Physics Letters A ••• (••••) •••–•••

Contents lists available at ScienceDirect

Physics Letters A

www.elsevier.com/locate/pla

Detecting nonlinearity in short and noisy time series using the permutation entropy

Luciano Zunino [a,b,*], Christopher W. Kulp [c]

[a] Centro de Investigaciones Ópticas (CONICET La Plata – CIC), C.C. 3, 1897 Gonnet, Argentina
[b] Departamento de Ciencias Básicas, Facultad de Ingeniería, Universidad Nacional de La Plata (UNLP), 1900 La Plata, Argentina
[c] The Department of Astronomy and Physics, Lycoming College, Williamsport, PA 17701, USA

A R T I C L E I N F O

Article history:
Received 3 July 2017
Received in revised form 17 September 2017
Accepted 18 September 2017
Available online xxxx
Communicated by C.R. Doering

Keywords:
Time series analysis
Permutation entropy
Nonlinearity
Surrogate method

A B S T R A C T

Permutation entropy contains the information about the temporal structure associated with the underlying dynamics of a time series. Its estimation is simple, and because it is based on the comparison of neighboring values, it becomes significantly robust to noise. It is also computationally efficient and invariant with respect to nonlinear monotonous transformations. For all these reasons, the permutation entropy seems to be particularly suitable as a discriminative measure for unveiling nonlinear dynamics in arbitrary real-world data. In this paper, we study the efficacy of a conventional surrogate method with a linear stochastic process as the null hypothesis but implementing the permutation entropy as a nonlinearity measure. Its discriminative power is tested by implementing several analyses on numerical signals whose dynamical properties are known *a priori* (linear discrete and continuous models, chaotic regimes of discrete and continuous systems). The performance of the proposed approach in real-world applications (chaotic laser data, monthly smoothed sunspot index and neuro-physiological recordings) is also demonstrated. The results obtained allow us to conclude that this symbolic tool is very useful for discriminating nonlinear characteristics in very short and noisy data.

1. Introduction

Determining whether a given time series comes from a deterministic chaotic or a stochastic system can be a big challenge [1–3]. It is well-known that nonlinearity is a necessary condition for chaos. Consequently, to determine if an arbitrary time series is compatible with chaotic dynamics for modeling and classification purposes, it is first necessary to demonstrate that the dynamics producing the time series is, in fact, nonlinear. Furthermore, the detection of nonlinearity is not a trivial task especially for experimental records that are often contaminated with unknown noise sources. Motivated by these facts, in the last decade, several techniques for identifying nonlinear processes in observational data have been introduced [4–8]. Despite the existing contributions, discriminating the nonlinear dynamics of a complex system from time series is still a challenging problem of current research [9].

In this paper, we implement and test the efficacy of the permutation entropy (PE) as a discriminating statistic in a standard surrogate framework [10] in order to detect nonlinearities in short and noisy time series data. The PE is the Shannon entropic measure evaluated using the successful encoding introduced by Bandt and Pompe (BP) [11] to extract the probability distribution associated with an input signal. Taking into account the widely recognized practical advantages of this symbolic information-theory quantifier, namely, *i)* simplicity, *ii)* low computational cost, *iii)* noise robustness, and *iv)* invariance with respect to nonlinear monotonous transformations, PE is demonstrated to be an alternative and/or complementary approach to more traditional techniques for unveiling nonlinear structures from complex systems. The proposed nonlinearity test relies on the well-established method of surrogate data [10] just as many other nonlinear discriminating approaches [4–6,12], and, obviously, the generation of proper surrogates is essential for the test's success. As will be shown below, linear and nonlinear short noisy scalar time series can be efficiently characterized supporting a remarkable reliability of PE as a discriminator in practical contexts.

Even though the permutation entropy has been used in countless applications, it has been rarely implemented within a surrogate framework for unveiling the nonlinear dynamics of complex systems from time series. The analysis developed by Tony et al. [13], to identify the deterministic nature of pressure measure-

* Corresponding author at: Departamento de Ciencias Básicas, Facultad de Ingeniería, Universidad Nacional de La Plata (UNLP), 1900 La Plata, Argentina.
E-mail addresses: lucianoz@ciop.unlp.edu.ar (L. Zunino), kulp@lycoming.edu (C.W. Kulp).

http://dx.doi.org/10.1016/j.physleta.2017.09.032
0375-9601/

ments from a turbulent combustor, is one of these rare exceptions. Taking into account that, to the best of our knowledge, the performance of both techniques (permutation entropy and surrogate testing) together, as an unified approach, has not been analyzed in depth before, in this work we try to fill this gap through several numerical and real-world data tests. As it will be shown below, this approach is able to unveil the presence of nonlinear dynamics even in very short and noisy time series.

The remainder of the paper is organized as follows. In the next section, PE and the surrogate data analysis are briefly presented. Numerical and experimental analyses for testing the performance of the proposed nonlinearity test are detailed in Sections 3 and 4, respectively. Finally, Section 5 summarizes the results and contains concluding remarks.

2. Methodology

2.1. Permutation entropy

The symbolic encoding scheme due to BP [11], based on the ordinal relation between the amplitude of neighboring values of a given data sequence, has been implemented for estimating several information-theory quantifiers from time series. The BP ordinal method of symbolization naturally arises from the time series, inherits the causal information that stems from the temporal structure of the system dynamics, and also, avoids amplitude threshold dependencies that affect other more conventional symbolization recipes based on range partitioning [14,15]. These traits could be the main reasons behind notable success, as is evidenced by the enormous amount of applications in heterogeneous fields (see, *e.g.*, [16–26]). Furthermore, the ordinal pattern distribution is invariant with respect to nonlinear monotonous transformations. Thus, nonlinear drifts or scalings artificially introduced by a measurement device do not modify the quantifiers' estimations. It appears to be better suited to cope with usual problems (non-stationarities, nonlinearities, noise distortions) encountered when studying real time series compared to range-based encoding methods. Within this appealing encoding procedure, PE is undoubtedly the most widely-used descriptor. It should be stressed here that this entropic measure is applicable to noisy real time series from all class of systems, deterministic and stochastic, without the need to require any knowledge of the underlying mechanisms. As stated by Garland et al. [27], "It does not rely on generating partitions, and thus does not introduce bias into the results if one does not know the dynamics or cannot compute the partition. Permutation entropy makes no assumptions about, and requires no knowledge of, the underlying generating process: linear, nonlinear, the Lyapunov spectrum, etc." Furthermore, the relationship between the PE and the Kolmogorov–Sinai (KS) entropy has been discussed by several authors before. Basically, the growth rate of the PE is often used as a proxy for the KS entropy [28]. For more details, please see Refs. [29–31]. The KS entropy, probably the most appropriate indicator for distinguishing irregular deterministic from stochastic dynamics, requires specific knowledge of the generating process for its correct estimation. Finding the generating partition is not feasible for experimental data since they are inevitably contaminated with noise [32]. Consequently, the PE, which does not rely on generating partitions, emerges as a practical alternative to characterize data sets generated by an unknown dynamic process with unknown levels of noise [9].

Here, we will illustrate how to create ordinal patterns from the time series data with a simple example. Let us assume that we start with the time series $X = \{3, 2, 5, 8, 9, 6, 1\}$. In order to symbolize the series into ordinal patterns, first, two parameters, the order of the permutation symbols $D > 1$ ($D \in \mathbb{N}$, number of elements that form the ordinal pattern) and the lag τ ($\tau \in \mathbb{N}$, time separation between elements) are chosen. Next, the time series is partitioned into subsets of length D with lag τ similarly to phase space reconstruction by means of time-delay-embedding. The elements in each new partition (of length D) are replaced by their rank in the subset. For example, if we set $D = 3$ and $\tau = 1$, there are five different three-dimensional vectors associated with X. The first one $(x_0, x_1, x_2) = (3, 2, 5)$ is mapped to the ordinal pattern (102). The second three-dimensional vector is $(x_0, x_1, x_2) = (2, 5, 8)$, and (012) will be its related permutation. The procedure continues so on until the last sequence, $(9, 6, 1)$, is mapped to its corresponding motif, (210). In the case of two elements in the vector having the same value, the elements are ranked by index, for example, a vector $(7, 8, 7)$, which does not appear in X, would be mapped to (021). Afterward, an ordinal pattern probability distribution, $\boldsymbol{P} = \{p(\pi_i), i = 1, \ldots, D!\}$, can be obtained from the time series by computing the relative frequencies of the $D!$ possible permutations π_i. Continuing with the example: $p(\pi_1) = p(012) = 2/5$, $p(\pi_2) = p(021) = 0$, $p(\pi_3) = p(102) = 1/5$, $p(\pi_4) = p(120) = 1/5$, $p(\pi_5) = p(201) = 0$, and $p(\pi_6) = p(210) = 1/5$. PE is just the Shannon entropy estimated by using this ordinal pattern probability distribution, $S[\boldsymbol{P}] = -\sum_{i=1}^{D!} p(\pi_i)\log(p(\pi_i))$ ($0\log(0)$ is set to zero in accordance with its mathematical limit). Coming back to the example, $S[\boldsymbol{P}(X)] = -(2/5)\log(2/5) - 3(1/5)\log(1/5) = 1.3322$. PE quantifies the temporal structural diversity of a time series. Technically speaking, the ordinal pattern probability distribution $\boldsymbol{P}$ is obtained once we fix the order D and the lag τ. The PE estimation does not require the optimal reconstruction of the phase space that is necessary for estimating other quantifiers of chaotic signals. Consequently, D and τ are not usually selected following the methodologies often employed in a conventional phase space reconstruction (*e.g.*, the first zero of the autocorrelation function, the first minimum of the average mutual information, and the false nearest neighbor algorithm). Taking into account that there are $D!$ potential permutations for a D-dimensional vector, the condition $N \gg D!$, with N the length of the time series, must be satisfied in order to obtain a reliable estimation of $\boldsymbol{P}$ [33]. For practical purposes, BP suggest in their seminal paper to estimate the frequency of ordinal patterns with $3 \le D \le 7$ and $\tau = 1$ (consecutive points). However, it has been demonstrated that the analysis with lagged data points, *i.e.* $\tau \ge 2$, can be useful for reaching a better comprehension of the underlying dynamics [34–36]. Essentially, by changing the value of the lag τ, different time scales are being considered because this parameter physically corresponds to multiples of the sampling time of the signal under analysis. For further details about the BP methodology, we recommend [34,37,38]. It is common to normalize the PE, and therefore in this paper, a normalized PE is used and is given by

$$\mathcal{H}_S = S[\boldsymbol{P}]/S_{\max} = S[\boldsymbol{P}]/\log(D!) \quad (1)$$

with $S_{\max} = \log(D!)$ the value obtained from an equiprobable ordinal pattern probability distribution, *i.e.* $\boldsymbol{P} = \{p(\pi_i) = 1/D!, i = 1, \ldots, D!\}$.

2.2. Testing nonlinear dynamics in time series with surrogate methods

The method of surrogate data, introduced by Theiler et al. [10], represents a cornerstone in nonlinear time series analysis. Briefly, a statistic sensitive to nonlinearities is estimated for the original univariate time series $\{x_i\}_{i=1}^{N}$ and for an ensemble of M generated surrogate time series, $\{\hat{x}_i^j\}_{i=1}^{N}$ with $j = 1, \ldots, M$. Each surrogate (indexed with j) is a constrained realization of the original data that mimics its linear properties (autocorrelation function/power spectrum) while potential higher order correlations are randomized. When a statistically significant difference is found between