Submitted to the Statistical Science

Confidence Regions for Information-Theoretic Descriptors of Time Series

Eduarda T. C. Chagas and Marcelo Queiroz and Osvaldo A. Rosso and Heitor S. Ramos and Christopher G. S. Freitas and Leonardo V. Pereira and Alejandro C. Frery

School of Mathematics and Statistics

Abstract. The Bandt and Pompe methodology has been used with success in time series analysis. It consists in computing Information-Theoretic descriptors, from a histogram of ordinal patterns, which lie in a 2D manifold: the Entropy-Complexity plane. So far, the analysis of the dynamics underlying the time series has been performed using two reference points: those corresponding to a deterministic and to white noise time series. In this paper, we provide confidence regions for white noise models, and we used these regions as a method for analyzing PRNGs randomness. These regions allow testing whether a sequence belongs to one of these models.

Primary 37M10; secondary 68Q30.

Key words and phrases: Nonparametric time series analysis, ordinal patterns, Information Theory.

CONTENTS

Victoria University of Wellington PO Box 600 Wellington 6140, New Zealand (e-mail: alejandro.frery@vuw.ac.nz)

1. INTRODUCTION

Time Series carry valuable information about the system which produces the data. Their analysis is usually based on two approaches Cryer and Chan [2008]: in the (natural) time and transformed domains (for instance, frequency and wavelet).

In the context of time-domain analysis, a new methodology was proposed by Bandt and Pompe 2002. Its approach is non-parametric and based on descriptors of Information Theory. Through this, the time series is transformed into ordinal patterns, with which a histogram is formed. Using such patterns, the resulting distribution becomes less sensitive to outliers and, as it does not depend on any model, can be applied to a variety of situations.

The Bandt-Pompe methodology and its variants have been used successfully in the analysis of many types of dynamics, receiving so far more than 2500 citations, according to the Journal of Citation Records. We found works using such an approach in multiple areas of scientific knowledge such as, for example, the study of electroencephalography signals using wavelet decomposition Rosso et al. [2006], characterization of household appliances through their energy consumption Aquino et al. [2017], online signature classification and verification Rosso et al. [2016] .

Each time series is described by a point in the range of $\mathbb{R}^2$, the entropy complexity plane. Two points are well known in this plane: those of white noise and a completely deterministic sequence. Through these references, we can characterize time series according to the dynamics of its generating process. Based on this premise, studies with different applications managed to obtain relevant results from time series through information on the nature of the data provided by the $H \times C$ plane. Examples include the work of Carpi et al. 2011 analyzed the evolution of dynamic networks by trajectories in the $H \times C$ plane, Schieber et al. 2016 checked the effect of attacks to complex networks by the displacement of their points in the $H \times C$ plane, Aquino et al. 2015 described the behavior of vehicles as a function of the topology of cities, and Cabral et al. 2013 employ Information-Theoretic measures to describe the evolution of wireless sensors networks.

Many results in this regard have been provided through the analysis and study of chaotic components. We can cite as highlights: Martin et al. 2006 analyzed the logistic chaotic map and discuss the boundaries of the $H \times C$ plane. De Micco et al. 2009 studied chaotic components in pseudo-random number generators. Ravetti et al. 2014 tackled the often hard problem of distinguishing chaos and noise. Zunino et al. 2012 used a multi-scale approach to analyze the interplay between chaotic and stochastic dynamics.

With the knowledge of the expected variability of such points, according to the underlying dynamics, we can make hypothesis tests for a wide variety of models. Some preliminary results can already be found in the literature. Larrondo et al. 2006 showed that the Entropy-Complexity plane ($H \times C$) is a good indicator of the results of Diehard tests for pseudorandom number generators. De Micco et al. 2008 assessed ways of improving pseudorandom sequences by their representation in this plane. Bariviera et al. 2015 identified spurious interventions in the Libor market using the $H \times C$ plane representation.

This paper advances the state-of-the-art by providing confidence regions for

points in the $H \times C$ plane for a diversity of situations of interest. The input is a sequence of actual random observations generated by a physical procedure, and these regions are obtained by the application of PCA in space of Entropy-Complexity Plane. We emphasize that this study aims at providing confidence regions for sequences of the size that usually appear in the literature, not tests for randomness of PRNGs.

The paper is structured as follows: Section 1 introduces the elements of the study (the Bandt and Pompe methodology, the random deviates, and model). The confidence regions are presented in Section 3 with some analysis of this application, and the conclusions are discussed in Section 4.

2. METHODOLOGY

2.1 The Bandt-Pompe Methodology

Let $\mathcal{X} \equiv \{x_t\}_{t=1}^{T}$ be a real valued time series of length T, without ties. As stated by Bandt and Pompe 2002in their seminal work:

> "If the $\{x_t\}_{t=1}^{T}$ attain infinitely many values, it is common to replace them by a symbol sequence $\Pi \equiv \{\pi_j\}$ with finitely many symbols, and calculate source entropy from it".

Also, as stressed by these authors,

> "The corresponding symbol sequence must come naturally from the $\{x_t\}_{t=1}^{T}$ without former model assumptions".

Let $\mathbb{A}_D$ (with $D \geq 2$ and $D \in \mathbb{Z}$) be the symmetric group of order $D!$ formed by all possible permutation of order D, and the symbol component vector $\boldsymbol{\pi}^{(D)} = (\pi_1, \pi_2, \ldots, \pi_D)$ so every element $\boldsymbol{\pi}^{(D)}$ is unique ($\pi_j \neq \pi_k \ \forall \ j \neq k$). Consider for the time series $\mathcal{X} \equiv \{x_t\}_{t=1}^{T}$ its time delay embedding representation, with embedding dimension $D \geq 2$ ($D \in \mathbb{Z}$) and time delay $\tau \geq 1$ ($\tau \in \mathbb{Z}$, also called "embedding time"):

$$\mathbf{X}_t^{(D,\tau)} \ = \ (x_t, x_{t+\tau}, \ldots, x_{t+(D-1)\tau}) \ , \tag{1}$$

for $t = 1, 2, \ldots, N$ with $N = T-(D-1)\tau$. Then the vector $\mathbf{X}_t^{(D,\tau)}$ can be mapped to a symbol vector $\boldsymbol{\pi}^{(D)} \in \mathbb{A}_D$. This mapping should be defined in a way that preserves the desired relation between the elements $x_t \in \mathbf{X}_t^{(D,\tau)}$, and all $t \in T$ that share this pattern (also called motif) have to mapped to the same $\boldsymbol{\pi}^{(D)}$. The two most frequent ways to define the mapping $\mathbf{X}^{(D,\tau)} \mapsto \boldsymbol{\pi}^{(D)}$ are:

a) ordering the ranks of the $x_t \in \mathbf{X}^{(D,\tau)}$ in chronological order (*Rank Permutation*) or,
b) ordering the time indexes according to the ranks of $x_t \in \mathbf{X}^{(D,\tau)}$ (*Chronological Index Permutation*);

see details in Traversaro et al. 272018Traversaro et al.Traversaro, Redelico, Risk, Frery, and Rosso272018Traversaro et al.Traversaro, Redelico, Risk, Frery, and Rosso(272018Traversaro et al.Traversaro, Redelico, Risk, Frery, and Rosso). Without loss of generality, in the following we will only use the latter.

Consider, for instance, the time series $\mathcal{X} = (1.8, 1.2, 3.2, 4.8, 4.2, 4.5, 2.3, 3.7, 1.2, .5)$ depicted in Fig. 1. Assume we are using patterns of length $D = 5$ with unitary time lag $\tau = 1$. The code associated to $\mathbf{X}_3^{(5,1)} = (x_3, \ldots, x_7) = (3.2, 4.8, 4.2, 4.5, 2.3)$,

shown in black, is formed by the indexes in $\boldsymbol{\pi}^{(5)} = (1, 2, 3, 4, 5)$ which sort the elements of $\mathbf{X}_3^{(5,1)}$ in increasing order: 51342. With this, $\widetilde{\pi}^{(5)} = 51342$, and we increase the counting related to this motif in the histogram of all possible patterns of size $D = 5$.

The dash-dot line in Fig. 1 illustrates $\mathbf{X}_1^{(5,2)}$, i.e. the sequence of length $D = 5$ starting at x_1 with lag $\tau = 2$. In this case, $\mathbf{X}_1^{(5,2)} = (1.8, 3.2, 4.2, 2.3, 1.2)$, and the corresponding motif is $\widetilde{\pi}^{(5)} = 51423$.

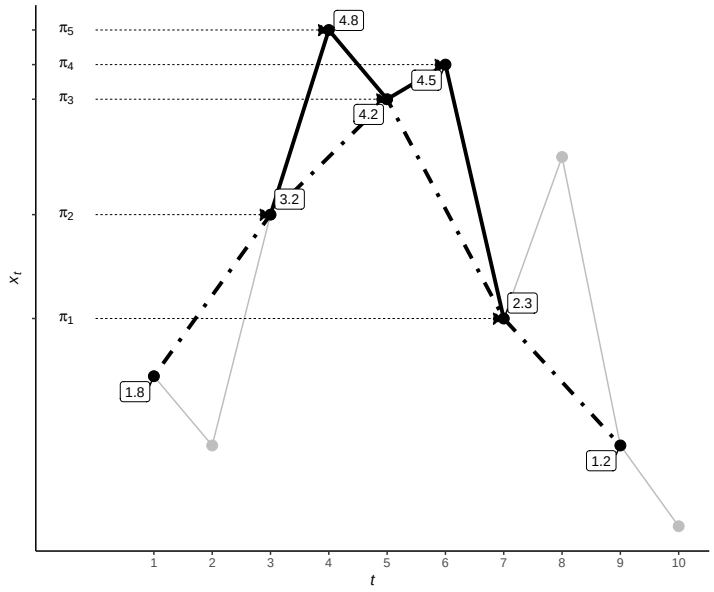


FIG 1. *Illustration of the Bandt and Pompe coding*

Once all symbols have been computed, one obtains the histogram of proportions $\boldsymbol{h} = (h(j))_{1 \leq j \leq D!}$. This is an estimate of the (unknown, in general) probability distribution function of these patterns. The next step into the characterization of the time series is computing descriptors from this histogram.

The first descriptor is a measure of the disorder of the system. The most frequently used feature for this is the Normalized Shannon entropy, defined as

$$H(\boldsymbol{h}) = -\frac{1}{\log D!} \sum_{j=1}^{D!} h(j) \log h(j), \tag{2}$$

with the convention that terms in the summation for which $h(j) = 0$ are null. This quantity is bounded in the unit interval, and is zero when $h(j) = 1$ for some j (and, thus, all other bins are zero), and one when $h(j) = 1/D!$ for every j (the uniform probability distribution function).

Although very expressive, the Normalized Shannon Entropy is not able to describe all possible underlying dynamics. In particular, for intermediate values of H, there is a wide variety of situations worth characterizing. To this aim, López-Ruiz et al. 1995 proposed using Q, the disequilibrium, a measure of how far $\boldsymbol{h}$ is from an equilibrium or noninformative distribution. They employed the Euclidean distance between $\boldsymbol{h}$ and the uniform probability distribution function.

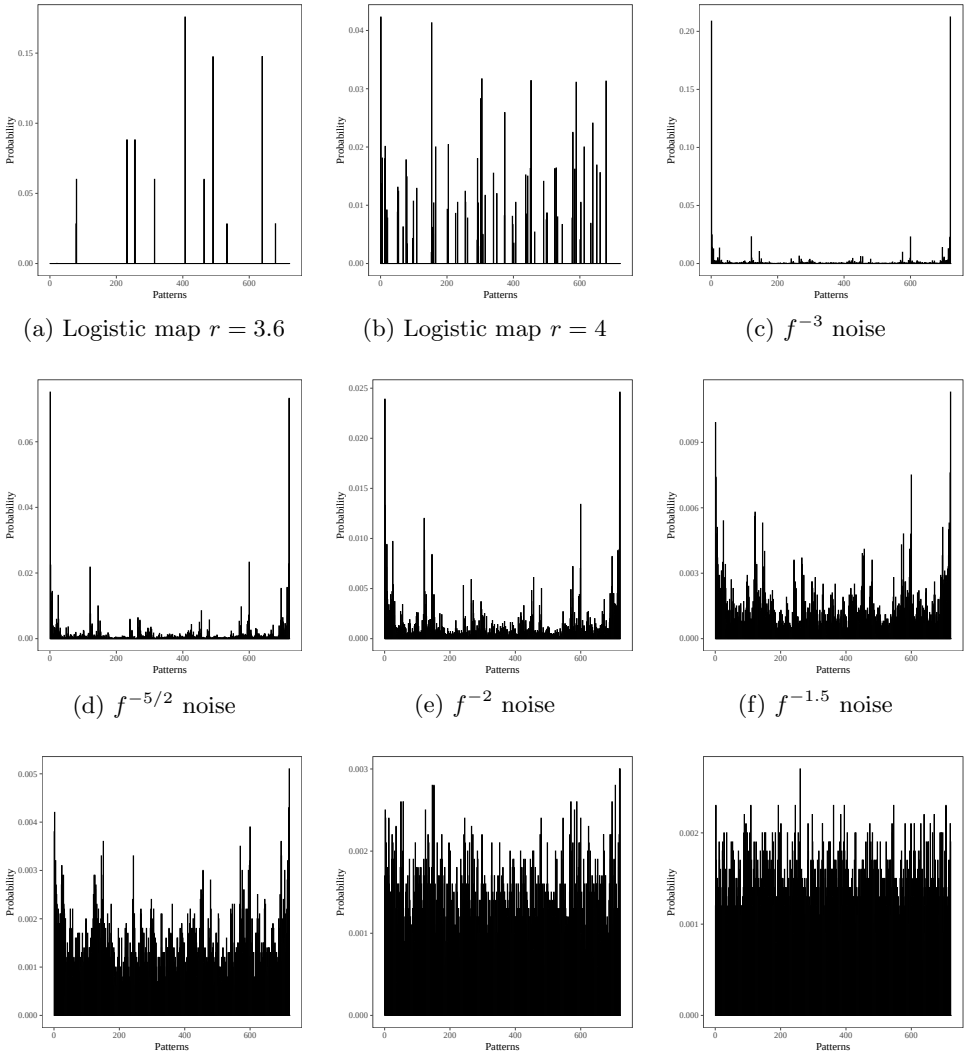
With this, they proposed $C = HQ$ as a measure of the Statistical Complexity of the underlying dynamics. A time series can then be mapped into a point in the $H \times C$ plane.

2.2 The Entropy-Complexity Plane

We illustrate the use of the Entropy-Complexity ($H \times C$) with the following time series:

- Colored k-noise: white ($k = 0$), $k = -1/2$, pink ($k = 1$), $k = 3/2$, red ($k = 2$), $= 5/2$, and $k = 3$;
- Chaotic logistic series $x_t = rx_{t-1}(1 - x_{t-1})$, with $r = 3.6$ and 4;
- Deterministic series: monotonic increasing ($\log(x_t+0.1)$, $x_t = \{1, 2, \ldots, 10^4\}$) and periodic ($\sin(2x_t)\cos(2x_t)$, with $0 \leq x_t \leq 2\pi$ over ten thousand equally spaced points).

In all cases, we used $D = 6$ and $\tau = 1$. Fig. 2 shows nine of the histograms produced by these series using the Mersenne-Twister pseudorandom number generator; we omitted those corresponding to the deterministic series, as they produce one and two nonzero bins.



(a) Logistic map $r = 3.6$ (b) Logistic map $r = 4$ (c) f^{-3} noise

(d) $f^{-5/2}$ noise (e) f^{-2} noise (f) $f^{-1.5}$ noise

(g) f^{-1} noise (h) $f^{-1/2}$ noise (i) White noise

FIG 2. *Patterns histograms of selected time series, with* $D = 6$ *and* $\tau = 1$

Fig. 3 shows the $H \times C$ plane with the bounds for $D = 6$, the time series and the points they were mapped onto. The points due to f^{-k} noises appear joined by dotted segments. It is noticeable that deterministic patterns have more complexity than random ones. Also, points related to f^{-k} noises tend to clutter for $k < 1$.

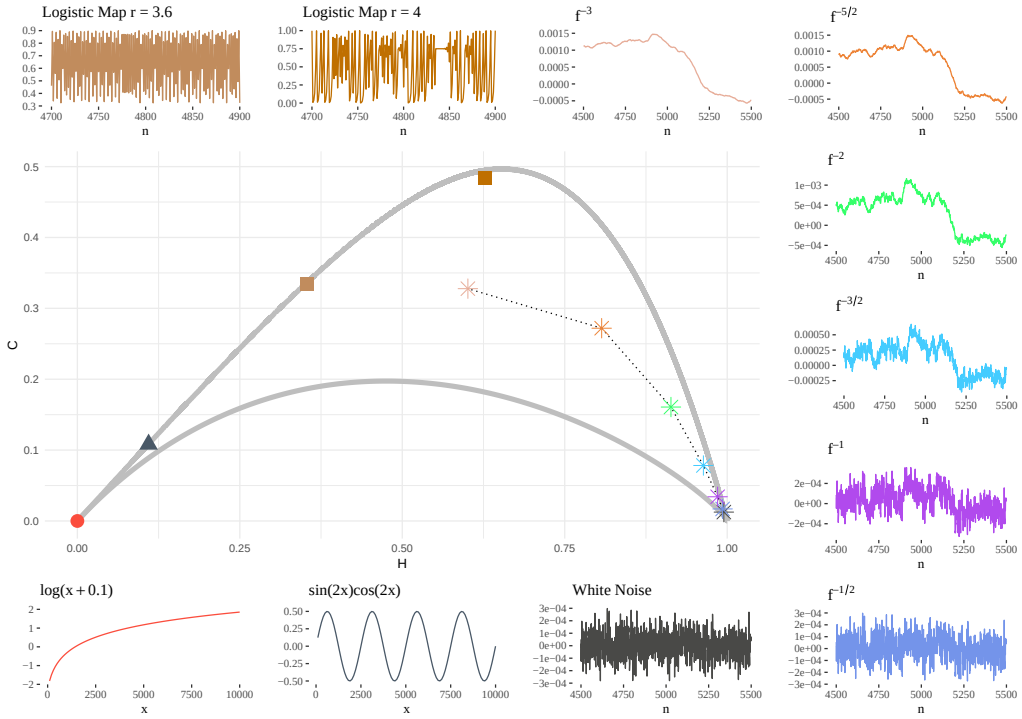


FIG 3. *Eleven systems and their points in the $H \times C$ plane*

Fig. 4 shows the rightmost lower corner of the $H \times C$ plane, emphasizing the location of the white ($k = 0$), $k = -1/2$, and pink ($k = 1$) noises.

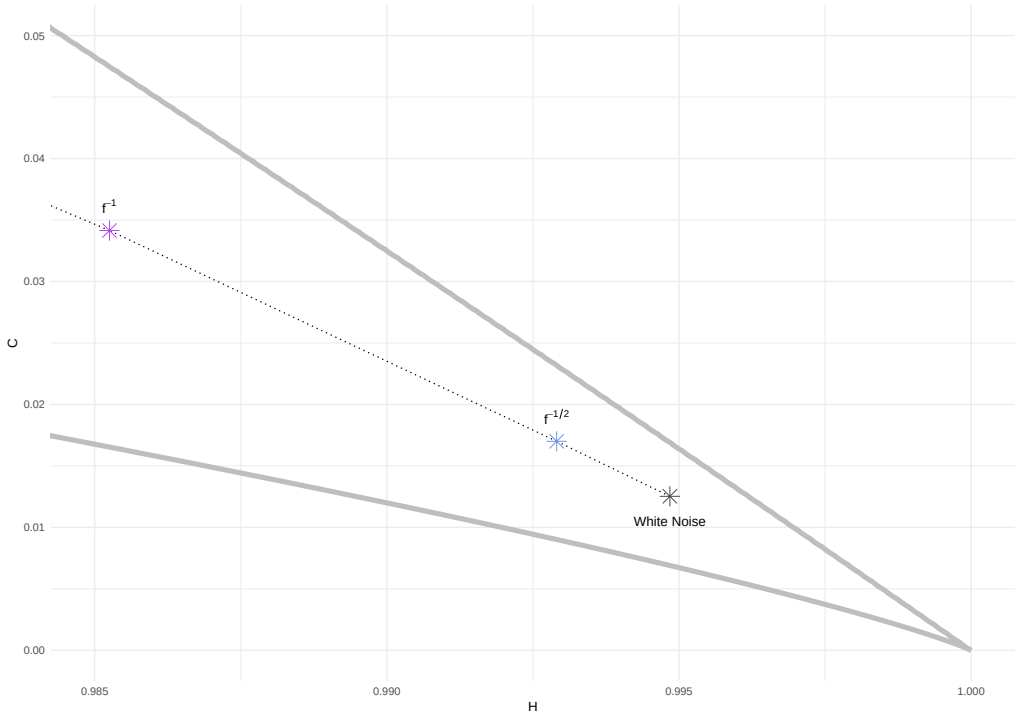


FIG 4. *White Noise, $f^{-1/2}$ and f^{-1} noise points*

The focus of our study is the empirical distribution of points produced by White and selected f^{-k} noises, with this, providing confidence regions in the $H \times C$ plane. In particular, we are interested in assessing pure randomness with this technique.

2.3 The $H \times C$ plane in the literature

Table 1 presents references in which the authors analyze time series of size N and embedding dimension D. The authors also attest whether the time series is white noise or not, according to its (h, c) feature.

The following widely used non-chaotic algorithmic generators are analyzed:

- RNG available in Intel fortran compiler (FOR)
- RNG available in Borland C++ compiler (CCC)
- Matlab RAND function (MAT)
- Mother RNG, available in Marsaglia website Marsaglia [1994] (MOT)
- Multiple with carry RNG (MWC) Marsaglia [1994]
- Combo RNG (COM) Marsaglia [1994]
- Lehmer RNG (LEH) Payne et al. [1969]
- Fractional Gaussian noise with $\alpha = 0$ (fGn)
- Fractional Brownian motion with $\alpha = 1.2$ (fBm)
- f^{-k} noise with $k = 0$
- Linear Congruential Generator (LCG) Knuth [1997]

TABLE 1
References which attest whether a time series of size N is white noise according to its (h, c) feature, and the empirical p-value of such hypothesis according to our results

Reference	PRNG	N	D	H	C	Is white noise?	p-value
Larrondo et al. (2002)	MOT	NA	6	$\cong 0.9969$	$\cong 0$	no	NA
	FOR	NA	6	$\cong 0.997$	$\cong 0$	no	NA
	CCC	NA	6	$\cong 0.997$	$\cong 0$	no	NA
	MAT	NA	6	$\cong 0.997$	$\cong 0$	no	NA
González et al. (2005)	MWC	65536	NA	$\cong 1$	0.3	yes	NA
	MOT	65536	NA	$\cong 1$	0.3	yes	NA
	COM	65536	NA	$\cong 1$	0.05	yes	NA
Larrondo et al. (2006)	LEH	5×10^6	5	NA	10^{-4}	yes	NA
	MOT	5×10^6	5	NA	10^{-4}	yes	NA
	MWC	5×10^6	5	NA	10^{-4}	yes	NA
Olivares et al. (2012)	fGn	2×10^{15}	6	$\cong 0.998$	NA	yes	NA
	fBm	2×10^{15}	6	$\cong 0.993$	NA	yes	NA
	f^{-k}	2×10^{15}	6	$\cong 0.997$	NA	yes	NA
Rosso et al. (2013)	LCG	1×10^7	6	0.997871	0.005101	no	NA
Xiong et al. (2020)	fGn	2×10^{17}	6	$\cong 1$	$\cong 0$	yes	NA
	f^{-k}	2×10^{17}	6	$\cong 1$	$\cong 0$	yes	NA

2.4 True Random Numbers

Gerador - Marcelo e Heitor

Random numbers are used in many fields, from gambling to criptography, aiming to guarantee a secure, realistic or unpredictable behavihor. Pseudo randomic results can be achieved by software in a deterministic way. But, some applications need actual random numbers (despite the somewhat elusive nature of actual randomness). Randomness can be observed in unpredictable real world phenomena like cathodic radiation or atmospheric noise. In this study we used two sources of real random numbers. The first is based on vacuum states to generate random quantum numbers described by Gabriel et al. (2010), the second one is based on atmospheric noise captured by a cheap radio receiver presented at `www.random.org`.

2.5 Empirical Confidence Regions

As we do not know the joint probability distribution of the pair (h, c) for a sequence of random variables collectively independent and identically distributed according to a uniform law, studies involving classical bi-variate analysis, linear regression, and generalized linear models become unfeasible at this first moment. Therefore, for the construction of our proposal, we adopted a non-parametric approach, making an empirical analysis of data obtained from physical sources and using them as our reference in the search for confidence regions.

The set of all feasible pairs in the $H \times C$ plane is found in a compact subset of $\mathbb{R}^2$, which has limits with explicit expressions for the boundaries of this closed manifold, dependent only on the dimension of the probability space considered, that is $D!$ in the traditional Bandt-Pompe method Martín et al. [2006]. Due to such quotas, some limitations are generated, such as the absence of a representative distance metric and the difficulty of proposing confidence regions. In view of this, it is necessary to apply an orthogonal projection in the data for a

new two-dimensional coordinate system to solve these restrictions. A classic proposal in these categories of problems is the principal component analysis (PCA) algorithm Wold et al. [1987].

Let $\mathcal{P} = \{p_n\}_{n=1}^{N}$ be a set of N observations in the $H \times C$ plane, where $p_n = (h_n, c_n)$ corresponds to a time series. We apply a principal component analysis to these points and obtain $\{(u_n, v_n)\}_{n=1}^{N}$. This transformation yields uncorrelated observations in the $\Omega_1 \times \Omega_2$ plane.

We will find the empirical confidence regions on the $\Omega_1 \times \Omega_2$ plane, and then they will be mapped back to the $H \times C$ plane. For simplicity, and without loss of generality, assume N odd; we find a parallelepiped that contains $100\,(1-\alpha)\%$ of the points in the $\Omega_1 \times \Omega_2$ as

1. Find the ranks that sort the values of the first principal component $\boldsymbol{u} = (u_1, u_2, \ldots, u_N)$ in ascending order: $\boldsymbol{r} = (r_1, r_2, \ldots, r_N)$, i.e., u_{r_1} is the minimum value, and u_{r_N} is the maximum value.
2. Find point (u, v) whose first principal component is the median: $(u_{r_{(N+1)/2}}, \cdot)$.
3. Find the point (u, v) whose first principal component is the quantile $\alpha/2$: $(u_{r_{[N\alpha/2]}}, \cdot)$.
4. Find the point (u, v) whose first principal component is the quantile $1-\alpha/2$: $(u_{r_{[N(1-\alpha/2)]}}, \cdot)$.
5. The values $u_{r_{[N\alpha/2]}}$ and $u_{r_{[N(1-\alpha/2)]}}$ are the rightmost and leftmost bounds of the box, respectively.
6. The top and bottom bounds of the box are the minimum and maximum values of the second principal component of the points whose first principal component is at least $u_{r_{[N\alpha/2]}}$ and at most $u_{r_{[N(1-\alpha/2)]}}$; denote these values $v_{\min}$ and $v_{\max}$, respectively.
7. The corners of the box are $(u_{r_{[N\alpha/2]}}, v_{\min})$, $(u_{r_{[N\alpha/2]}}, v_{\max})$, $(u_{r_{[N(1-\alpha/2)]}}, v_{\min})$ and $(u_{r_{[N(1-\alpha/2)]}}, v_{\max})$.
8. Apply the inverse PCA transform to these corners obtaining (h_{v_1}, c_{v_1}), (h_{v_2}, h_{v_2}), (h_{v_3}, c_{v_3}) and (h_{v_4}, c_{v_4}).

These steps are also depicted in Fig. 5.

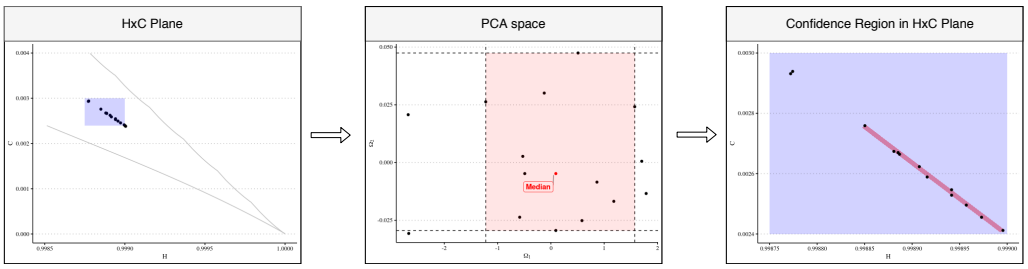


FIG 5. *Outline of the methodology used for the construction of confidence regions.*

3. RESULTS

In this section, we describe the descriptive analysis of representative points in relation to the confidence regions, as well as revisiting hypotheses of white noise present in the literature. Empirical experiments on the behavior of an emblematic sample when injecting correlation structures are also presented.

TABLE 2. *Empirical confidence regions for time series of length $T = 5 \times 10^4$ in the $H \times C$ plane obtained by the proposed methodology*

Confidence Regions	D	(h_{v_1}, c_{v_1})	(h_{v_2}, h_{v_2})	(h_{v_3}, c_{v_3})	(h_{v_4}, c_{v_4})
90 %	3	$(0.9999489, 5.06 \times 10^{-5})$	$(0.9999487, 5.04 \times 10^{-5})$	$(0.9999998, 4 \times 10^{-7})$	$(0.9999996, 2 \times 10^{-7})$
95 %	3	$(0.9999384, 6.11 \times 10^{-5})$	$(0.9999382, 6.09 \times 10^{-5})$	$(0.9999994, 9 \times 10^{-7})$	$(0.9999991, 7 \times 10^{-7})$
99 %	3	$(0.9999079, 9.11 \times 10^{-5})$	$(0.9999077, 9.09 \times 10^{-5})$	$(0.9999982, 2 \times 10^{-6})$	$(0.999998, 1.8 \times 10^{-6})$
99.9 %	3	$(0.9998625, 0.0001361)$	$(0.9998622, 0.0001358)$	$(0.9999973, 3 \times 10^{-6})$	$(0.999997, 2.7 \times 10^{-6})$
90 %	4	$(0.9999684, 3.98 \times 10^{-5})$	$(0.9999696, 4.13 \times 10^{-5})$	$(0.9998075, 0.0002508)$	$(0.9998087, 0.0002524)$
95 %	4	$(0.9999725, 3.44 \times 10^{-5})$	$(0.9999737, 3.6 \times 10^{-5})$	$(0.9998506, 0.0001942)$	$(0.9998518, 0.0001958)$
99 %	4	$(0.9999783, 2.68 \times 10^{-5})$	$(0.9999795, 2.83 \times 10^{-5})$	$(0.9998756, 0.0001615)$	$(0.9998768, 0.000163)$
99.9 %	4	$(0.9999833, 2.02 \times 10^{-5})$	$(0.9999845, 2.18 \times 10^{-5})$	$(0.9998889, 0.000144)$	$(0.9998901, 0.0001456)$
90 %	5	$(0.9998172, 0.0003232)$	$(0.9998194, 0.0003273)$	$(0.9994812, 0.0009246)$	$(0.9994834, 0.0009286)$
95 %	5	$(0.9998259, 0.0003075)$	$(0.9998282, 0.0003116)$	$(0.9996371, 0.0006455)$	$(0.9996394, 0.0006495)$
99 %	5	$(0.9998428, 0.0002774)$	$(0.999845, 0.0002814)$	$(0.9996703, 0.0005862)$	$(0.9996725, 0.0005901)$
99.9 %	5	$(0.9998573, 0.0002517)$	$(0.9998593, 0.0002553)$	$(0.9996884, 0.000554)$	$(0.9996904, 0.0005576)$
90 %	6	$(0.9990169, 0.002336)$	$(0.9990249, 0.0023554)$	$(0.9978983, 0.0050219)$	$(0.9979064, 0.0050413)$
95 %	6	$(0.9990368, 0.002288)$	$(0.9990449, 0.0023074)$	$(0.998714, 0.0030633)$	$(0.998722, 0.0030827)$
99 %	6	$(0.9990736, 0.0021997)$	$(0.9990817, 0.0022191)$	$(0.998765, 0.0029407)$	$(0.9987731, 0.0029601)$
99.9 %	6	$(0.9991069, 0.0021197)$	$(0.999115, 0.0021392)$	$(0.9987884, 0.0028845)$	$(0.9987965, 0.0029039)$

TABLE 3. *Empirical confidence regions for time series of length $T = 1 \times 10^3$ in the $H \times C$ plane obtained by the proposed methodology*

Confidence Region	D	(h_{v_1}, c_{v_1})	(h_{v_2}, h_{v_2})	(h_{v_3}, c_{v_3})	(h_{v_4}, c_{v_4})
90 %	3	(0.9973334, 0.0025601)	(0.9974047, 0.0026304)	(0.9999497, 0)	$(1, 5.17 \times 10^{-5})$
95 %	3	(0.9967311, 0.0031343)	(0.9968219, 0.0032238)	(0.9999398, 0)	$(1, 6.12 \times 10^{-5})$
99 %	3	(0.9953009, 0.0045054)	(0.9954349, 0.0046375)	(0.9999203, 0)	$(1, 8.45 \times 10^{-5})$
99.9 %	3	(0.9931825, 0.0065387)	(0.9933704, 0.006724)	(0.9998925, 0)	(1, 0.0001104)
90 %	4	(0.994364, 0.0081246)	(0.9939234, 0.0075452)	(0.9994791, 0.0013982)	(0.9990385, 0.0008188)
95 %	4	(0.9937138, 0.0089796)	(0.9932534, 0.0083741)	(0.9991609, 0.0018166)	(0.9987005, 0.0012111)
99 %	4	(0.9922575, 0.0108947)	(0.9917308, 0.0102022)	(0.9987924, 0.0023012)	(0.9982658, 0.0016087)
99.9 %	4	(0.9902578, 0.0135243)	(0.9897312, 0.0128318)	(0.9985727, 0.0025901)	(0.9980461, 0.0018976)
90 %	5	(0.9811818, 0.0321294)	(0.9827429, 0.0350291)	(0.9919707, 0.0120896)	(0.9935319, 0.0149893)
95 %	5	(0.9801289, 0.0340045)	(0.9817117, 0.0369446)	(0.9909376, 0.0139279)	(0.9925204, 0.016868)
99 %	5	(0.977917, 0.0377295)	(0.9796031, 0.0408613)	(0.9898161, 0.0156277)	(0.9915021, 0.0187595)
99.9 %	5	(0.9753326, 0.0425299)	(0.9770187, 0.0456617)	(0.9892599, 0.0166608)	(0.9909459, 0.0197926)
90 %	6	(0.9121895, 0.2201993)	(0.9146048, 0.2260776)	(0.9443868, 0.1418373)	(0.9468021, 0.1477156)
95 %	6	(0.9105951, 0.2239294)	(0.9130413, 0.2298829)	(0.9419202, 0.1476904)	(0.9443663, 0.1536439)
99 %	6	(0.9105951, 0.2239294)	(0.9130413, 0.2298829)	(0.9396577, 0.1531967)	(0.9421039, 0.1591502)
99.9 %	6	(0.9077672, 0.2305874)	(0.9102595, 0.2366531)	(0.9383611, 0.1561279)	(0.9408534, 0.1621937)

We employed the following factors in this study:

- Time series length $T \in \mathcal{T} = \{10^3, 5 \times 10^4\}$,
- Embedding dimension $D \in \mathcal{D} = \{3, 4, 5, 6\}$,and
- Exponent of the f^{-k} noise $k \in \mathcal{K} = \{0, 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.5, 1, 1.5, 2, 2.5, 3\}$.
- Time delay $\tau = \{1, 10, 30, 50\}$.

The data generation and analyses were performed using the `R` platform R Core Team [2017] v. 3.6.3. We used the `ggplot2` library Wickham [2009] for generating the plots.

3.1 Descriptive analysis of representative data

For the construction of the confidence regions presented in this work we use:

- An set of 104596 points in the $H \times C$ plane, referring to sequences of length $T = 1000$, for each combination of the factors $\mathcal{T} \times \mathcal{D} \times \tau$, and
- An another set of 2093 points in the $H \times C$ plane, referring to sequences of length $T = 50000$, for each combination of factors $\mathcal{T} \times \mathcal{D} \times \tau$.

Since the results involving the variation of the time delay parameter did not show a significant difference in repeated experiments, we don't consider it as a determinating factor during the execution of the algorithm. On the other hand, we consider two determining variables during the generation of such sub-spaces: the embedding dimension and the length of the sequence.

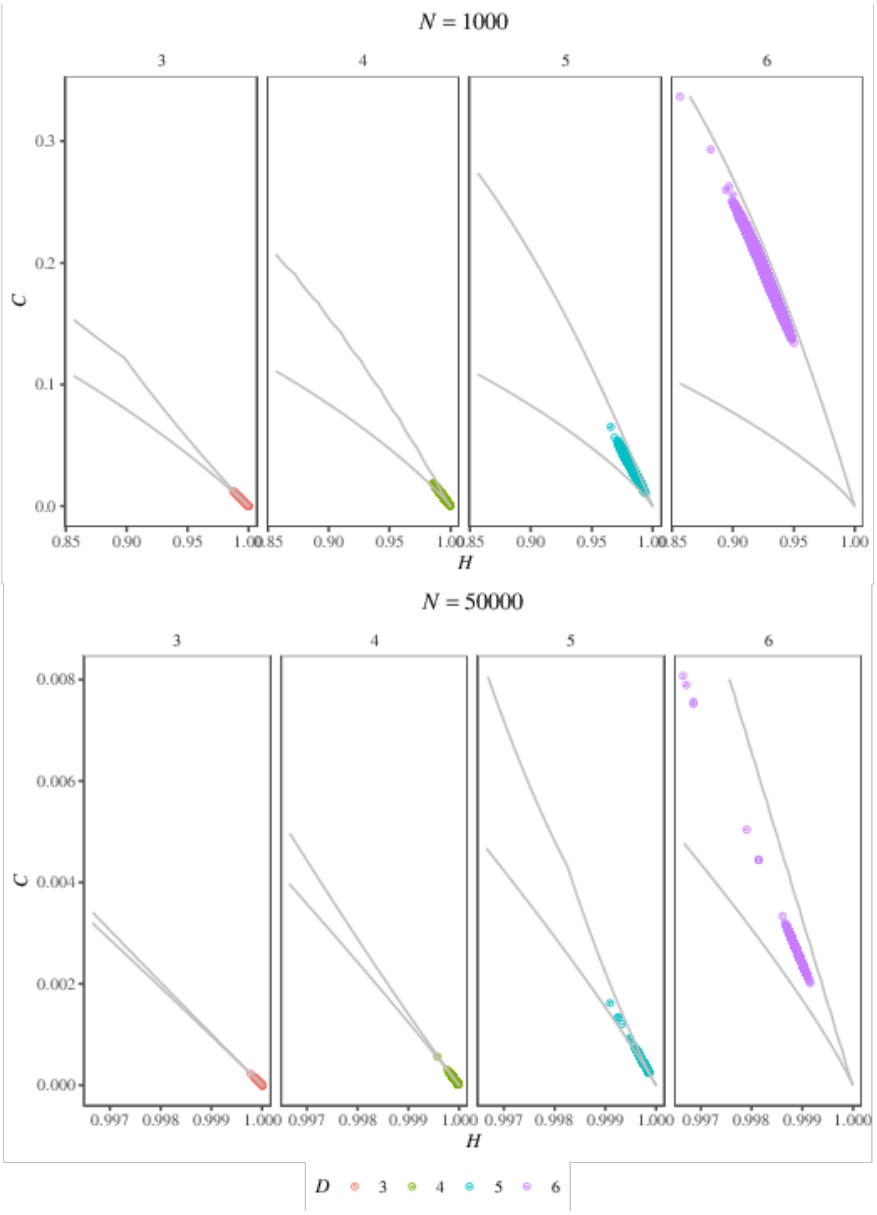


FIG 6. *White noise samples considered during the construction of the proposed confidence regions.*

In Fig. 7 we show the results obtained for $T = 50000$ in the scenarios of $D = 3$ and $D = 6$ in the new space defined by the PCA, together with the quantiles of 90%, 95%, 99% and 99.9%. We also show the projection of the $H \times C$ plane limits in this new representation space, as well as identifying the median of each data set, the latter being represented as the red dots present in the graphs.

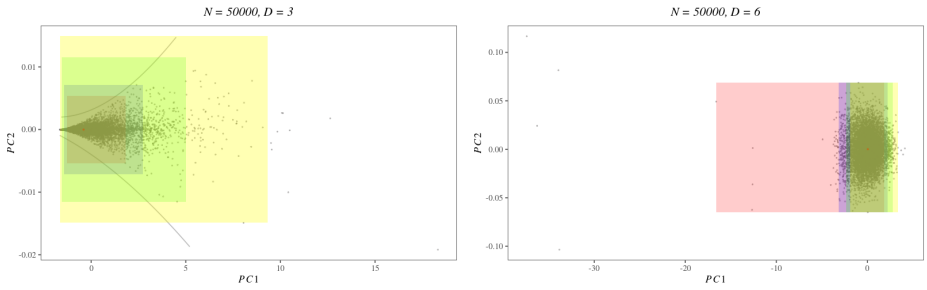


FIG 7.

As we can see in Fig. 8 in the new representation space produced by the PCA, the data are not evenly distributed among the axes of the first main component, maintaining the character presented in the $H \times C$ plane, since such points tend to be concentrated close to the point $(1, 0)$.

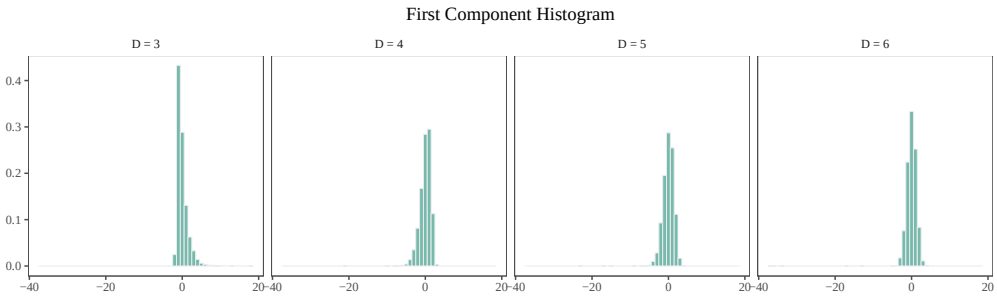


FIG 8. *Histogram of the PCA first component*

3.2 Testing White Noise in the confidence regions

To test the efficiency of the confidence region calculated, we tested its applicability on a set of truly random data generated physically not used by the algorithm during its construction. The results can be seen in Fig. 9.

For small series, $N = 1000$, and $D = 3$ we managed to maintain exactly 99% of the data in the confidence region of the same value, and as the dimension increased reach 96% of the points. On the other hand, there was a very large loss of points located in the region with 95% confidence as the dimension increased. A reasonable explanation for this event is given in the choice of the parameter itself. It is known by definition that $D! << N$, which does not happen for such a sample size, thus presenting many missing patterns that lead to an unrepresentative probability distribution. For larger series, $N = 50000$, although we observed a small drop in the percentage of data present in the region with 99% confidence, there was a significant increase in points in the region with 95% confidence, showing between 90% and 88% of the points when we vary the embedding dimension.

TABLE 4
Results found it for samples of true random numbers

Algorithm	N	D	95 %	99 %
True-Random	1000	3	0.98	1
		4	0.98	0.96
		5	1	0.94
		6	0.97	0.87
True-Random	50000	3	0.97	0.96
		4	0.94	0.95
		5	0.97	0.96
		6	0.98	0.99

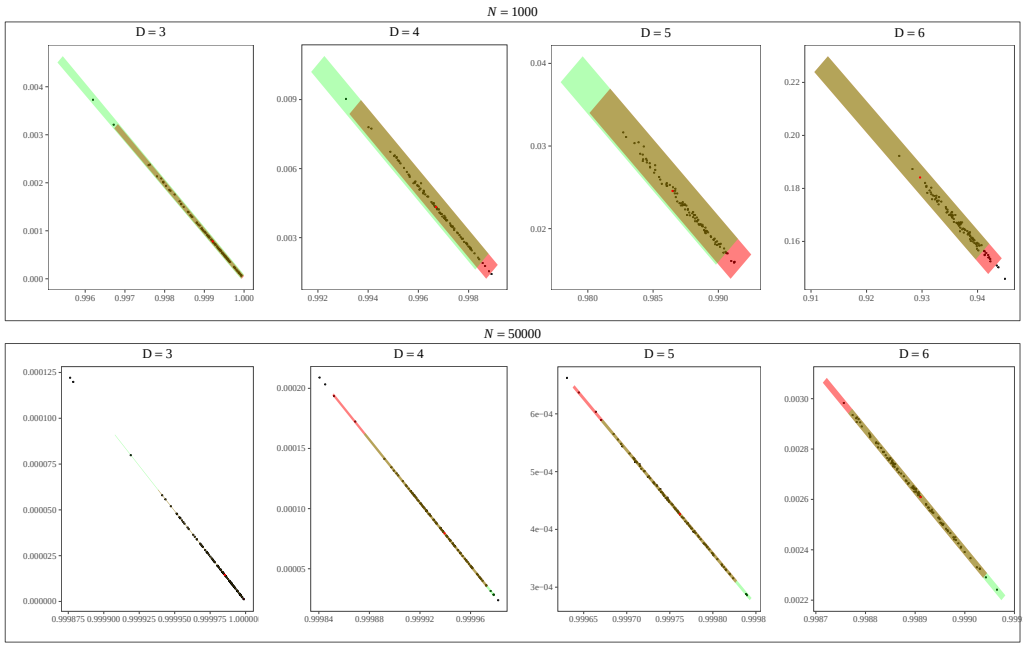


FIG 9. *White Noise in HC Plane*

3.3 Analyzing of the Empirical Confidence Regions Injecting Correlation

Fig. 3 shows the behavior of random time series with different levels of correlation (by means of the f^{-k} model) in the (h, c) plane. Knowing that such plane can discriminate between different system dynamics, several studies in the literature have used this approach in the investigation of methods of identification and characterization of randomness. Although this same strategy can be used to characterize different levels of correlation structures, in our case, we want to analyze the impact of injecting such dynamics into noise under the aspect of confidence regions.

For carrying out the experiment, we used an "emblematic" time series as a basis. This series consists of the sample corresponding to the median of the (h, c) points used to build the confidence regions, thus expressing a representative sample of the dataset. Fig. 10a. illustrates, respectively, the effect of a white noise time series when adding correlation structures related to the f^{-k} series for

$k \in \{0, 1, 2, 3\}$. As we can observe in the plane as the correlation between the observations increases, that is, $k > 0$, the randomness decreases and the entropy presented decreases, informing the loss of its stochastic characteristic.

Fig. 10b. illustrates the degree of limit correlation structure that can be added in white noise to eliminate it from the regions of confidence, where the red dot represents the original "emblematic" series. When we have $k = 0$, the features of the sequence have a small variation, corroborating the premise that series of noise $f^{-k} = 0$ have a minimum correlation measure, not significantly changing the dynamics of the system.

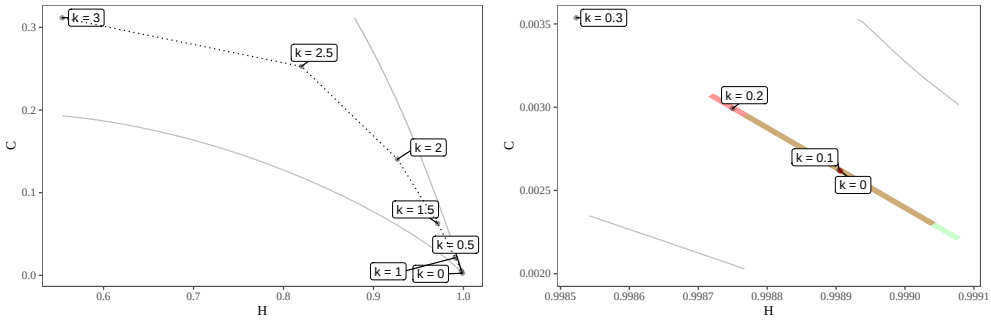


FIG 10. *Correlation Structure Analysis*

3.4 Revisiting the White Noise Hypothesis in the Literature

TABLE 5
Points inside the confidence regions

Algorithm	D	95 %	99 %	Algorithm	D	95 %	99 %
Wichmann-Hill	3	0.97	0.96	Xoshiro256+ (FOR)	3	0.98	0.98
Wichmann-Hill	4	0.9675	0.9675	Xoshiro256+ (FOR)	4	0.9675	0.965
Wichmann-Hill	5	0.9725	0.965	Xoshiro256+ (FOR)	5	0.9525	0.96
Wichmann-Hill	6	0.97	0.97	Xoshiro256+ (FOR)	6	0.9775	0.975
Super-Duper	3	0.9675	0.975	Mersenne-Twister	3	0.9725	0.97
Super-Duper	4	0.9775	0.9725	Mersenne-Twister	4	0.98	0.97
Super-Duper	5	0.9675	0.985	Mersenne-Twister	5	0.9725	0.9775
Super-Duper	6	0.9825	0.965	Mersenne-Twister	6	0.975	0.9725
Knuth-TAOCP-2002	3	0.9775	0.9625	Randu	3	0.2275	0.42
Knuth-TAOCP-2002	4	0.965	0.98	Randu	4	0	0
Knuth-TAOCP-2002	5	0.9625	0.9825	Randu	5	0	0
Knuth-TAOCP-2002	6	0.9725	0.96	Randu	6	0	0
Knuth-TAOCP	3	0.975	0.9775	LCG	3	0.07	0.09
Knuth-TAOCP	4	0.965	0.9725	LCG	4	0	0
Knuth-TAOCP	5	0.9875	0.985	LCG	5	0	0
Knuth-TAOCP	6	0.9725	0.975	LCG	6	0	0
LEcuyer-CMRG	3	0.98	0.98	LEH	3	0.9825	0.98
LEcuyer-CMRG	4	0.9675	0.9725	LEH	4	0.9675	0.9775
LEcuyer-CMRG	5	0.965	0.9725	LEH	5	0.9725	0.9775
LEcuyer-CMRG	6	0.9625	0.97	LEH	6	0.9775	0.9825
pcg64	3	0.97	0.9675	fBm	3	0	0
pcg64	4	0.97	0.97	fBm	4	0	0
pcg64	5	0.9775	0.9625	fBm	5	0	0
pcg64	6	0.9575	0.9575	fBm	6	0	0
Threefry	3	0.9625	0.97	fGn	3	0.9722	0.9874
Threefry	4	0.975	0.97	fGn	4	0.9798	0.9773
Threefry	5	0.96	0.9625	fGn	5	0.9646	0.9773
Threefry	6	0.965	0.955	fGn	6	0.9722	0.9545
Xoroshiro128+	3	0.975	0.9825	f^{-k}	3	0.9525	0.96
Xoroshiro128+	4	0.995	0.9825	f^{-k}	4	0.975	0.9725
Xoroshiro128+	5	0.98	0.9725	f^{-k}	5	0.98	0.9675
Xoroshiro128+	6	0.9725	0.955	f^{-k}	6	0.9775	0.975
MWC	3	1	0.75	COM	3	0.25	0.5
MWC	4	0.25	0.25	COM	4	0	0
MWC	5	1	1	COM	5	0	0
MWC	6	0	0	COM	6	0	0
MOT	3	1	1	CCC	3	0	0
MOT	4	1	1	CCC	4	0	0
MOT	5	1	1	CCC	5	0	0
MOT	6	0.75	1	CCC	6	0	0

4. CONCLUSIONS

5. SOURCE CODE AVAILABILITY

The text, source code, and data used in this study are available at the *Confidence-Regions* repository https://github.com/EduardaChagas/ConfidenceRegions.

6. ACKNOWLEDGEMENTS

This work was partially funded by the Coordination for the Improvement of Higher Education Personnel (CAPES) and National Council for Scientific and Technological Development (CNPq).

REFERENCES

A. L. L. Aquino, T. S. G. Cavalcante, E. S. Almeida, A. C. Frery, and O. A. Rosso. Characterization of vehicle behavior with information theory. *The European Physical Journal B: Condensed Matter and Complex Systems*, 88(10):257–269, 2015. ISSN 1434-6036. . URL http://dx.doi.org/10.1140/epjb/e2015-60384-x.

A. L. L. Aquino, H. S. Ramos, A. C. Frery, L. P. Viana, T. S. G. Cavalcante, and O. A. Rosso. Characterization of electric load with information theory quantifiers. *Physica A*, 465:277–284, 2017. .

C. Bandt and B. Pompe. Permutation entropy: A natural complexity measure for time series. *Physical Review Letters*, 88:174102–1–174102–4, 2002. .

A. F. Bariviera, M. B. Guercio, L. B. Martinez, and O. A. Rosso. The (in)visible hand in the Libor market: an information theory approach. *The European Physical Journal B*, 88(8), 2015. .

R. S. Cabral, A. L. L. Aquino, A. C. Frery, O. A. Rosso, and J. A. Ramírez. Structural changes in data communication in wireless sensor networks. *Central European Journal of Physics*, 11 (12):1645–1652, 2013. .

L. C. Carpi, O. A. Rosso, P. M. Saco, and M. Gómez Ravetti. Analyzing complex networks evolution through Information Theory quantifiers. *Physics Letters A*, 375:801–804, 2011. .

J. D. Cryer and K.-S. Chan. *Time Series Analysis With Applications in R*. Springer, 2 edition, 2008.

L. De Micco, C. M. González, H. A. Larrondo, M. T. Martin, A. Plastino, and O. A. Rosso. Randomizing nonlinear maps via symbolic dynamics. *Physica A: Statistical Mechanics and its Applications*, 387(14):3373–3383, 2008. ISSN 0378-4371. . URL http://dx.doi.org/10.1016/j.physa.2008.02.037.

L. De Micco, H. A. Larrondo, A. Plastino, and O. A. Rosso. Quantifiers for randomness of chaotic pseudo-random number generators. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 367(1901):3281–3296, 2009. ISSN 1471-2962. . URL http://dx.doi.org/10.1098/rsta.2009.0075.

C. Gabriel, C. Wittmann, D. Sych, R. Dong, W. Mauerer, U. L. Andersen, C. Marquardt, and G. Leuchs. A generator for unique quantum random numbers based on vacuum states. *Nature Photonics*, 4(10):711–715, 2010. ISSN 1749-4893. . URL http://dx.doi.org/10.1038/NPHOTON.2010.197.

C. González, H. Larrondo, and O. Rosso. Statistical complexity measure of pseudorandom bit generators. *Physica A: Statistical Mechanics and its Applications*, 354:281–300, 2005.

D. Knuth. Sorting and searching, vol. 3 of the art of computer programming, section 6.2. 2, 1997.

H. A. Larrondo, L. De Micco, and C. M. González. Statistical complexity of chaotic pseudorandom number generators. *Phys. Rev. Lett*, 88:174102, 2002.

H. A. Larrondo, M. T. Martín, C. M. González, A. Plastino, and O. A. Rosso. Random number generators and causality. *Physics Letters A*, 352(4–5):421–425, 2006. ISSN 0375-9601. . URL http://www.sciencedirect.com/science/article/pii/S0375960105018232.

R. López-Ruiz, H. Mancini, and X. Calbet. A statistical measure of complexity. *Physics Letters A*, 209(5-6):321–326, 1995. ISSN 0375-9601. . URL http://www.sciencedirect.com/science/article/pii/0375960195008675.

G. Marsaglia. Yet another rng. *Posted to the electronic billboard sci. stat. math, August*, 1, 1994.

M. T. Martin, A. Plastino, and O. A. Rosso. Generalized statistical complexity measures: Geometrical and analytical properties. *Physica A*, 369:439–462, 2006. .

M. T. Martín, A. Plastino, and O. A. Rosso. Generalized statistical complexity measures: Geometrical and analytical properties. *Physica A: Statistical Mechanics and its Applications*, 369(2):439–462, 2006. .

F. Olivares, A. Plastino, and O. A. Rosso. Contrasting chaos with noise via local versus global information quantifiers. *Physics Letters A*, 376(19):1577–1583, 2012.

W. Payne, J. R. Rabung, and T. Bogyo. Coding the lehmer pseudo-random number generator. *Communications of the ACM*, 12(2):85–86, 1969.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2017. URL https://www.R-project.org/.

M. G. Ravetti, L. C. Carpi, B. A. Gonçalves, A. C. Frery, and O. A. Rosso. Distinguishing noise from chaos: objective versus subjective criteria using Horizontal Visibility Graph. *PLOS One*, 9(9):1–15, 2014. .

O. A. Rosso, M. T. Martin, A. Figliola, K. Keller, and A. Plastino. EEG analysis using wavelet-based information tools. *Journal of Neuroscience Methods*, 153:163–182, 2006. .

O. A. Rosso, F. Olivares, L. Zunino, L. De Micco, A. L. Aquino, A. Plastino, and H. A. Larrondo. Characterization of chaotic maps using the permutation bandt-pompe probability distribution. *The European Physical Journal B*, 86(4):116, 2013.

O. A. Rosso, R. Ospina, and A. C. Frery. Classification and verification of handwritten signatures with time causal information theory quantifiers. *PLoS ONE*, 11(12):e0166868, 2016. .

T. A. Schieber, L. Carpi, A. C. Frery, O. A. Rosso, P. M. Pardalos, and M. G. Ravetti. Information theory perspective on network robustness. *Physics Letters A*, 380:359–364, 2016. .

F. Traversaro, F. O. Redelico, M. R. Risk, A. C. Frery, and O. A. Rosso. Bandt-Pompe symbolization dynamics for time series with tied values: a data-driven approach. *Chaos*, page in press, 2018.

H. Wickham. *ggplot2: Elegant Graphics for Data Analysis*. Springer, 2009. ISBN 978-0-387-98140-6. .

S. Wold, K. Esbensen, and P. Geladi. Principal component analysis. *Chemometrics and intelligent laboratory systems*, 2(1-3):37–52, 1987.

H. Xiong, P. Shang, J. He, and Y. Zhang. Complexity and information measures in planar characterization of chaos and noise. *Nonlinear Dynamics*, pages 1–15, 2020.

L. Zunino, M. C. Soriano, and O. A. Rosso. Distinguishing chaotic and stochastic dynamics from time series by using a multiscale symbolic approach. *Phys. Rev. E*, 86:046210, 2012. . URL http://link.aps.org/doi/10.1103/PhysRevE.86.046210.