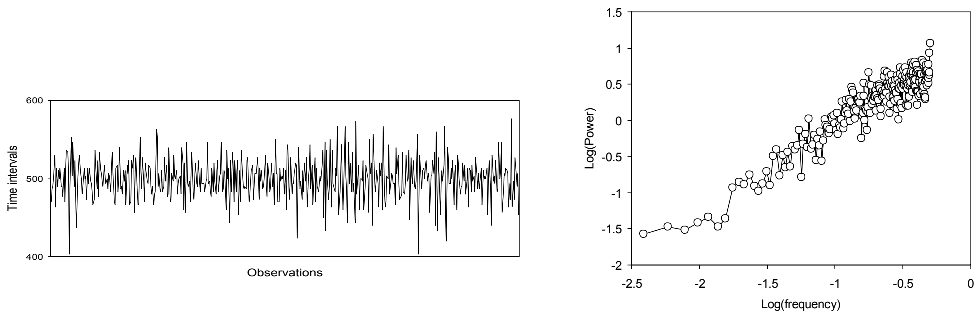

Figure 8: Right: An example time series of inter-tap intervals collected in a synchronized finger tapping task. Left: averaged log-log power spectrum ($N = 13$, data from Torre & Delignières, 2008).

Torre and Delignières (2008) interpreted this result on the basis of a model initially proposed by Vorberg and Wing (1996).

$$A_{i+1} = (1 - \alpha)\, A_i + C_i + M_{i+1} - M_i - \tau \qquad (13)$$

In this model τ represents the fixed period of the metronome, and C_i and M_i are the components of the aforementioned Wing-Kristofferson model (internal timekeeper and motor delays, respectively). A_i represents asynchronies to the metronome and the model incorporates an autoregressive correction process with parameter (1 - α). Time intervals are derived from a linear combination of successive asynchronies (Eq. 14).

$$I_i = A_i - A_{i-1} + \tau \qquad (14)$$

In the initial formulation of the model the timekeeper (C_i) was considered a white noise source (Vorberg & Wing, 1996), but this hypothesis was unable to explain the pattern of correlations observed in interval series. Torre and Delignières (2008) showed that providing the timekeeper with $1/f$ properties allowed to satisfactorily reproducing the experimental results.

These two examples show that even in very simple laboratory tasks, LRC do not appear in isolation in experimental series, and could sometimes be difficultly discernable. This is important because apparent alterations in long-range correlations could in fact be due to the superimposition of short-term correlations.

An interesting example has been described in the domain of walking. Hausdorff et al. (1996) showed that the series of stride intervals during walking contained LRC. However, they suggested that correlations tended to disappear when participants had to walk in synchrony with a metronome. The authors applied DFA of stride interval series and obtained α exponents close to one in self-paced walking, but close to 0.5 in metronomic walking. They concluded that "supraspinal influences [could] override the normally present long-range correlations" (p. 1456). In other words, the intention to walk in synchrony with the metronome was supposed to induce a kind of loss of complexity in the locomotor system, walking pace being externally regulated by the metronome. However, Delignières and Torre (2009), in a reassessment of the original data of Hausdorff and colleagues, showed that the series of stride intervals, when participants had to walk in synchronization with a metronome, where not uncorrelated as postulated by the authors, but presented a rather complex pattern of correlation. Especially the application of PSD revealed a log-log power spectrum with positive slope in low frequencies, and a slightly negative slope in the high frequency region (Figure 9).

Delignières and Torre (2009) tried to account for this result with a model initially proposed by West and Scafetta (2003), and composed of a forced van der Pol oscillator:

$$\ddot{x} = \gamma\dot{x} - \lambda\dot{x}x^2 - \omega_i^{\,2}x + A\sin(\omega_0) + \sqrt{Q\xi_t} \qquad (15)$$

This model is supposed to represent the dynamics of an oscillator, driven by a periodic forcing function. In this equation the stiffness parameter (ω_i) is cycle-dependent and possesses fractal properties. The forcing term $\sin(\omega_0)$ represents the metronome with a fixed frequency ω_0, and A represents the coupling strength. The authors showed that setting A to 1 for self-paced walking and to 10 for synchronized walking allowed reproducing the patterns of correlations experimentally observed. These results show that during synchronization, LRC are still present in the system, but that the process of synchronization induces a short-term corrective process that could hide their presence.

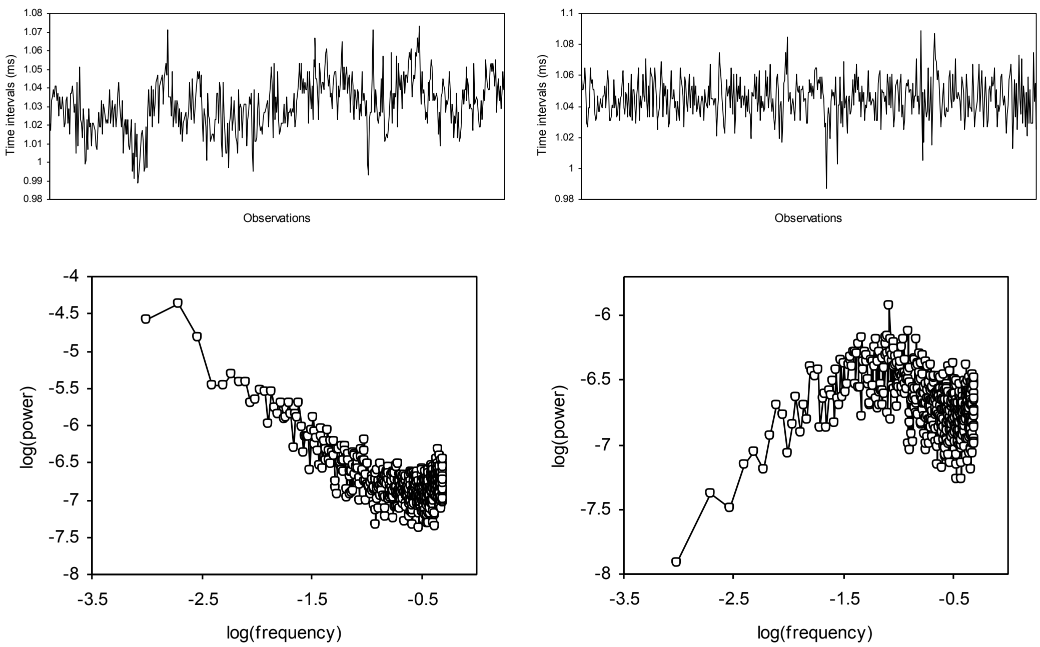

Figure 9: Analysis of serial correlations in stride interval series during self-paced walking (left column), and during walking in synchrony with a metronome (right column). Top raw: example time series, bottom raw: averaged log-log power spectra ($N = 10$, from Delignières & Torre, 2009).

Another example concerns finger tapping. Chen, Ding and Kelso (2001) showed that in synchronized finger tapping, the series of asynchronies to the metronome presented LRC. They also showed that the strength of LRC tended to increase when taps had to be performed in syncopation, i.e. in between metronome signals. The application of PSD revealed a steeper negative slope for syncopation than for synchronization (Figure 10). The authors attributed the strengthening of LRC to the increased task difficulty of tapping between the beats. In terms of system complexity, task difficulty could be supposed to solicit more complex coordination between system's components, yielding a more robust $1/f$ scaling.

However, the relationship between task difficulty and LRC is not so univocal: Sometimes an increased difficulty induces a strengthening of LRC, but in other experiments the opposite result can be obtained (Kello, Beltz, Holden & van Orden, 2007). Delignières, Torre and Lemoine (2009) tried to provide an alternative explanation for the increase of LRC in syncopation tapping. They proposed to account for asynchronies during tapping in synchronization by the previously presented model (Eq. 13), derived from an original model by Vorberg and Wing (1996) and enriched by providing the timekeeper with $1/f$ properties

The authors proposed that in syncopation conditions, another process should be added in the model, for accounting for the necessary estimation of the half-period of the metronome, and they suggested that this process could be considered correlated over time (for further details see Delignières, Torre & Lemoine, 2009). Simulation results showed that the combination of this short-range correlated process with the previous model induced an increase of correlations in the resulting series, similar to that empirically observed.

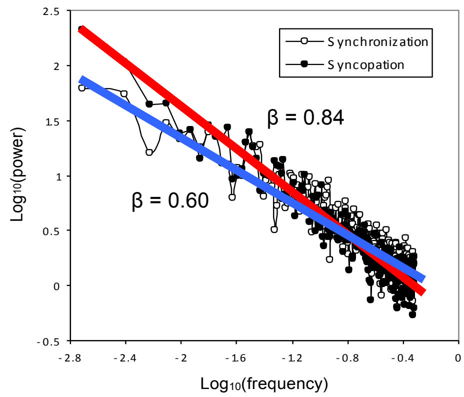

Figure 10: Averaged log-log power spectra for series of asynchronies obtained during tapping in synchronization and in syncopation with a metronome ($N = 11$; from Delignières, Torre & Lemoine, 2009).

In sum, these experiments suggest that short-range processes can either decrease or conversely increase correlations in the signal. Fractal analysis methods are often unable to discern and identify these short-range processes, and can often yield erroneous interpretations. The application of different methods, in the time domain and in the frequency domain, and additional studies based on modeling and simulation are generally necessary for identifying the multiple sources of serial correlation that combine in experimental series.

False detection of long-range correlations

An additional problem is that short-range correlation processes can sometimes mimic long-range correlations. In order to illustrate the possible misinterpretations that could arise with classical methods such as DFA or PSD, we simulated three time series, possessing distinct correlation properties. The first series was generated by a one-order autoregressive model:

$$y_i = \varphi y_{i-1} + \varepsilon_i \qquad (18)$$

In this equation φ is the autoregressive parameter and was set to 0.85. ε_i is a white noise process with zero mean and unit variance. The second series was generated by an integrated one-order moving average model:

$$y_i = y_{i-1} - \theta\varepsilon_{i-1} + \varepsilon_i \qquad (19)$$

In this equation θ is the first-order moving average parameter and was set to 0.8. ε_i is a white noise process with zero mean and unit variance. Finally we used the Harte-Davies algorithm for simulating a series of fractional Gaussian noise, with H = 0.9 (Harte & Davies, 1987). The three series are presented in Figure 11. By construction, the two first series present only short-term correlations, while the third one presents LRC.

The corresponding DFA diffusion plots are presented in the middle row of Figure 11. As can be seen in all cases the diffusion plot presents a linear slope close to 0.9. Obviously the best linear fit is observed for the fGn series, which contains genuine long-range correlations. However the diffusion plots obtained for auto-regressive and moving average series mimic the typical shape expected from long-range correlated series, and could lead to erroneous interpretations.

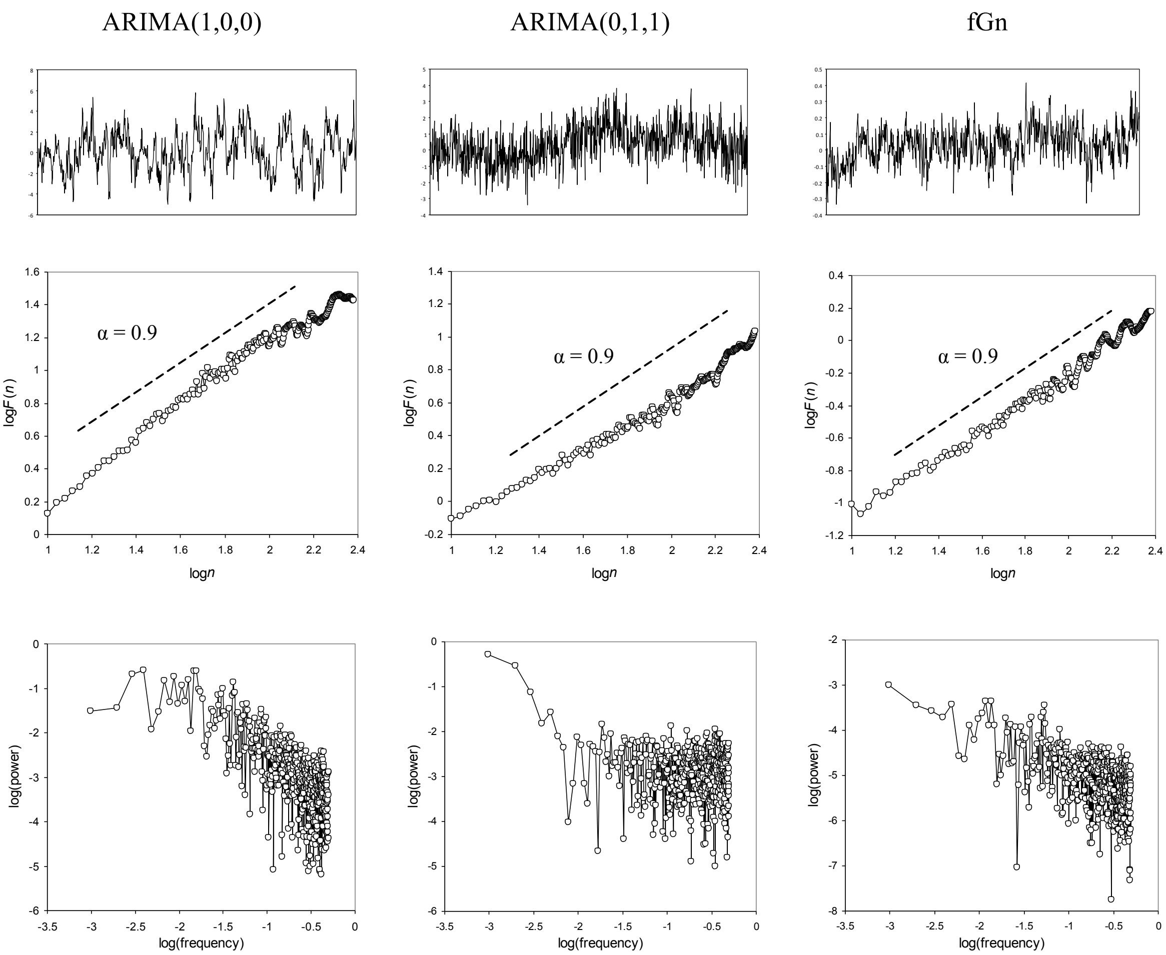

Figure 11: Example series simulated with a one-order autoregressive model ($y_i = 0.85y_{i-1} + \varepsilon_i$, left column), a one-order moving average model ($y_i = y_{i-1} - 0.8\varepsilon_{i-1} + \varepsilon_i$, central column), and the Harte-Davies algorithm (fGn with H = 0.9, right column). The corresponding DFA diffusion plots are represented in the median row, and the log-log power spectra in the bottom row.

Spectral analysis yields less ambiguous results (see Figure 11, bottom panels): for the auto-regressive series, the log-log spectrum presents a typical flattening in low frequencies, revealing the absence of correlation beyond a given time lag. The log-log spectrum of the integrated moving average series presents a flattening in high frequencies, due to the strong contribution of white noise terms in the generation of the process. In contrast, the spectrum of the fGn series exhibits a perfect linear slope over the whole range of frequencies. However, each spectrum presents a negative linear slope in a broad range of frequencies, and could induce an erroneous detection of LRC. Especially in the case of auto-regressive series, the flattening is often restricted to few points in the low frequency region, and the spectrum almost perfectly mimics a $1/f$ shape (Wagenmakers et al., 2004).

Evidencing the presence of long-range correlation

The main problem with graphical methods such as PSD or DFA is that these methods cannot provide a statistical evidence for the effective presence of LRC in a series. Sometimes surrogate tests are performed, consisting in comparing the result obtained with an experimental series with those obtained from a set of shuffled versions of the original series (Figure 12). In that case, the null hypothesis that is tested is the absence of correlation in the series. The test can eventually show that the series contains correlations, but cannot give any information about the (short or long-range) nature of these correlations.

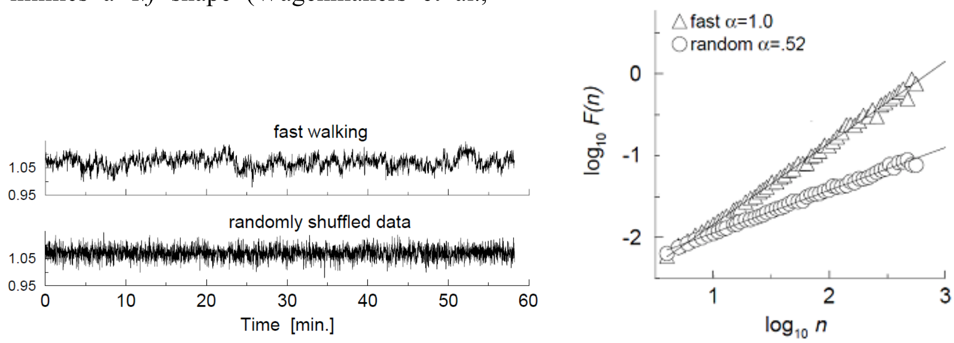

Figure 12: Surrogate tests. Left: an example walking stride interval series (top) and a randomly shuffled version (bottom). Right: DFA diffusion plots of the original and the shuffled series. The original series exhibits a $1/f$–like behavior, while the shuffled series yields a slope close to 0.5. From Hausdorff et al. (1996).

However, as stated by Slifkin and Newell (1998) the presence of correlations in empirical time series is more the rule than the exception. As such, the proper null hypothesis for testing for the presence of LRC is not the absence of serial correlation, but, the short-range nature of correlations in the series.

Some specific methods have been developed for testing for the effective presence of LRC. In particular Wagenmakers, Farrell and Ratcliff (2005) and Torre, Delignières and Lemoine (2007b) have proposed a method based on ARMA and ARFIMA modeling.

ARMA models have been introduced by Box and Jenkins (1976). An ARMA model is the combination of autoregressive and moving average processes, and is noted (p,q), in which p is the order of the auto-regressive process, and q the order of the moving average process. A more complex class of models (ARIMA) is characterized by the inclusion of an integration process allowing to represent trends in the series. An ARIMA model is noted (p,d,q), d representing the order of the integration process. In the original formulation p, d, and q are integers. Granger and Joyeux (1980) proposed to allow the integrating parameter d to take on fractional values. The resulting ARFIMA models (for *auto-regressive-fractionally integrated-moving average* models) present long-range correlation properties.

The method proposed by Wagenmakers et al. (2005) consists in fitting 18 models to the studied series. Nine of these models are ARMA (p,q) models, p and q varying systematically from 0 to 2. These ARMA models do not contain any long-range serial correlations. The other nine models are the corresponding ARFIMA (p,d,q) models, differing from the previous ARMA models by the inclusion of the fractional integration parameter d representing persistent serial correlations. One supposes that if the analyzed series contains LRC, ARFIMA models should present a better fit than their ARMA counterparts. The method selects the best model on the basis of a goodness-of-fit criterion (Bayes Information Criterion) based on a trade-off between accuracy and parsimony.

Torre et al. (2007b) proposed two complementary statistical criteria for ascertaining the presence of LRC: (1) The percentage of series better fitted by an ARFIMA rather than by an ARMA model, and (2) the mean sum of weights captured by ARFIMA models (the weight of a model is derived from the value of the goodness-of-fit criterion and represents the probability that this model is the best among the 18 candidate models for a given series). The authors proposed to accept the LRC hypothesis if at least 90% of the series are best fitted by an ARFIMA model, and if the mean sum of ARFIMA weights exceeds 0.90 (note that the total sum of weights for the 18 models is 1). The authors showed that this method allowed to correctly detecting LRC in simulated fGn series (see Figure 13, top panel). They also showed that the percentage of false detections, in pure autoregressive and moving average series, remained moderate, except for very low values of p and q (see Figure 13, bottom panels).

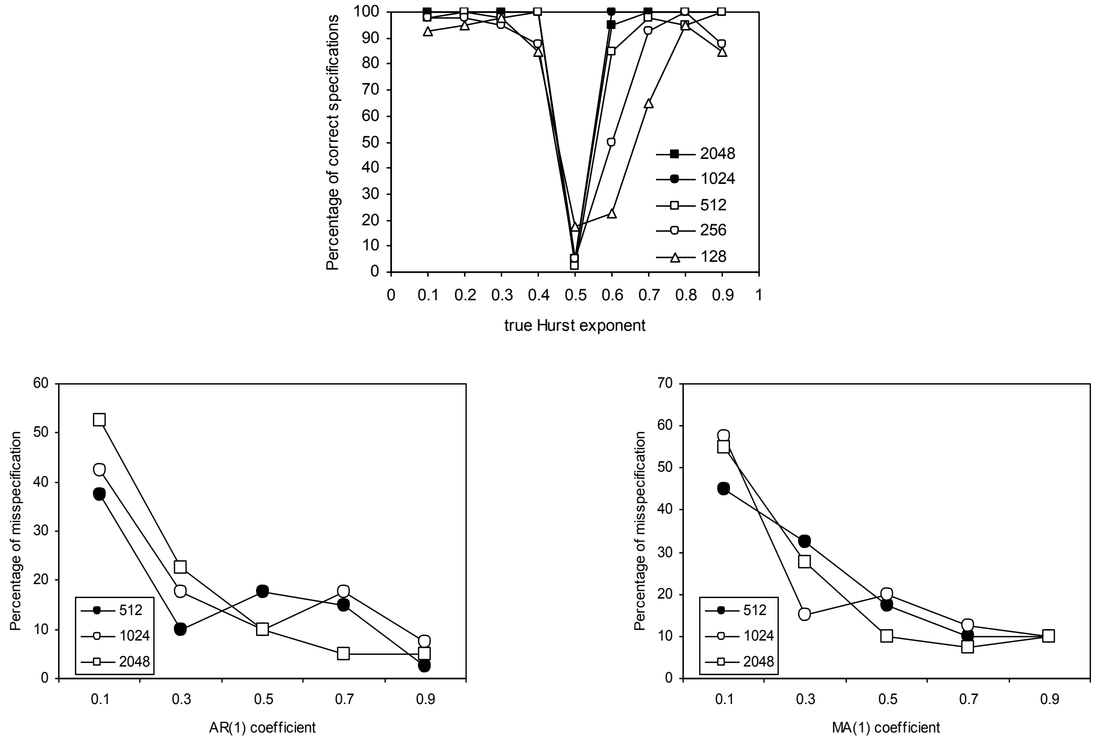

Figure 13: ARMA/ARFIMA modeling. Top panel: percentage of correct specifications of exact fGn series (from H = 0.1 to H = 0.9), for different series lengths (from N = 128 to N = 2048). The percentage of correct specifications exceeds 90%, except for very short, persistent series (H > 0.5, N = 128 and N = 256). Bottom panels: percentage of misspecifications for auto-regressive series (left) and moving average series (right), according to series length (from N = 512 to N = 2048) and to the value of the auto-regressive (AR(1)) and the moving average (MA(1)) parameters. The percentage of misspecifications remains moderate, except for very low values of AR(1) and MA(1). From Torre, Delignières and Lemoine, 2007.

Thornton and Gilden (2005) proposed an alternative method, a spectral likelihood classifier that used the shape of the power spectrum to decide among short- and long-range descriptions of data. Farrell, Wagenmakers and Ratcliff (2006) showed that the two methods gave roughly equivalent results.

As an example, we would like to present the results of an experiment during which participants had to walk on a treadmill at four different velocities (Marmelat & Delignières, 2011). These four velocities were individually determined, on the basis on a preliminary estimation of preferred velocity (the most comfortable and efficient velocity), and critical velocity (the velocity that induced a spontaneous transition towards running). v_1 was fixed at 80% of preferred velocity, v_2 corresponded to preferred velocity, v_3 was fixed at equal distance between preferred and critical velocities, and v_4 was the critical velocity. We collected 512 stride intervals in each conditions, and applied DFA and ARMA/ARFIMA modeling on the series.

DFA revealed an effect of walking velocity, with a gradual increase of serial correlations with increasing velocity (Figure 14), that could lead to a first conclusion: LRC tend to increase at high velocities. This first conclusion is consistent with those proposed in some previous studies (Hausdorff, Purdon, Peng, Ladin, Wei, & Goldberger, 1996; Jordan, Challis & Newell, 2006). However, the application of ARMA/ARFIMA modeling showed that the global likelihood of ARFIMA models exceeded the threshold of 90% and that the number of series best fitted by ARFIMA models reached 90% only at slow and preferred speeds (Figure 14). This yields a completely different conclusion: LRC tend to disappear at high velocities. This example clearly shows that the strength of serial correlation is not predictive of the genuine presence of LRC. Sometimes systems can present moderate levels of *effective* LRC, whereas in others cases, series can present high correlation levels but are not long-range correlated. Measuring strong correlations in a series does not guaranty that these correlations are long-range in nature.

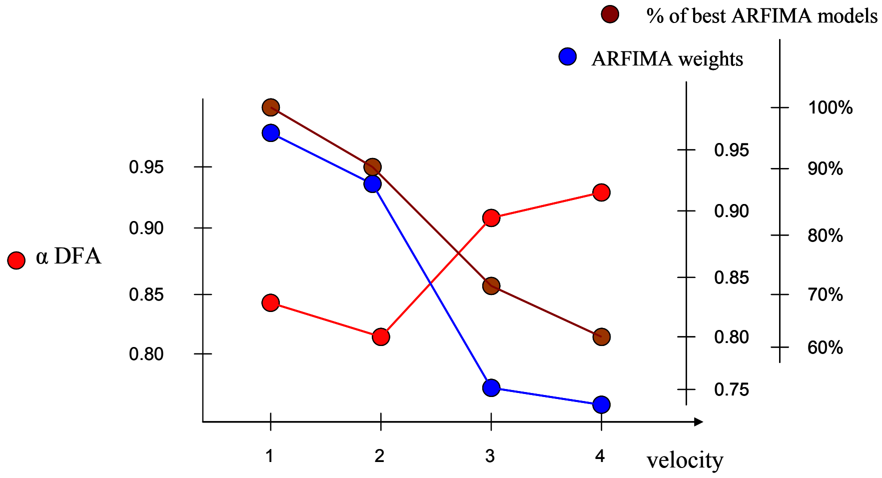

Figure 14: Analysis of stride interval series during walking on a treadmill. Effect of walking velocity on the scaling exponent (αDFA), the sum of ARFIMA weights, and the percentage of series best fitted by ARFIMA models.

Conclusion

The concept of LRC is currently particularly appealing for researchers. This approach allows to reveal new properties of the systems under study, and to question traditional theories (Torre & Wagenmakers, 2009). Especially the monofractal framework, considering 1/*f* fluctuations as an optimal compromise between order and disorder, suggests the adoption of new points of view concerning various key topics, including health (Goldberger, Amaral, Hausdorff, Ivanov, Peng, Stanley, 2002), adaptability (Harbourne & Stergiou, 2009), or learning (Wijnants, Bosman, Hasselman, Cox & Van Orden, 2010). Learning, for example, is generally conceived as the progressive installation of order in the system: with practice, trajectories become smoother, inter-trial variability decreases, etc. Wijnants et al. (2010) analyzed the effects of practice on LRC in the series of movement times in a cyclical aiming task. Their results showed that LRC tended to progressively increase, and reached levels close to 1/*f* noise after a sufficient amount of practice. This suggests that learning could be essentially conceived as an optimization of interaction dynamics within the system, and explains why LRC are generally evidenced in series produced in usually practiced and overlearned tasks.

LRC analyses could appear easy to apply and their results easy to interpret, in terms of deviation from the optimal 1/*f* noise. The examples developed in the present paper suggest, however, that the results provided by these analyses should be interpreted with caution. Especially 1/*f* fluctuations appear rarely in isolation in experimental series, and could be contaminated by various short-range correlated processes. The resulting correlation pattern is often complex, and can lead to erroneous conclusions.

References

Bassingthwaighte, J. B., Liebovitch, L. S., & West, B. J. (1994). *Fractal physiology* New York: Oxford University Press.

Beltz, B. B. & Kello, C. T. (2006). On the intrinsic fluctuations of human behavior. In M. Vanchevsky (Ed), *Focus on Cognitive Psychology Research* (pp. 25-41). Hauppauge, NY: Nova Science Publishers.

Box, G.E.P., & Jenkins, G.M. (1976). *Time series analysis: Forecasting and control.* Oakland: Holden-Day.

Caccia, D.C., Percival, D., Cannon, M.J., Raymond, G., & Bassingthwaigthe, J.B. (1997). Analyzing exact fractal time series: evaluating dispersional analysis and rescaled range methods. *Physica A, 246*, 609-632.

Cannon, M.J., Percival, D.B., Caccia, D.C., Raymond, G.M., & Bassingthwaighte, J.B. (1997). Evaluating scaled windowed variance methods for estimating the Hurst coefficient of time series. *Physica A, 241*, 606-626.

Chen, Y., Ding, M. & Kelso, J.A.S. (2001). Origins of timing errors in human sensorimotor coordination. Journal of Motor Behavior, 33, 3-8.

Delignières, D., & Torre, K (2009). Fractal dynamics of human gait: a reassessment of the 1996 data of Hausdorff et al. *Journal of Applied Physiology, 106,* 1272-1279.

Delignières, D., Ramdani, S., Lemoine, L., Torre, K., Fortes, M. & Ninot, G. (2006). Fractal analysis for short time series : A reassessement of classical methods. *Journal of Mathematical Psychology, 50,* 525-544.

Delignières, D., Fortes, M., & Ninot, G. (2004). The fractal dynamics of self-esteem and physical self. *Nonlinear Dynamics in Psychology and Life Science, 8,* 479-510.

Delignières, D., Lemoine, L. & Torre, K. (2004) Time intervals production in tapping and oscillatory motion. *Human Movement Science, 23*, 87-103.

Delignières, D., Marmelat, V., & Torre, K. (2011). *Degeneracy and long-range correlation: A simulation study*. Paper presented to the SKILLS Conference 2011, Montpellier, December 15-16, 2011.

Delignières, D. Torre, K. & Lemoine, L. (2009). Long-range correlation in synchronization and syncopation tapping: a linear phase correction model. *PLoS ONE* 4, 11, e7822.

Diniz, A., Wijnants, M.L., Torre, K., Barreiros, J., Crato, N., Bosman, A.M.T., Hasselman, F., Cox, R.F.A., Van Orden, G.C. & Delignières, D. (2011). Contemporary theories of 1/*f* noise in motor control. *Human Movement Science, 30,* 889-905.

Einstein, A. (1905). Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendieren Teilchen. *Annalen der Physik, 322*, 549-560.

Eke, A., Herman, P., Bassingthwaighte, J.B., Raymond, G.M., Percival, D.B., Cannon, M., Balla, I., & Ikrényi, C. (2000). Physiological time series: distinguishing fractal noises from motions. *Pflügers Archives, 439*, 403-415.

Eke, A., Hermann, P., Kocsis, L. & Kozak, L.R. (2002). Fractal characterization of complexity in temporal physiological signals. *Physiological Measurement, 23*, R1-R38.

Farrell, S., Wagenmakers, E.-J., & Ratcliff, R. (2006). Methods for detecting 1/f noise. Bristol, UK: University of Bristol. Available at: http://seis.bris.ac.uk/~pssaf/tgreplysims.pdf. Accessed on Septembre 3, 2011.

Gilden, D.L. (1997). Fluctuations in the time required for elementary decisions. *Psychological Science, 8*, 296-301.

Gilden, D.L., Thornton, T., & Mallon, M.W. (1995). 1/f noise in human cognition*. Science, 267*, 1837-1839.

Goldberger, A.L., Amaral, L.A.N., Hausdorff, J.M., Ivanov, P. Ch., Peng, C.-K. & Stanley, H.E. (2002). Fractal dynamics in physiology: Alterations with disease and aging. *PNAS, 99*, 2466–2472.

Goldberger, A.L., Peng, C.-K., & Lipsitz, L.A. (2002). What is physiologic complexity and how does it change with aging and disease? Neurobiology of Aging, 23, 23–26.

Granger, C.W.J., & Joyeux, R. (1980). An introduction to long-memory models and fractional differencing. *Journal of time Series Analysis, 1*, 15-29.

Harbourne, R.T. & Stergiou, N. (2009). Movement Variability and the Use of Nonlinear Tools: Principles to Guide Physical Therapist Practice. *Physical Ther*apy, 89, 267–282.

Hausdorff, J.M. (2005). Gait variability: methods, modeling and meaning. *Journal of NeuroEngineering and Rehabilitation* 2005, 2,19, doi:10.1186/1743-0003-2-19

Hausdorff, J.M., Purdon, P.L., Peng, C.K., Ladin, Z., Wei, J.Y. et Goldberger, A.R. (1996). Fractal dynamics of human gait: stability of long-range correlations in stride interval fluctuation. *Journal of Applied Physiology, 80*, 1448-1457.

Hausdorff, J.M., Mitchell, S.L., Firtion, R., Peng, C.K., Cudkowicz, M.E., Wei, J.Y. et Goldberger, A.L. (1997). Altered fractal dynamics of gait: reduced stride-interval correlations with aging and Huntington's disease*. Journal of Applied Physiology, 82*, 262-269.

Hausdorff, J.M., Peng, C.K., Ladin, Z., Wei, J.Y., & Goldberger, A.R. (1995). Is walking a random walk? Evidence for long-range correlations in stride interval of human gait. *Journal of Applied Physiology, 78*, 349-358.

Hurst, H.E. (1951). Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineering, 116*, 770-799.

Hurst, H.E. (1965). *Long-term storage: An experimental study*. London: Constable.

Ihlen, E. A. F., & Vereijken, B. (2010). Beyond $1/f^{\alpha}$ fluctuation in cognitive performance. *Journal of Experimental Psychology: General,* 139, 436–463.

Jordan, K., Challis, J.H., & Newell, K.M. (2006). Long range correlations in the stride interval of running. *Gait & Posture, 24*, 120–125.

Kello, C. T., Beltz, B. C., Holden, J. G., & Van Orden, G. C. (2007). The emergent coordination of cognitive function. *Journal of Experimental Psychology: General, 136*, 551-568.

Kello, C.T., Brown, G.D.A., Ferrer-i-Cancho, R., Holden, J.G., Linkenkaer-Hansen, K., Rhodes, T., & Van Orden, G. (2010). Scaling laws in cognitive sciences. Trends in Cognitive Sciences, 14, 223-232.

Leland, W.E., Taqqu, M.S., Willinger, W. & Wilson, D.V. (1994). On the self-similar nature of Ethernet traffic (extended version). IEEE/ACM Transactions on Networking, 2, 1-15.

Lemoine, L., Torre, K, & Delignières, D. (2006). Testing for the presence of 1/*f* noise in continuation tapping data. *Canadian Journal of Experimental Psychology, 60,* 247-257.

Mandelbrot, B.B., & van Ness, J.W. (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Review, 10*, 422-437.

Marmelat, V. & Delignières, D. (2011). Complexity, coordination and health: Avoiding pitfalls and erroneous interpretations in fractal analyses. *Medicina Lithuania*, *47,* 393-398.

Matsuzaki , M. (1994). Fractals in Earthquakes. *Philosophical Transactions: Physical Sciences and Engineering, 348,* 449-457.

Peng, C.-K., Mietus, J., Hausdorff, J.M., Havlin, S., Stanley, H.E., et Goldberger, A.L. (1993). Long-range anti-correlations and non-Gaussian behavior of the heartbeat. *Physical Review Letter*, 70, 1343-1346.

Pressing, J., & Jolley-Rogers, G. (1997). Spectral properties of human cognition and skill. *Biological Cybernetics, 76*, 339-347.

Slifkin, A.B,. & Newell, K.M. (1998). Is variability in human performance a reflection of system noise? *Current Directions in Psychological Science, 7*, 170-177.

Thornton, T. L., & Gilden, D. L. (2005). Provenance of correlations in psychophysical data. *Psychonomic Bulletin & Review, 12*, 409-441

Torre, K. & Wagenmakers, E.J. (2009). Theories and models for 1/*f* noise in human movement science. *Human Movement Science, 28,* 297-318.

Torre, K., & Delignières, D. (2008). Unraveling the finding of $1/f^{\beta}$ noise in self-paced and synchronized tapping: A unifying mechanistic model. *Biological Cybernetics, 99,* 159-170.

Torre, K, Delignières, D., & Lemoine, L. (2007a). $1/f^{\beta}$ fluctuations in bimanual coordination: An additional challenge for modeling. *Experimental Brain Research, 183,* 225-234.

Torre, K, Delignières, D., & Lemoine, L. (2007b). Detection of long-range dependence and estimation of fractal exponents through ARFIMA modeling. *British Journal of Mathematical and Statistical Psychology, 60,* 85-106.

Van Orden, G. C. (2007, February). The fractal picture of health and wellbeing. *Psychological Science Agenda, 21*(2). Available at http://www.apa.org/science/psa/vanorden.html

Van Orden, G.C., Holden, J.C., & Turvey, M.T. (2003). Self-organization of cognitive performance. *Journal of Experimental Psychology: General, 132*, 331-350.

Vorberg, D. & Wing, A. (1996). Modeling variability and dependence in timing. In H. Heuer & S.W. Keele (Eds.), *Handbook of perception and action, vol 2* (pp. 181-262). London: Academic Press.

Wagenmakers, E.-J., Farrell, S., & Ratcliff, R. (2005). Human cognition and a pile of sand: A discussion on serial correlations and self-organized criticality. *Journal of Experimental Psychology: General, 134*, 108-116.

Wagenmakers, E.-J., Farrell, S., & Ratcliff, R. (2004). Estimation and interpretation of $1/f^{\alpha}$ noise in human cognition. *Psychonomic Bulletin & Review, 11*, 579–615.

West, B.J. (1990). *Fractal Physiology and Chaos in Medicine,* World Scientific: Singapore

West, B. J. & Scafetta, N. (2003). Nonlinear dynamical model of human gait. *Physical Review E, 67*, 051917.

Wijnants, M.L., Bosman, A.M.T., Hasselman, F., Cox, R.F.A., et Van Orden, G.C. (2009). 1/*f* scaling in movement time changes with practice in precision aiming. *Nonlinear Dynamics in Psychology and Life Science, 13*, 79-98.

Wing, A.M. & Kristofferson, A.B. (1973). The timing of interresponse intervals. *Perception and Psychophysics, 13*, 455-460.