Medical Engineering & Physics 36 (2014) 761–767

Contents lists available at ScienceDirect

Medical Engineering & Physics

journal homepage: www.elsevier.com/locate/medengphy

Technical note

Performance analysis of four nonlinearity analysis methods using a model with variable complexity and application to uterine EMG signals

Ahmad Diab [a,b,*], Mahmoud Hassan [c], Catherine Marque [a], Brynjar Karlsson [b]

[a] UMR CNRS 7338, Biomécanique et Bio-ingénierie, Université de Technologie de Compiègne, Compiègne, France
[b] School of Science and Engineering, Reykjavik University, Reykjavik, Iceland
[c] Laboratoire Traitement du Signal et de L'Image, INSERM, Université de Rennes 1, Campus de Beaulieu, Rennes, France

ARTICLE INFO

Article history:
Received 29 November 2012
Received in revised form 4 October 2013
Accepted 27 January 2014

Keywords:
Nonlinear time series analysis
Uterine electromyogram
Contraction discrimination
Surrogates

ABSTRACT

Several measures have been proposed to detect nonlinear characteristics in time series. Results on time series, multiple surrogates and their *z*-score are used to statistically test for the presence or absence of non-linearity. The *z*-score itself has sometimes been used as a measure of nonlinearity. The sensitivity of nonlinear methods to the nonlinearity level and their robustness to noise have rarely been evaluated in the past. While surrogates are important tools to rigorously detect nonlinearity, their usefulness for evaluating the level of nonlinearity is not clear. In this paper we investigate the performance of four methods arising from three families that are widely used in non-linearity detection: statistics (time reversibility), predictability (sample entropy, delay vector variance) and chaos theory (Lyapunov exponents). We used sensitivity to increasing complexity and the mean square error (MSE) of Monte Carlo instances for quantitative comparison of their performances. These methods were applied to a Henon nonlinear synthetic model in which we can vary the complexity degree (*CD*). This was done first by applying the methods directly to the signal and then using the *z*-score (surrogates) with and without added noise. The methods were then applied to real uterine EMG signals and used to distinguish between pregnancy and labor contraction bursts. The discrimination performances were compared to linear frequency based methods classically used for the same purpose such as mean power frequency (MPF), peak frequency (PF) and median frequency (MF). The results show noticeable difference between different methods, with a clear superiority of some of the nonlinear methods (time reversibility, Lyapunov exponents) over the linear methods. Applying the methods directly to the signals gave better results than using the *z*-score, except for sample entropy.

1. Introduction

One of the most common ways of obtaining information on neurophysiologic systems is to study the features of the signal(s) using time series analysis techniques. This traditionally rely on linear methods in both time and frequency domains [1]. Unfortunately, these methods cannot give information about purely nonlinear features of the signal. Due to the intrinsic nonlinearity of most biological systems, these nonlinear features may be present in physiological data and even be a characteristic of major interest. Recently, much attention has been paid to the use of nonlinear analysis techniques for the characterization of a biological signal [2]. Indeed, this type of analysis gives information about the nonlinear features of these signals, which arise from the underlying physiological processes, many of which have complex behavior. There is a growing literature reporting nonlinear analysis of various biosignal types (EEG [3], ECG [4], HRV [5] and EMG [6]).

The EHG or electrohysterogram (electrical uterine activity recorded on woman's abdomen) has been widely studied [7–11]. Nonlinear characteristics have been observed in the EHG and some success has been achieved by using these characteristics to obtain information of potential clinical usefulness. Radomski et al. show that nonlinear analysis of EHG based on the sample entropy statistic could differentiate dynamic states of uterine contractions [12]. A comparison between linear and nonlinear analysis with

* Corresponding author at: UMR CNRS 7338, Biomécanique et Bio-ingénierie, Université de Technologie de Compiègne, Compiègne, France. Tel.: +33 612295638.
E-mail addresses: ahmad.diab@utc.fr, ahmaddiab_87@hotmail.com (A. Diab), mahmoud.hassan@univ-rennes1.fr (M. Hassan), catherine.marque@utc.fr (C. Marque), brynjar@ru.is (B. Karlsson).

1350-4533/$ – see front matter.
http://dx.doi.org/10.1016/j.medengphy.2014.01.009

different conditions was done in [13]. It was concluded that median frequency is the best method among linear methods and that sample entropy is the best method among nonlinear methods for term/preterm EHG contractions classification. Sample entropy is superior to median frequency, which indicates that nonlinear analysis is more suitable than linear analysis for studying EHG signals. In [14] the progress of labor was evaluated using sample entropy. Our team has examined nonlinear EHG analysis methods. Our results confirm the presence of nonlinearity in EHG signals. This character of the signals is useful in discriminating between pregnancy and labor contractions [15,2,16]. Practical disadvantages of the nonlinear analysis methods have been reported in [16]. They include excessive calculation time due to surrogates analysis and promising but inconclusive results due to the small amount of data that can practically be used due to heavy calculation times.

This paper presents work that extends previous work done in our group in comparing approximate entropy, correntropy and time reversibility [16]. In this work we implemented additional nonlinear analysis methods (delay vector variance, Lyapunov exponents) and new ways of testing them. We also used a larger database of real signals than in the previous work and we investigated the sensitivity of the methods to the varying complexity of signals and their robustness. The kind of sensitivity and robustness analyses of non-linearity measures presented in this paper, are rare or absent in the literature.

Four nonlinear methods: time reversibility [17], sample entropy [18], delay vector variance [19] and Lyapunov exponents [20] were used in this work. Sensitivity of these methods to the complexity degree (*CD*) of a signal as well as robustness analysis was done on Henon model synthetic signals where *CD* can be controlled. The sensitivity to *CD* was first studied using the direct value provided by the method. It was then studied using surrogates and *z*-score, as the measure permitting evaluation of the nonlinearity. One objective of this study is to show which method(s) is most sensitive to the change of signal complexity. A second objective is to determine whether the use of surrogates gives better overall results than the direct application of the methods. This is of major practical importance for clinical application, as the generation of surrogates is very computationally expensive. The methods are also compared using the Mean square error (MSE) of the method results for 30 Monte Carlo instances of the signal. Finally, these non-linear methods are compared to three linear frequency based characteristics of the signal, MPF, PF and MF, when applied to real EHG signals, in order to discriminate pregnancy and labor contractions.

2. Materials and methods

2.1. Data

2.1.1. Synthetic signals

The Henon map is a well-known two-dimensional discrete-time system given by:

$$Y_{t+1} = c - Y_t^2 + CD \times X_t,$$

$$X_{t+1} = Y_t,$$

where Y_t and X_t represent dynamical variables, *CD* is the complexity degree and *c* is the dissipation parameter. In this paper we use $c = 1$ as in [21] and $CD \in [0,1]$ to change the model complexity [22] (Fig. 1). The number of generated points is fixed to 1000. For the robustness analysis, we add to the synthetic signals a white Gaussian noise with the same duration, with a fixed 5db SNR with *CD* varying between 0 and 1 with a step 0.1. In the Monte Carlo analysis, we use 30 signals generated for each *CD* value.

2.1.2. Real signals

EHG signals were recorded from 38 subjects using a 4×4 electrode matrix located on the subject's abdomen (Fig. 2), during 1 h either at rest (woman lying on a bed) or during labor. One signal channel (bipolar vertical 7: BP7), located on the median vertical axis of the uterus was used for subsequent analysis (see [23] for details). After segmentation we obtained 115 labor bursts (recorded during delivery) and 174 pregnancy bursts (recorded more than 24 h before delivery).

2.2. Non-linear analysis methods

2.2.1. Statistics family

2.2.1.1. Time reversibility. A time series is said to be reversible only if its probabilistic properties are invariant with respect to time reversal. Time irreversibility can be taken as a strong signature of nonlinearity [17]. In this paper we used the simplest method, described in [24] to compute the time reversibility of a signal S_n:

$$Tr(\tau) = \left(\frac{1}{N-\tau}\right) \sum_{n=\tau+1}^{N} (S_n - S_{n-\tau})^3$$

where *N* is the signal length and τ is the time delay.

2.2.2. Chaos theory family

2.2.2.1. Lyapunov exponents. Lyapunov exponent (LE) is a quantitative indicator of system dynamics, which characterizes the average convergence or divergence rate between adjacent tracks in phase space [20]. We used the method described in [13] to compute LE:

$$\lambda = \lim_{t \to \infty} \lim_{||\Delta y_0|| \to 0} \left(\frac{1}{t}\right) \log \frac{||\Delta y_t||}{||\Delta y_0||},$$

where $||\Delta y_0||$ and $||\Delta y_t||$ represent the Euclidean distance between two states of the system, respectively to an arbitrary time t_0 and a later time *t*.

2.2.3. Predictability family

2.2.3.1. Sample entropy. Sample entropy (*SampEn*) is the negative natural logarithm of the conditional probability that a dataset of length *N*, having repeated itself for *m* samples within a tolerance *r*, will also repeat itself for $m+1$ samples. Thus, a lower value of *SampEn* indicates more regularity in the time series [18]. We used the method described in [12] to compute *SampEn*:

For a time series of *N* points, $x_1, x_2, \ldots, x_N$, we define subsequences, also called template vectors, of length *m*, given by: $y_i(m) = (x_i, x_{i+1}, \ldots, x_{i+m-1})$ where $i = 1, 2, \ldots, N-m+1$.

Then the following quantity is defined: $B_i^m(r)$ as $(N-m-1)^{-1}$ times the number of vectors X_j^m within *r* of X_i^m, where *j* ranges from 1 to $N-m$, and $j \neq i$, to exclude self-matches, and then define:

$$B^m(r) = \frac{1}{N-m} \sum_{i=1}^{N-m} B_i^m(r)$$

Similarly, we define $A_i^m(r)$ as $(N-m-1)^{-1}$ times the number of vectors X_j^{m+1} within *r* of X_i^{m+1}, where *j* ranges from 1 to $N-m$, where $j \neq i$, and set

$$A^m(r) = \frac{1}{N-m} \sum_{i=1}^{N-m} A_i^m(r)$$