A non linear analysis of human gait time series based on multifractal analysis and cross correlations

A Muñoz-Diosdado
Department of Mathematics, Unidad Profesional Interdisciplinaria de Biotecnología, Instituto Politécnico Nacional, Av. Acueducto s/n, 07340, Mexico City, México
E-mail: amunoz@acei.upibi.ipn.mx

Abstract. We analyzed databases with gait time series of adults and persons with Parkinson, Huntington and amyotrophic lateral sclerosis (ALS) diseases. We obtained the staircase graphs of accumulated events that can be bounded by a straight line whose slope can be used to distinguish between gait time series from healthy and ill persons. The global Hurst exponent of these series do not show tendencies, we intend that this is because some gait time series have monofractal behavior and others have multifractal behavior so they cannot be characterized with a single Hurst exponent. We calculated the multifractal spectra, obtained the spectra width and found that the spectra of the healthy young persons are almost monofractal. The spectra of ill persons are wider than the spectra of healthy persons. In opposition to the interbeat time series where the pathology implies loss of multifractality, in the gait time series the multifractal behavior emerges with the pathology. Data were collected from healthy and ill subjects as they walked in a roughly circular path and they have sensors in both feet, so we have one time series for the left foot and other for the right foot. First, we analyzed these time series separately, and then we compared both results, with direct comparison and with a cross correlation analysis. We tried to find differences in both time series that can be used as indicators of equilibrium problems.

1. Introduction
A relevant extension of the fractal concept was the recognition that it applies not just to irregular geometric or anatomic forms that lack a characteristic scale of length, but also to complex processes that lack a single scale of time [1, 2]. The stride interval has been shown to fluctuate from one strike to the next in a complex fashion [3]. Hausdorff et al. detected fractal dynamics in the stride interval time series of human walking [4]. The analysis of these fluctuations revealed a self-similar pattern: fluctuations at one time scale are statistically similar to those at multiple other time scales, while healthy subjects walk at their normal rate. In healthy subjects, this variability is not simply attributable to random fluctuations or solely to short-range influences.

A lot of research on this field has been centered in the effect of the age on the human gait. For example, the children gait has been studied to detect some problem and to help them to acquire a healthy and mature gait. The gait of healthy adults has also been extensively studied to be able to evaluate the deviations of this healthy gait caused by neuro-degenerative diseases or lesions that affect the locomotor system. The knowledge of the factors that determine a healthy gait can help to the rehabilitation of ill or injured persons. This research has also been carried out for old persons.
When we analyze physiological time series we can obtain monofractal or multifractal behavior. Monofractals are characterized with the same scaling features throughout the entire signal; they can be indexed by a single global exponent $h_0$ (the Hurst exponent) or by a single fractal dimension $[5, 6]$, which indicates that they are stationary from the viewpoint of their local scaling properties. Multifractals have higher complexity, they can be decomposed into many subsets characterized by different local Hurst exponents $h$ or different fractal dimensions, which quantify the local singular behavior and thus relate to the local scaling of the time series. Multifractals require many exponents to characterize their scaling properties; they are intrinsically more complex and inhomogeneous than monofractals $[7, 8]$. Ivanov et al. $[9]$ have established the relevance of the multifractal formalism for the description of human interbeat time series. They also have discovered multifractality in other physiological time series and its breakdown with pathology. More recently, West and Scafetta $[10]$ have determined that gait time series, rather than being monofractal, are weakly multifractal. Indeed, we have checked this fact, because the multifractal spectra of the gait time series are always narrow.

In this paper we discuss multifractal properties of the stride interval time series. We found the opposite of that observed in the interbeat time series. The gait time series of healthy young persons have narrow multifractal spectra; these spectra almost have a monofractal behavior. The multifractal spectra get wider for old persons and they are also wider for time series of persons that have a disease affecting the gait (as Parkinson, Huntington or amyotrophic lateral sclerosis (ALS)).

2. The gait databases

We analyzed three gait time series of the reference $[11]$. In this site there are four gait databases: I. The Neuro-Degenerative Disease Database has records from patients with Parkinson’s disease ($n = 15$), Huntington’s disease ($n = 20$), or amyotrophic lateral sclerosis (ALS) ($n = 13$). The database has also records from 16 healthy control subjects. II. The gait in Aging and Disease Database is a "mini-collection" of human gait data. Walking stride interval time series are from 15 subjects: 5 healthy young adults (23 - 29 years old), 5 healthy old adults (71 - 77 years old), and 5 older adults (60 - 77 years old) with Parkinson’s disease. III. The Gait Maturation Database has the records of the gait cycle duration of 50 healthy children ranging in age from 40 months to 163 months. IV. The fourth database has gait long-term recordings of ten young, healthy men whose mean age was 21.7 years (range: 18-29 years). We considered for this analysis only the databases I, II and IV.

First, we analyzed the time series of these databases using the graphs of accumulated events and we calculated the Hurst exponent. Then we made the multifractal analysis of the series.

3. Analysis of cumulative plots and Hurst exponent

3.1. Cumulative plots

For a time series $x = \{x_k\}_{k=1}^N$, we obtain for $1 \leq n \leq N$, $\sigma(n) = \sum_{i=1}^n x_i$, and we plot $\sigma(n)$ versus $n$. These plots are like staircases and we found that in the long-term situation they can be bounded by a straight line in all cases as we show in figure 1. We obtain the slope of this straight line using the least squares linear fit. For the first database the results are shown in Table 1. We can see that the average slope is different for healthy people and ill people. These average slopes are different also for young and old persons. However, if we see with more detail the data, we note that some individual data are overlapped, this means that some slope values in the control group are close to slope values of persons with the Huntington, Parkinson or ALS diseases. This overlapping is also observed in the other databases, so in Table 2 we only show the average slope values for the other databases.
Figure 1. Cumulative plot for a time series of a 75 years old healthy patient.

Table 1. Slopes of the straight lines that bound the cumulative plots for time series of healthy persons (control) and with Parkinson’s and Huntington’s diseases and amyotrophic lateral sclerosis. First database.

|                | Slope     | Mean   |
|----------------|-----------|--------|
| Control        | 1.08      | 1.10   |
|                | 1.00      | 1.12   |
|                | 1.03      | 1.16   |
|                | 1.14      |        |
| Huntington     | 0.90      | 1.05   |
|                | 1.22      | 1.06   |
|                | 1.21      | 1.20   |
|                | 1.04      | 1.09   |
| Parkinson      | 1.14      | 1.24   |
|                | 1.01      | 1.06   |
|                | 1.22      | 1.03   |
| Sclerosis      | 1.30      | 2.01   |
|                | 1.15      | 1.32   |

3.2. Hurst exponent

Hurst [5] introduced the rescaled range analysis (R/S) and later Mandelbrot [1] and Feder [12] transplanted this analysis to the fractal analysis. For a time series \(1\) and any \(2 \leq n \leq N\), we can define, as usual [6],
\begin{align}
\langle x \rangle_n &= \frac{1}{n} \sum_{i=1}^{n} x_i,
\end{align}

\begin{align}
X(i,n) &= \sum_{i=1}^{n} [x_i - \langle x \rangle_n],
\end{align}

\begin{align}
R(n) &= \max_{i \leq n} X(i,n) - \min_{i \leq n} X(i,n),
\end{align}

\begin{align}
S(n) &= \left[ \frac{1}{n} \sum_{i=1}^{n} (x_i - \langle x \rangle_n)^2 \right]^{1/2}.
\end{align}

Hurst found that

\begin{align}
R(n)/S(n) \approx \left( \frac{n}{2} \right)^{H},
\end{align}

where \( H \) is the Hurst exponent. As \( n \) changes from 2 to \( N \), we obtain \( N-1 \) points in the \( \ln(n) \) versus \( \ln(R(n)/S(n)) \) plane; the plot is approximately a straight line. Then we calculate the Hurst exponent using the least square linear fit. The Hurst exponent is usually used as a measure of complexity, it has been found that there is a relation between \( H \) and the fractal dimension \( D \), expressed as \( D = 2 - H \) \[6\]. Hence a smaller \( H \) means a more complex system. Exponents greater than 0.5 indicate persistence (past trends persist into the future), whereas exponents less than 0.5 indicate antipersistence (past trends tend to reverse in the future).

**Table 2.** Slopes of the straight lines that bound the cumulative plots for time series of young and old healthy persons and with Parkinson’s disease (second database) and healthy youths of the fourth database.

| Database   | Average slopes |
|------------|----------------|
| Young (5)  | 1.10           |
| Old (5)    | 1.03           |
| Parkinson (5) | 1.16       |
| IV 10 youths | 1.12           |

We obtained the average Hurst exponent for the groups described earlier (Table 3). We obtained few information of this Table, there is antipersistence in all the gait time series and it seemed that the Hurst exponents of the time series belonging to the ill and old persons have the lowest values. However, when the \( \ln(n) \) versus \( \ln(R(n)/S(n)) \) graphs are obtained we see they have many irregularities (Figure 2).

The best Hurst exponent calculations were obtained for the series in the fourth database, this can be due to two factors: First, these series have more data, but also for other time series (in the control group of the first database) we obtained graphs with few irregularities, so we think this factor does not seem to be decisive. Our second hypothesis is that the behavior of these series is monofractal while those of more irregular aspect would be multifractal, we will test this hypothesis in the following section.

4. **Multifractal analysis**

Monofractals can be characterized by a single global Hurst exponent \( h_0 \), which indicates that they are stationary from the viewpoint of their local scaling properties. Multifractals can be decomposed into many subsets characterized by different local Hurst exponent \( h \). The statistical properties of the different subsets characterized by these different exponents can be quantified by the function \( f(h) \) or \( f(\alpha) \), \( (\alpha \) is the Lipschitz-Hölder exponent and \( f(\alpha) \) is the Hausdorff dimension of each subset), so
multifractals require many exponents to characterize their scaling properties and have a higher complexity than monofractals [1, 2]. In a recent study Ivanov et al. [9] have found that the healthy heartbeat is described by a broad range of spectral exponents $\alpha$ with a well-defined set of bounding parameters $\alpha_{\text{min}}$ and $\alpha_{\text{max}}$. The exponents occur with a given structure characterized by the function $f(\alpha)$. In contrast in subjects with a pathological condition (congestive heart failure), there is a loss of multifractality, $f(\alpha)$ is nonzero only over a very narrow range of exponents $\alpha$ indicating monofractal behavior.

### Table 3. Slopes of the straight lines that bound the cumulative plots for time series of young and old healthy persons and with Parkinson’s disease (second database), children (third database) and healthy youths of the fourth database.

| Database | Average Hurst exponents |
|----------|-------------------------|
| Control  | 0.126                   |
| I        | Huntingdon 0.046        |
|          | Parkinson 0.033         |
|          | ALS 0.083               |
|          | Young (5) 0.096         |
| II       | Old (5) 0.019          |
|          | Parkinson (5) 0.012     |
| IV       | 10 youths 0.062        |

![Figure 2](image.png)

**Figure 2.** Calculation of the Hurst exponent for a 66 year-old person with the Huntington disease. For real data, this plot falls only approximately along a straight line.

We obtained the multifractal spectra of the gait time series. If the spectra are narrow we can assure that the behaviour is monofractal or almost monofractal. But if the spectra are wide we can think that a multifractal description is needed. To calculate the multifractal spectrum we used the method
proposed by Chhabra and Jensen [13, 14, 15]. We can consider the time series as a singular measure \( P(x) \) if we normalize the series. We calculate the \( f(\alpha) \) curve first covering the measure with boxes of length \( L=2^{-n} \) and computing the probabilities \( P_j(L) \) in each of the boxes. We then construct the one-parameter family of normalized measures with

\[
\mu_i(q, L) = \frac{[P_i(L)]^q}{\sum_j [P_j(L)]^q}
\]

Finally, for each value of \( q \) we evaluate the numerators on the right-hand sides of the equations:

\[
f(q) = \lim_{L \to 0} \frac{\sum \mu_i(q, L) \ln[\mu_i(q, L)]}{\ln L},
\]

\[
\alpha(q) = \lim_{L \to 0} \frac{\sum \mu_i(q, L) \ln[P_i(L)]}{\ln L},
\]

for decreasing box sizes (increasing \( n \)). We extract \( f(q) \) and \( \alpha(q) \) from the numerator slopes versus \( \ln L \). The parameter \( q \) provides a microscope for exploring different regions of the singular measure. For \( q>1 \), \( \mu(q) \) amplifies the more singular regions of \( P \), while for \( q<1 \) it accentuates the less singular regions, and for \( q=1 \) the measure \( \mu(1) \) replicates the original measure.

We first analysed the fourth database to have a reference, because as we mentioned in the last section it is very probable that only only a Hurst exponent can describe these series. We obtained for the ten series very narrow multifractal spectra, \( q \) from \(-30 \) to \( 30 \), and we obtained \( \alpha_{\min} \) and \( \alpha_{\max} \); \( \Delta \alpha = \alpha_{\max} - \alpha_{\min} \) is the width of the spectra. The average \( \Delta \alpha \) is 0.016.

We calculated the multifractal spectra of the time series of the other databases and obtained the width of the spectra, the results are shown in table 4. In figure 3 we show in the top the multifractal spectrum for a gait time series of a young person and in the bottom the spectrum for a gait time series of a person with the Parkinson disease. Both spectra are narrow, but it is still more narrow the one that belongs to the healthy person. This suggests that the gait time series corresponding to the ill person are more complex, since they are necessary more fractal dimensions to describe them. For example, we calculated the \( \alpha_{\min} \) and the \( \alpha_{\max} \) for both spectra showed in figure 4. For gait time series of the young person \( \alpha_{\min}=0.988 \), \( \alpha_{\max}=1.010 \), so \( \Delta \alpha = 0.022 \); and for gait time series of the person with the Parkinson disease \( \alpha_{\min}=0.937 \), \( \alpha_{\max}=1.085 \), so \( \Delta \alpha = 0.148 \), and this is almost 7 times the other \( \Delta \alpha \), i.e. the first multifractal spectrum is almost 7 times wider the second one.

**Table 4.** Average \( \Delta \alpha \) for different groups of persons in the databases.

| Group                                      | Average \( \Delta \alpha \) |
|--------------------------------------------|-------------------------------|
| 10 young healthy persons (fourth database) | 0.016                         |
| 5 healthy elderly persons (second database)| 0.034                         |
| 16 healthy persons (control group in first database) | 0.048                     |
| 15 patients with Parkinson's disease (first database) | 0.092                     |
| 20 patients with Huntington's disease (first database) | 0.108                     |
| 13 patients with ALS (first database)       | 0.211                         |

We can show the multifractal spectra with color panels by assigning a color to each value of the Lipschitz-Hölder exponent, if we assign blue color to \( \alpha_{\min} \) and red color to \( \alpha_{\max} \) and other colors to intermediate values in a sequence similar to the disposition in the visible electromagnetic spectrum then we can obtain color panels [7,15]. These panels are “optical” spectra of the series; these spectral lines have associated a fractal dimension, which is the support dimension of the probability, which is characterized by the Lipschitz-Hölder exponent.
Figure 3. Multifractal spectra for a gait time series of a young person (top) and a person with the Parkinson disease (bottom).

5. Discussion
The study of gait time series by using non-multifractal analysis, as the graphs of accumulated events and the calculation of the global Hurst exponent allow us to make the distinction of the time series. Although the slope average values of the straight line that bounds the graphs of accumulated events are different for healthy and ill persons, when we consider the individual values, some of them are overlapped. We can see in table 4, that the multifractal spectra of persons with a pathology that affects the gait are wider than the multifractal spectra of healthy persons. The spectra of healthy young
persons are almost monofractal. So in human gait time series the multifractal behavior emerges with the pathology.

The study of cross correlations did not give significant results. As me mentioned before, the data were collected from healthy and ill subjects as they walked in a roughly circular path and they have sensors in both feet, so we have one time series for the left foot and other for the right foot. First, we analyzed these time series separately, using cumulative plots, Hurst exponent and multifractal analysis, then we compared both results, with direct comparison and with a cross correlation analysis. There are some differences, mainly in the time series of persons with ALS, but these differences are not statistically significant. We tried to find these differences in both time series that can be used as indicators of equilibrium problems. These differences are very small, so we can conclude that the aging and these three diseases (Huntington, Parkinson and ALS) affect the gait, because they deteriorate the walking ability, but it seems to be that they do not cause equilibrium problems.

6. Conclusions
We have shown that the gait time series of healthy individuals are less complex than those of ill persons. This trend up to the gain of complexity with the illness is somewhat in the opposite direction present in the analysis of the interbeat cardiac series. Maybe this is related with the fact that in this case we are mainly treating with the degradation of a mechanical system. When we carry out the calculations of the suitable parameters if we find an anomalous width multifractal spectrum, it can indicate some problem in the gait, mainly in the children. We can calculate other parameters and use other methodologies, but seemingly the multifractal approach can separate between the ill and healthy gait time series. Making this fractal and multifractal analysis, could be useful for a bigger understanding of the human gait, it can also help in the evaluation and improvement of rehabilitation treatments and in prosthesis adjustment.

Acknowledgements
We acknowledge EDI, EDD and COFFA-IPN Mexico for partial support.

References
[1] Mandelbrot B B 1982 The Fractal Geometry of Nature ed W H Freeman (New York)
[2] Bolis C L and Licinio J eds 1999 The Autonomic Nervous System (Geneva: World Health Organization)
[3] Hausdorff J M, Peng C K, Ladin Z, Wei J Y and Goldberger A L 1995 Is walking a random walk? Evidence for long-range correlations in the stride interval of human gait J Appl Physiol 78 349-358
[4] Hausdorff J M, Lertratanakul A., Cudkowicz M E, Peterson A L., Kaliton D and Goldberger A L 2000 Dynamics markers of altered gait rhythm in amyotrophic lateral sclerosis J Appl Physiol 88 2045-2053
[5] Hurst H E 1951 Long-term storage capacity of reservoirs Trans Amer Civil Energ 116 770-808
[6] Yu Z G and Wang B 2001 A time series model for CDS sequences in complete genome Chaos, Solitons and Fractals 12 519-526
[7] Bunde A, Knopp J and Schellnhuber H J eds 2002 The Science of Disasters (Germany: Springer-Verlag)
[8] Malamud B D, Turcotte D L 1999 Self-affine time series: I. Generation and analysis Advances in Geophysics 40 1-90
[9] Ivanov P Ch, Nunes Amaral A, Goldberger A L and Stanley H E Stochastic feedback and the regulation of biological rhythms Europhys Lett 43 (4) 363-368
[10] West B J, Scafetta N 2003 Nonlinear dynamical model of human gait Phys Rev E 67 051917.
[11] Goldberger A L, Amaral L A N, Glass L, Hausdorff J M, Ivanov P Ch, Mark R G, Mietus J E, Moody G B, Peng C K and Stanley H E 2000 PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals Circulation 101 (23) e215-e220. http://circ.ahajournals.org/cgi/content/full/101/23/e215
[12] Feder J 1988 Fractals (New York: Plenum Press)
[13] Chhabra A, Jensen R V 1989 Direct determination of the $f(\alpha)$ singularity spectrum Phys Rev Lett 62 (12) 1327-1330
[14] Chhabra A, Menevau C, Jensen R V and Sreenivasan K R 1989 Direct determination of the $f(\alpha)$ singularity spectrum and its application to fully developed turbulence Phys. Rev. A 40 5284-5294
[15] Del Río Correa J L and Muñoz-Diosdado A 2002 Multifractality in physiological time series AIP Conference Proceedings 630 (1) 191-201