Accepted Manuscript

Characterization of autoregressive processes using entropic quantifiers

Francisco Traversaro, Francisco O. Redelico

PII: S0378-4371(17)30713-6
DOI: http://dx.doi.org/10.1016/j.physa.2017.07.025
Reference: PHYSA 18442

To appear in: *Physica A*

Received date : 10 April 2017
Revised date : 19 July 2017

Please cite this article as: F. Traversaro, F.O. Redelico, Characterization of autoregressive processes using entropic quantifiers, *Physica A* (2017), http://dx.doi.org/10.1016/j.physa.2017.07.025

Highlights

- We present a novel informational plane that can lead to better characterization of the investigated time series in both, the correlation structure and the probability distribution shape.
- We show the insensibility to the probability distribution of the Bandt and Pompe symbolization used to estimate the Shannon Entropy.
- Two examples are presented and our plane gives a deeper understanding of them.

Characterization of autoregressive processes using entropic quantifiers

Francisco Traversaro[a], Francisco O. Redelico[b,c]

[a] *Instituto Tecnológico de Buenos Aires (ITBA), C1106ACD, Av. Eduardo Madero 399, Ciudad Autónoma de Buenos Aires, Argentina.*
[b] *CONICET - Hospital Italiano de Buenos Aires, Departamento de Informática en Salud Tte. Gral. Juan Domingo Perón 4190,, Ciudad Autónoma de Buenos Aires, Argentina.*
[c] *Departamento de Ciencia y Tecnología, Universidad Nacional de Quilmes*

Abstract

The aim of the contribution is to introduce a novel information plane, the causal-amplitude informational plane. As previous works seems to indicate, Bandt and Pompe methodology for estimating entropy does not allow to distinguish between probability distributions which could be fundamental for simulation or for probability analysis purposes. Once a time series is identified as stochastic by the causal complexity-entropy informational plane, the novel causal-amplitude gives a deeper understanding of the time series, quantifying both, the autocorrelation strength and the probability distribution of the data extracted from the generating processes. Two examples are presented, one from climate change model and the other from financial markets.

Keywords: Permutation entropy; Time series analysis.

1. Introduction

The dynamical behavior of a complex system is usually recorded as a time series (TS). One important task is to understand the nature (i.e linear, chaotic, periodic, stochastic, etc.), along with others features that can help modeling the

Email addresses: francisco.redelico@gmail.com (Francisco Traversaro), francisco.redelico@gmail.com (Francisco O. Redelico)

data generator process of the time series at hand.

The use of quantifiers based on Information Theory has led to interesting results regarding the characterization of nonlinear chaotic dynamics, improving the understanding of their associated time series. *Permutation Entropy* (PE), which is a information theory quantifier, that has the same functional form as *Shannon Entropy* [1] but uses, in order to estimate the required probabilities, the symbolic methodology proposed by Bandt and Pompe [2] that reveals internal nonlinear causality, understood as autocorrelation, for a given time series. In [3, 4, 5, 6, 7] was found that the use of the *Permutation Entropy* and the *Statistical Complexity* in an informational plane is very useful to distinguish between chaotic (i.e deterministic) and stochastic dynamics. But once this discrimination is completed, it is time to look for a better understanding by analyzing the distribution of the data. In this manner, several contribution were presented before, in [3, 8] PE was applied for general non Gaussian $1/f^{\alpha}$ noise and the fractional Brownian motion and, in [9], theoretical curves for the PE of the fractional Brownian motion and fractional Gaussian noise were developed. In [3] PE was applied in general non Gaussian stochastic processes, and therefore no knowledge about their moments were obtained in order to compare with chaotic time series, and in [8] PE was applied in Gaussian time series. At this point the following question arise: *Is the Bandt and Pompe methodology useful to characterize the shape of the marginal probability distribution of a stochastic process?*. This is the question this paper addresses by means of simulating Gaussian and non-Gaussian autoregressive processes of order 1 and comparing their *Shannon Amplitude Entropy* (i.e the Shannon Entropy of the histogram of the data) and *Shannon Permutation Entropy*. Autoregressive processes are simulated within this paper because both the associate correlation structure and the probability distribution are well established and the correlation structure is easily manipulable through the correlation parameter. Although Gaussian autoregressive processes are well known [10], this is not the case of non Gaussian autoregressive processes. However, the later impacts in diverse fields of science and technology as diverse as *random number generators* [11], *modeling*

irregularly spaced transaction financial data [12], *foreign exchange rate volatility modeling* [13], *studying nervous systems mechanism* (*Spike sorting*)[14] and *speech signal analysis* [15] among others, inciting their study. The determination of the *Probability Density Function* (PDF) of the data generator process is a fundamental task in several areas like simulation modeling where a real time series should be reproduce numerically.

The paper reads as follows: Section 2 presents a brief introduction to *Permutation Entropy* and *Amplitude Entropy* and the stochastic processes simulated, Section 3 presents numerical results of the application of the entropies over the simulated stochastic processes, Section 4 is devoted to present some application cases and finally Section 5 is a discussion about the results.

2. Information quantifiers

A brief introduction to the entropy quantifiers considered within this paper is presented in order to make this paper self-contained. Given a continuous probability distribution function (PDF) $-f(x)-$ the entropy of this distribution is given by:

$$S[f] \;=\; -\int_{\Delta} f\;\ln(f)\;dx\;, \tag{1}$$

and is a measure of *global character* that it is not too sensitive to strong changes in the distribution taking place on a small-sized region. Let now $\boldsymbol{P} = \mathbb{P}(i) = \{p_i;\; i = 1, \cdots, N\}$, with $\sum_{i=1}^{N} p_i = 1$, be a discrete PDF, with N the number of possible states of the system under study. In the discrete case, we define a "normalized" Shannon entropy ($0 \leq \mathcal{H} \leq 1$) as

$$\mathcal{H}[\boldsymbol{P}] \;=\; S[\boldsymbol{P}]/S_{max} \;=\; \left\{-\sum_{i=1}^{N}\; p_i\;\ln(p_i)\right\}/S_{max}\;, \tag{2}$$

where the denominator $S_{max} = S[\boldsymbol{P_e}] = \ln N$ is that obtained by an uniform probability distribution $\boldsymbol{P_e} = \{p_i = 1/N,\; \forall i = 1, \cdots, N\}$.

There is no a universal way to compute the set P for every system. One possibility is to estimate each element p_i is by using the empirical histogram

of the values of the time series using some method that wishfully leads to a good estimation of the underlying probability distribution function of the data. This is done by using the histogram as the maximum likelihood estimator of the PDF, and computing the probabilities p_i as:

$$p_i = \frac{\sum_{l=1}^{T-m+1} \mathbb{1}(x_t \text{ belongs to bin } i)}{T-m+1}, \tag{3}$$

leading to the *Normalized Amplitude Entropy* $\mathcal{H}^{ML}$ (note: ML stands for the Maximum Likelihood method). Note that this entropy does not take account of the order of the data appearance in the time series.

Other way to compute the probabilities p_i is through the symbolization proposed by Bandt and Pompe [2]. Let $X_m(t) = (x_t, x_{t+1}, \ldots, x_{t+m-1})$ with $0 \leq t \leq T-m+1$ be a non-disjoint partition containing the vectors of real values of length m of the time series $\{X_t\}_{t\in T}$. Let $S_{m\geq 3}$ the symmetric group of order $m!$ form by all possible permutation of order m, $\pi_i = (i_1, i_2, \ldots, i_m) \in S_m$ ($i_j \neq i_k \ \forall j \neq k$ so every element in π_i is unique). We will call an element π_i in S_m a symbol or a motiv as well. Then $X_m(t)$ can be mapped to a symbol π_i in S_m for a given but otherwise arbitrary t. The m number of real values $X_m(t) = (x_t, x_{t+1}, \ldots, x_{t+m-1})$ are mapped onto their rank. The rank function is defined as:

$$R(x_{t+n}) = \sum_{k=0}^{m-1} \mathbb{1}(x_{t+k} < x_{t+n}) \tag{4}$$

where $\mathbb{1}$ is the indicator function (i.e $\mathbb{1}(Z) = 1$ if Z is true and 0 otherwise), $x_{t+n} \in X_m(t)$ with $0 < n \leq m-1$ and $1 \leq R(x_{t+n}) \leq m$. So the rank $R(min(x_{t+k})) = 1$ and $R(max(x_{t+k})) = m$. The complete alphabet is all the possible permutation of the ranks. Hence, any vector $X_m(t)$ is uniquely mapped onto $\pi_i = (R(x_t), R(x_{t+1}), \ldots, R(x_{t+m-1})) \in S_m$. With this Rank Permutation Mapping one simply maps each value x_i in $X_m(t)$ placing its rank $R(x_i) \in \{1, 2, \ldots, m\}$ in chronological order to form π_i in S_m. Using the rank permutation Mapping we compute $pi = \mathbb{P}(\pi_i)$ (see Fig. 1, bottom) ,

$$p_i = \frac{\sum_{l=1}^{T-m+1} \mathbb{1}(X_m(l) \text{ has ordinal patter } \pi_i \text{ in } S_m)}{T-m+1}, \tag{5}$$

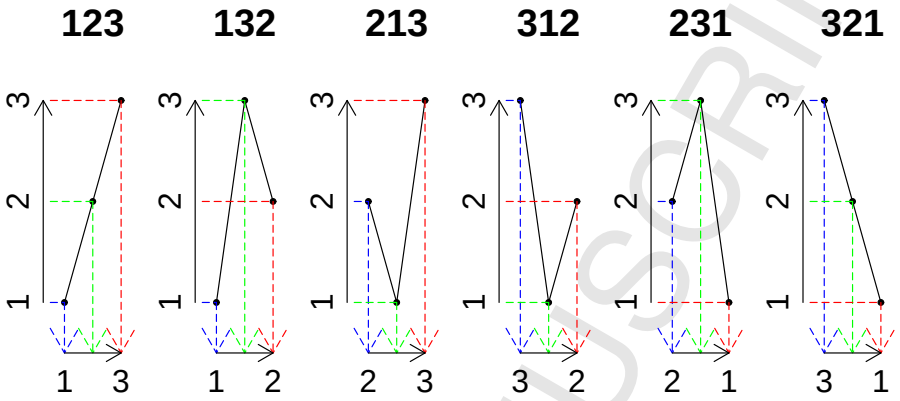

Figure 1: **Rank Permutation Mapping**: All symbols for $m = 3$ are shown. With this Rank Permutation Mapping one simply maps each value x_i in $X_m(t)$ placing its rank $R(x_i) \in \{1, 2, \ldots, m\}$ in chronological order to form π_i in S_m. Using the rank permutation Mapping we compute $p_i = P(\pi_i)$. It can be seen that the indexes of the vertical axis are fixed, ordered by amplitude (i.e ranks), and they are mapped onto the time axis. For each vector $\mathbf{X}_t^{(3)} = (x_t, x_{t+1}, x_{t+2})$, the resultant pattern $\tilde{\pi}^{(3)}$ can be obtained reading the labels in the horizontal axis from left to right (in chronological order).

where $\mathbb{1}$ is the indicator function and $i = 1, \ldots, m!$.

Using these probabilities, the *Normalized Permutation Entropy* $\mathcal{H}^{BP}$ for each m can be computed using Eq. 2 where $N = m!$ is the order of the symmetric group S_m and $S_{max} = log(N)$. Note that this entropy does not take account of the amplitude of the observations besides the relative amplitude between neighbor values. The set of $\boldsymbol{P} = \{p_i \; i = 1 \ldots N\}$ obtained as in Eq. 2 is what we will call from now onwards the Bandt and Pompe Distribution (BP PDF). The BP PDF has two free parameters: the embedding dimension m and the time delay τ but within this paper $\mathcal{H}^{BP}$ is calculated using a fixed $\tau = 1$ parameter in order to not confuse with temporal scales and all results are shown for the embedding dimension $m = 4$ because there was no significant difference in the results when the m parameter varied from $m = 3$ to $m = 6$.

Another useful informational quantifier is the statistical complexity [8] $-\mathcal{C}-$

defined as,

$$\mathcal{C}[\boldsymbol{P}] \;=\; \mathcal{Q}_J[\boldsymbol{P}, \boldsymbol{P_e}] \cdot \mathcal{H}[\boldsymbol{P}] \tag{6}$$

where $\mathcal{H}$ is the *Normalized Permutation Entropy*, $\mathcal{Q}_J$ is the disequilibrium, $\boldsymbol{P}$ is the BP PDF and $\boldsymbol{P_e} = \{p_i = 1/N,\ \forall i = 1, \cdots, N\}$. $\mathcal{Q}_J$ is defined in terms of the Jensen–Shannon divergence. That is, $\mathcal{Q}_J[\boldsymbol{P}, \boldsymbol{P_e}] \;= Q_0\ \mathcal{J}[\boldsymbol{P}, \boldsymbol{P_e}]$, with $\mathcal{J}[\boldsymbol{P}, \boldsymbol{P_e}] \;=\; S[(\boldsymbol{P}+\boldsymbol{P_e})/2] - S[\boldsymbol{P}]/2 - S[\boldsymbol{P_e}]/2$, and Q_0 a normalization constant.

3. Stochastic processes

A general linear stochastic model is described that supposes a time series to be generated by a linear aggregation of random shocks. [10]. A widely used stochastic model is the autoregressive model or order p,

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + ... + \phi_p z_{t-p} + a_t \tag{7}$$

where the current value z_t is expressed as a finite, linear aggregate of previous values of the process $\{z_{t-1}, z_{t-2}, \cdots, z_{t-p}\}$ and a shock a_t, distributed with mean 0 and finite variance σ^2.

If z_t is a wide-sense stationary process (i.e the mean μ and the variance σ_z^2 do not depend on time), the autocorrelation between z_t ans z_s depends only on the lag between t and s. This implies that the autocorrelation can be expressed as a function of the time lag ($\nu = t - s$) between observations:

$$R(\nu) = \frac{E(z_t - \mu)E(z_{t-\nu} - \mu)}{\sigma_z^2} \tag{8}$$

For simplicity the autocorrelation of lag $\nu = 1$ is expressed as $R(1) = \rho(r_t, r_{t+1}) = \rho_1$.

We consider three kind of autoregressive process: Gaussian, Exponential and Uniform. The difference between them is the form of the random shock a_t. Within this paper, for sake of simplicity we set the order of the autoregressive process $p = 1$ leading to the first-order autoregressive process AR(1).

Gaussian autoregressive model of order 1 AR(1): In the Gaussian process, Ec. 7 takes the form:

$$z_t = \phi_1 z_{t-1} + a_t \tag{9}$$

where the current value z_t is expressed as a finite, linear aggregate of the previous values of the process z_{t-1} and an independent and identically distributed (i.i.d.) random shocks a_t that has marginal Gaussian distribution with mean 0 and variance σ^2. For this process $\rho_1 = \phi_1$.

Exponential autoregressive model of order 1 NEARA(1): Many positive-valued time series have an exponential marginal distribution [13,14]. When the z_t random variable has an Exponential marginal distribution with λ parameter, the linear autoregressive model Ec. 7 takes the form:

$$z_t = a_t + \begin{cases} \beta.z_{t-1} & \text{w.p } \alpha \\ 0 & \text{w.p } 1-\alpha \end{cases} \tag{10}$$

with

$$a_t = \begin{cases} e_t & \text{w.p } \frac{1-\beta}{1-(1-\alpha)\beta} \\ (1-\alpha).\beta.e_t & \text{w.p } \frac{\alpha\beta}{1-(1-\alpha)\beta} \end{cases} \tag{11}$$

where "w.p." stands by "with probability". $\alpha > 0$ and $\beta > 0$ are free correlation parameters such as $\rho_1 = \alpha\beta$, providing that α and β are not both equal to one. a_t has a particular mixed exponential distribution, where e_t $\{t = 0, 1, 2..\}$ are independent identically distributed exponential variables with parameter $\lambda > 0$ in order to make the marginal distribution of z_t exponential with λ parameter. Note that with α or β equal to zero z_t are exponential i.i.d. and both α and β are nonnegative, so with this method the autocorrelations $\rho_k = (\alpha\beta)^k$ are positive and geometrically decreasing. This is unlike the Gaussian AR(1) model where ρ_1 can be negative. To extend the exponential models to this possibility, two sequences z_t and z'_t are cross-couple constructed, involving antithetic variables, developing the NEARA(1) model in [16].

a_t	0	$\frac{1}{k}$	$\frac{2}{k}$	...	$\frac{(k-1)}{k}$
$P(a_t)$	$\frac{1}{k}$	$\frac{1}{k}$	$\frac{1}{k}$	$\frac{1}{k}$	$\frac{1}{k}$

Table 1: Discrete Uniform Distribution for the random shock a_t in the AUR(1) model according to [17] where $k \in \{2, 3, ...\}$.

Uniform autoregressive model of order 1 UAR(1): Another useful model construction for non-Gaussian variate time series is the first order autoregressive process with uniform marginal distribution. The model responds to

$$z_t = \frac{1}{k} z_{t-1} + a_t \tag{12}$$

with $k \geq 2$. It has been shown in [17] that z_t would shield continuous uniform (0,1) marginal distribution if the i.i.d.random shocks a_t has the marginal discrete uniform distribution presented in Table 1 where $k \in \{2, 3, ...\}$ and $\rho_1 = 1/k$. This model is called UAR(1) and has the lag r autocorrelation $\rho_r = \rho^r = (1/k)^r$. If ρ_1 is $-1/k$ there are similar results for negatively autocorrelated models.

4. Numerical Results

The processes presented in the previous Section were simulated, varying the autocorrelation.

For the Gaussian processes nine positively and nine negatively autocorrelated time series were simulated, with $\rho_1 = \pm\{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$ and in addition with $\rho_1 = 0$, the well known Gaussian white noise. All these processes were simulated with shocks a_t with $\sigma = 1$ and $\mu = 0$. The exponential processes where simulated with α and β such $\rho_1 = \{-0.2, -0.1, 0, 0.125, 0.25, 0.5, 0.75\}$. The autoregressive uniform process was simulated for $\rho_1 = \pm\{1/2, 1/3, ...1/9\}$ plus the uncorrelated data with uniform marginal distribution. All series have length $N = 10^6$ points.

Using the *Information Thoery Quantifiers* presented in Section 2, we first analyze the localization in the $H^{BP} \times C$ informational plane (Fig. 1) of these

processes. Their planar localization is in accordance with others previously reported [3]. It should be pointed out that the planar localization of the Gaussian and non-Gaussian autoregressive process coincide. Uncorrelated time series, i.e. Uniform, Gaussian and Exponential random numbers, are located in the same region $(\mathcal{H}^{BP}, \mathcal{C}) = (1, 0)$, with maximum entropy and minimum complexity. As the autocorrelation increases in absolute value it moves to a higher $\mathcal{C}$ and a lower $\mathcal{H}^{BP}$ planar localization. The three processes are mostly indistinguishable in the $\mathcal{H}^{BP} \times \mathcal{C}$ plane, specially for lower correlation parameters. Thus, the underlying marginal density distribution is not a factor of the location of the process in the $\mathcal{H}^{BP} \times \mathcal{C}$ informational plane, and this is because the Bandt and Pompe methodology for calculating p_i takes account about the correlation structure of the time series but not about the probability distribution of the data process generator. This fact can be observed in Fig. 3 where the three histograms for the BP PDF from non-correlated Uniform, Gaussian and Exponential standard distributions are presented. As it can be seen, the three histograms are mostly indistinguishable, confirming the first insight that the PE is not capable to discriminate between different probability distributions.

To get a better characterization of time series in a simple manner, we present a *novel information plane*: the $\mathcal{H}^{\mathcal{BP}} \times \mathcal{H}^{\mathcal{ML}}$ plane (Fig. 4). In this plane the simulated stochastic processes are clearly differentiated both in the x-axis , $\mathcal{H}^{\mathcal{BP}}$ and in the y-axis $\mathcal{H}^{\mathcal{ML}}$. The higher correlation coefficient is, the lower $\mathcal{H}^{\mathcal{BP}}$ value and $\mathcal{H}^{\mathcal{ML}}$ remains constant. In contrast, as the symmetry increases and the kurtosis decreases, the value of $\mathcal{H}^{\mathcal{ML}}$ also increases and the value for $\mathcal{H}^{\mathcal{BP}}$ remains constant. Hence, the planar localization of a given time series in this plane can discriminate not only the autocorrelation structure ($\mathcal{H}^{\mathcal{BP}}$) but also the probability distribution of the time series ($\mathcal{H}^{ML}$).

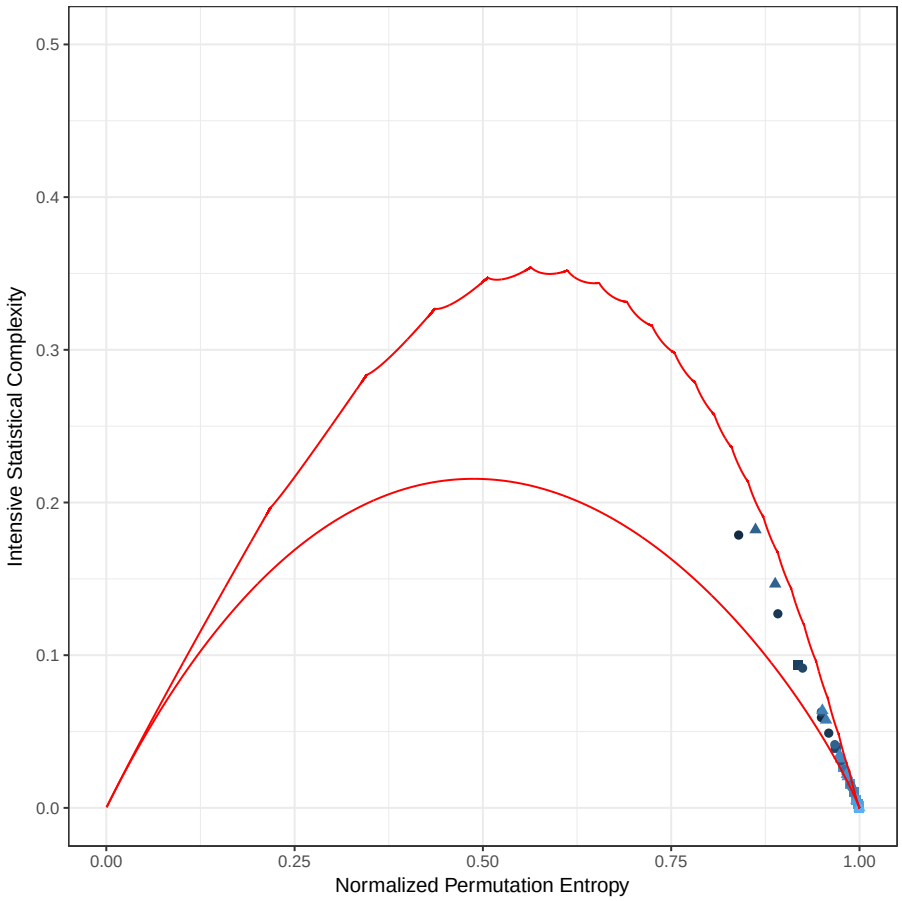

Figure 2: Localization in the causal entropy-complexity plane,considering $m = 4, \tau = 1$, of the stochastic processes simulated in the present work, in squares: exponential with parameter $\lambda = 1$, circles: standard Gaussian distribution and triangles: discrete uniform with support $[0, 1]$. Darker color represents higher values in the autocorrelation. Average values over ten realizations with different seeds are presented. The continuous lines represent the curves of maximum and minimum statistical complexity, C_{max} and C_{min}, for a given normalized permutation entropy $\mathcal{H}^{BP}$ [8]. As it was expected, the different location in the plane of the process depends, in first place of the kind of process (i.e deterministic, stochastic), that in this paper all three are autoregressive stochastic processes, and in second place, depends on the strength of the autocorrelation. It can be noticed that the underlying marginal density distribution is not a factor of the location of the process in this plane. For that reason the $\mathcal{H}^{BP} \times \mathcal{H}^{ML}$ plane is presented in Fig. 4.

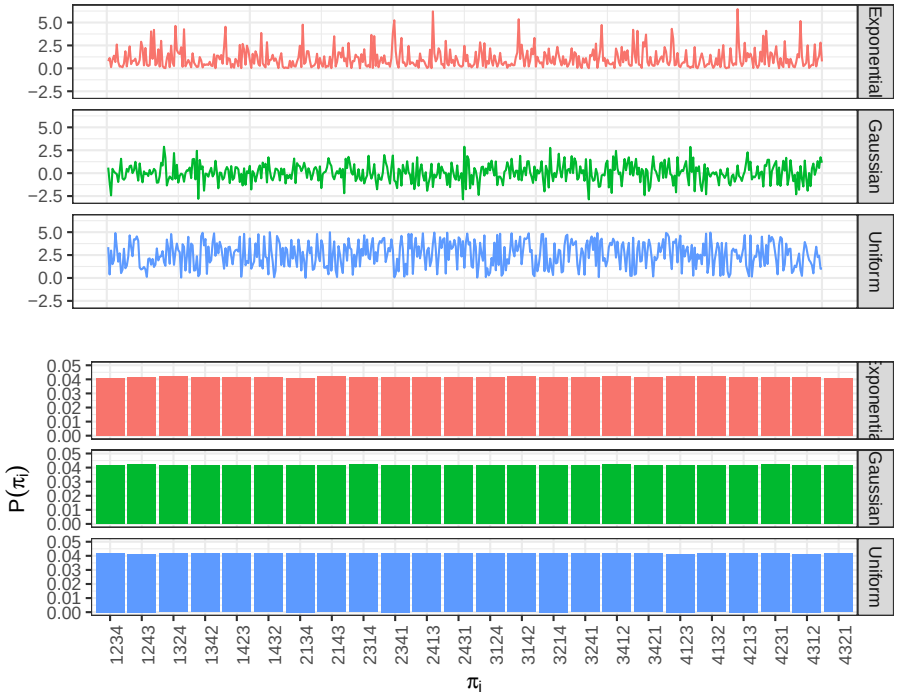

Figure 3: Time Series and their histogram for the symbolic sequence generated using Bandt and Pompe methodology for the three uncorrelated stochastic processes, in red: exponential with parameter $\lambda = 1$, green: standard Gaussian distribution and light blue: discrete uniform with support $[0, 1]$. This Figure depicts the fact that Shannon entropy endowed with probabilities using the Bandt and Pompe methodology fails to distinguish between probability distributions.

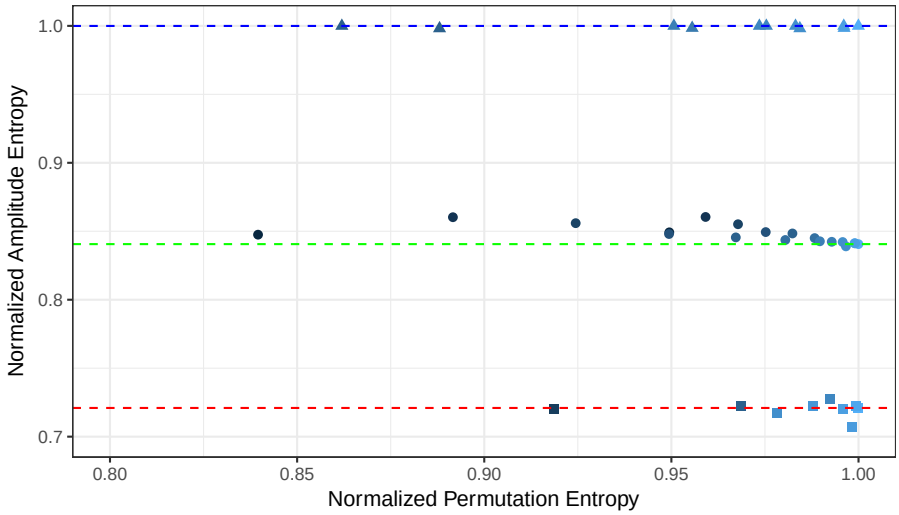

Figure 4: Causal $\mathcal{H}^{BP} \times \mathcal{H}^{ML}$ considering $m = 4, \tau = 1$, for the stochastic processes simulated in the present work, in squares: exponential with parameter $\lambda = 1$, circles: standard Gaussian distribution and triangles: discrete uniform with support $[0, 1]$. The dashed lines indicates the theoretical entropy of the given distribution (exponential red, Gaussian green and uniform blue). Darker color represents higher values in the autocorrelation. The stochastic processes are clearly differentiated both in the x-axis , $\mathcal{H}^{BP}$ and in the y-axis $\mathcal{H}^{ML}$. The Normalized Amplitude Shannon Entropy depends mainly on the shape of the distribution, but in a lesser way on the form that the histogram was constructed. The histogram construction of the uniform distribution only depends on the length of the support interval of the random variable, in this case, for all the UAR(1) processes is the same: $\mathcal{I}(0, 1)$. A similar case occurs with the exponential variable but in this case depends on the rate. The NEARA(1) processes were all simulated with $\lambda = 1$. On the other hand the histogram of a Gaussian AR(1) process depends on the variance σ_z^2 that depends on the autoregressive coefficient (ϕ) and on the shock variance σ^2. The processes simulated in this paper have all the same variance $\sigma^2 = 1$, but differs in ϕ. This is how the variation in the y-axis for these processes can be interpreted. Nevertheless this disturbance does not interferes with the discrimination of the processes with different marginal distributions.

5. Real applications

Two real applications are presented for the novel $\mathcal{H}^{BP} \times \mathcal{H}^{ML}$ plane. One referred to climate and the other to the stock market.

5.1. Climate change time series

Paleoclimatic data, taken from fossil corals, ice sheet, etc., make possible the study of the history of climate change within the Earth evolution. Studies using paleoclimatic data show that the El Niño/Southern Oscillation (ENSO) has been present in the Earth climate for at least the past 130,000 years [18]. We study the dynamics of the Holocene proxy ENSO record corresponding to Pallcacocha Lake sedimentary data, presented in [19]. The Holocene period is the geological epoch that began after the Pleistocene approximately 11,700 years before present [20, 21] [22, 19]. The proxy record was obtained from the analysis of clastic laminae deposition in sediment two $8 - m$ cores retrieved from the Pallcacocha Lake in Ecuador, interpolated to a sample time of one year using a cubic Hermite polynomial. This data set was previously analyzed in [23, 24, 25] among others. In Fig. 5 the $\mathcal{H}^{\mathcal{BP}} \times \mathcal{H}^{\mathcal{ML}}$ plane is presented for the Pallcacocha Lake sedimentary data and its evolution over time. It shows that the correlation dynamics (measured by $\mathcal{H}^{\mathcal{BP}}$) varies over the years during the Holocene leading to chaotic and hyper-chaotic behavior, see [25] and the probability distribution of the data generation process also changes. It can be noticed from the Figure that most of the variation during the Holeocene is due to change in the autocorrelation structure reflected in the Normalized Permutation Entropy that ranges between $0.332 - 0.916$ but there is also a change in the probability distribution, suggested by the interval between $0.639 - 0.836$ for the Normalized Amplitude Entropy. Probability distribution remains between the Gaussian and the Exponential distribution almost the time, but between 2500 and 1000 years before present the $\mathcal{H}^{\mathcal{ML}}$ decreases below the theoretical exponential line given an indication of extreme events on the period. In Figure 6 the main statistical indicators that affect both entropies are presented.

Normalized Amplitude Entropy $\mathcal{H}^{\mathcal{ML}}$ depends on the shape of the PDF of the data, represented by the skewness and kurtosis and in a lesser extent, the standard deviation (the coefficient of variation -CV- in this case) and Normalized Permutation Entropy $\mathcal{H}^{\mathcal{BP}}$ depends on the structure of the autocorrelation, represented in this case by $\rho(r_t, r_{t+1})$. Near 9000 thousands years ago the data was almost Gaussian (i.e low CV, kurtosis near 3 and skewness 0) but with strong autocorrelation, near 1, and this is reflected in the $\mathcal{H}^{\mathcal{BP}} \times \mathcal{H}^{\mathcal{ML}}$ plane of Figure 5 in a high $\mathcal{H}^{\mathcal{ML}}$ and a low $\mathcal{H}^{\mathcal{BP}}$. According to this analysis, the probability of extreme events increased during a time in the Holeoscene period but returned in the present near the starting point. This is a new information about ENSO evolution, up to our knowledge, that could bring a new insight in the research in climate change, studying the increase of extreme events by the probability distribution that generates the data.

5.2. Financial time series

We employed daily data beginning in January 2, 1995 and ending in July 23, 2007. There is on average 3100 observations for each index. All country indexes were studied for the same time period.

Let z_t be the equity index of a stock on a time t, the continuously compounded return or log return r_t of an asset is defined as the natural logarithm of its simple gross return:

$$r_t = \log\left(\frac{z_{t+1}}{z_t}\right) \tag{13}$$

We study the log return time series of the stock market of 30 countries, 13 emerging, 15 developed and 2 frontiers countries, according the Morgan Stanley Capital Index methodology (MSCI). These data were downloaded from https://finance.yahoo.com/

The Efficient Market Hypothesis (EMH) states that the returns of an efficient financial market is characterized as white noise. However it is known that stock markets indexes have multifractal structure [26]. In [27] was found that emerging markets have greater correlation than developed markets and in

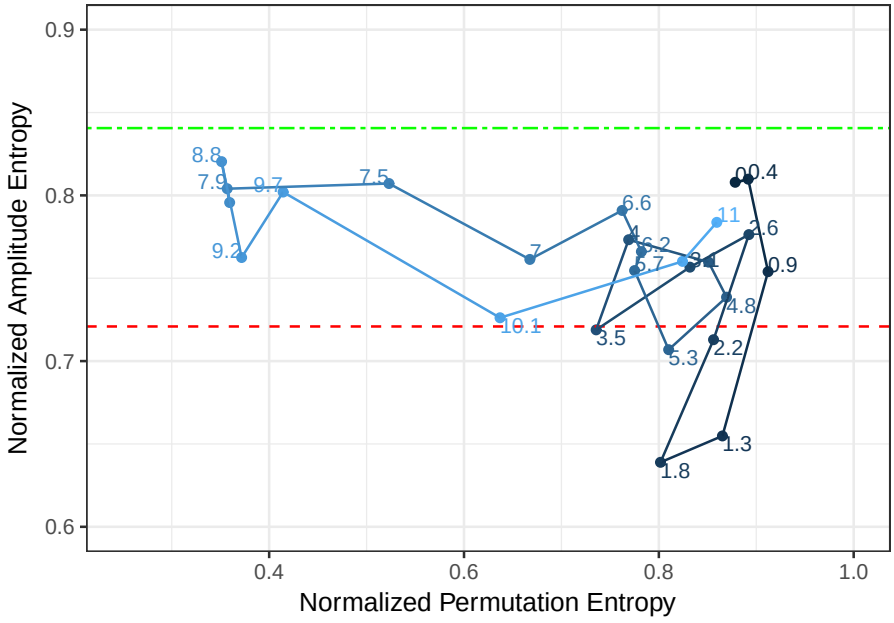

Figure 5: $\mathcal{H}^{\mathcal{BP}} \times \mathcal{H}^{\mathcal{ML}}$ for the ENSO evolution during the Holocene period. When darker, closer in time. The labels stand for the years (in thousands) before present. The dashed lines indicates the theoretical entropy of the given distribution, in red: exponential with parameter $\lambda = 1$ and green: standard Gaussian distribution. It can be noticed from the Figure that most of the variation during the Holocene is due to change in the autocorrelation structure and not in the probability distribution. Probability distribution remains almost constant and has a lower entropy than an exponential distribution almost all the time, but between 2500 and 1000 years before present the $\mathcal{H}^{\mathcal{ML}}$ decreases below the theoretical exponential line given an indication of extreme events on this period.

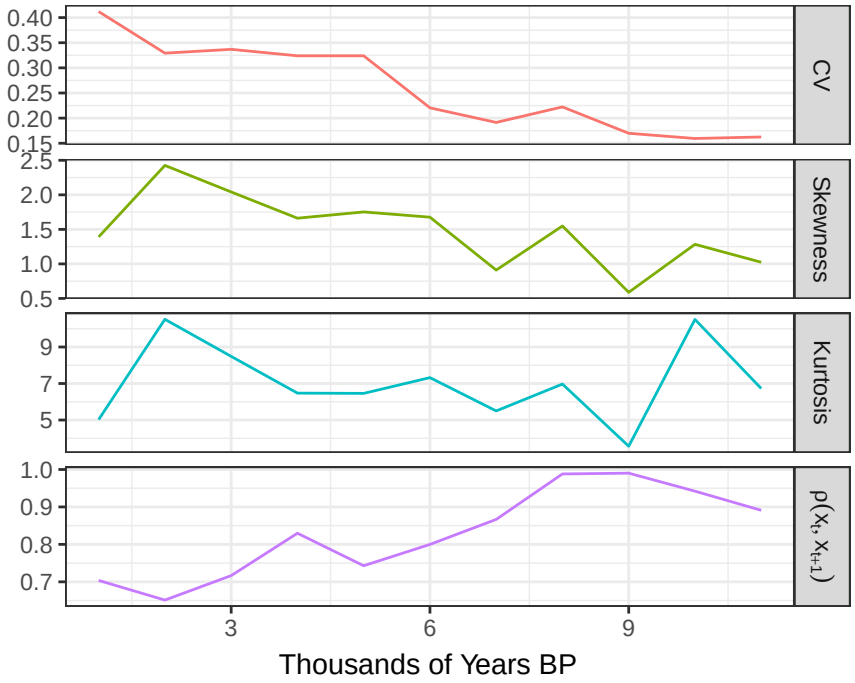

Figure 6: Main statistical indicators that affect both entropies are presented for the dynamics of the Holocene proxy ENSO record corresponding to Pallcacocha Lake sedimentary data. Normalized Amplitude Entropy $\mathcal{H}^{\mathcal{ML}}$ depends on the shape of the PDF of the data, represented by the skewness and kurtosis and in a lesser extent, the standard deviation (the coefficient of variation -CV- in this case) and Normalized Permutation Entropy $\mathcal{H}^{\mathcal{BP}}$ depends on the structure of the autocorrelation, represented in this case by $\rho(r_t, r_{t+1})$. Near 9000 thousands years ago the data was almost Gaussian (i.e low CV, kurtosis near 3 and skewness 0) but with strong autocorrelation, near 1, and this is reflected in the $\mathcal{H}^{\mathcal{BP}} \times \mathcal{H}^{\mathcal{ML}}$ plane of Figure 5 in a high $\mathcal{H}^{\mathcal{ML}}$ and a low $\mathcal{H}^{\mathcal{BP}}$.

[28] it is shown that deviations from the EMH could be associated with the degree of development. As [29, 30] pointed out, deviation from the EMH, and therefore the multi-fractal structure, could be explained in two ways: by the departure from Gaussianity and by the presence of autocorrelations. In Fig.7 the $\mathcal{H}^{\mathcal{BP}} \times \mathcal{H}^{\mathcal{ML}}$ plane is presented for financial markets for all 30 countries studied in this contribution. This Figure agrees with the results in [28, 29, 30] but also reflects the departure from Gaussianity. Even all time series are below the exponential distribution. While Argentina (MERVAL index) is located near the emergent countries, Venezuela (IBVC index) is far from both groups, emergent and developed; this fact shows that even the Efficient Market Hypothesis is contrasted using this plane but the Morgan Stanley Capital Index (MSCI) includes other sources of information, as political features, for the classification. Markets that have lower $\mathcal{H}^{\mathcal{ML}}$ in general reflect more probability for extreme events. With this plane a description of the inefficiency, measured as the departure from the white noise in the probability distribution and in the correlation structure, of the financial markets is provided, and it could be used as an important tool for the decision maker in this field as it can grasp no only the risk by the dynamical behavior of the market but also the probability of extreme events by the shape of the distribution.

In Tables 5.2 and 5.2 the main statistical indicators that affect both entropies are presented. Normalized Amplitude Entropy $\mathcal{H}^{\mathcal{ML}}$ depends on the shape of the PDF of the data, represented by the skewness and kurtosis and in a lesser extent, the standard deviation (this is because the entropy is normalized) and Normalized Permutation Entropy $\mathcal{H}^{\mathcal{BP}}$ depends on the structure of the autocorrelation, represented in this case by $\rho(r_t, r_{t+1})$. The excess of kurtosis from 3 is an indicator of the departure from Gaussianity, present in all the markets. A positive skewness suggest that the is an inclination in the index to close in a downward trend. The different values of these indicators impact in the value of $\mathcal{H}^{\mathcal{ML}}$. On the other hand while the autocorrelation coefficient varies between indexes, the $\mathcal{H}^{\mathcal{BP}}$ remains constant along all the countries, implying that the correlation structure is far more complex than an AR(1) process.

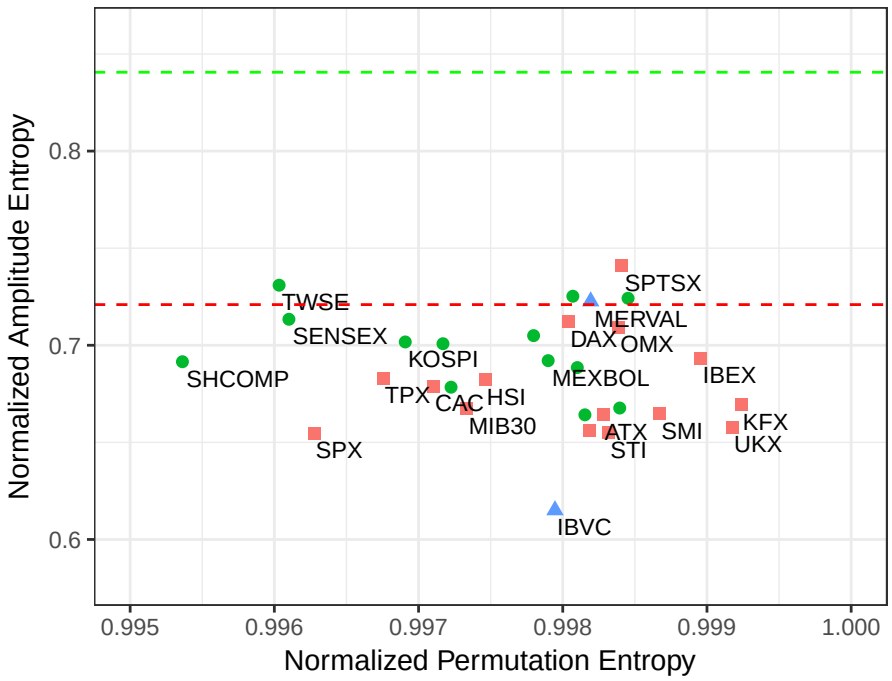

Figure 7: $\mathcal{H}^{\mathcal{BP}} \times \mathcal{H}^{\mathcal{ML}}$ for 30 countries. In red squares are the 15 developed countries, in green circles, 13 emerging countries, and in light blue triangles the 2 frontier countries. The dashed lines indicates the theoretical entropy of the given distribution, in red: exponential with parameter $\lambda = 1$ and green: standard Gaussian distribution. Even all time series are below the exponential distribution entropy (red dashed line) , developed countries have lower $\mathcal{H}^{\mathcal{ML}}$ than emerging ones in general. Two frontier countries were added in the plot: Argentina and Venezuela. While Argentina is located near the emergent countries, Venezuela is far from both groups, emergent and developed; this fact shows that even the Efficient Market Hypothesis is contrasted using this plane and the Morgan Stanley Capital Index (MSCI) includes other sources of information, as political features, for the classification. That can be understood as noise for the purpose of this plot. This new plane shows insights from financial series that are very useful for decision making.

Country	Code	Index	Std. Deviation	Skewness	Kurtosis	$\mathcal{H}^{\mathcal{ML}}$	$\rho(r_t, r_{t+1})$	$\mathcal{H}^{\mathcal{BP}}$
Emergent								
Brazil	IBOV	Brazil Bovespa Stock	0.01878	-0.09017	6.74441	0.725240878	0.01545	0.998069939
China	SHCOMP	Shangai Stock Exchange Composite	0.01571	-0.09326	7.56212	0.691524927	0.00504	0.995362536
Greece	ASE	ASE General	0.01765	-0.00937	7.19686	0.704972496	0.07700	0.997797882
India	SENSEX	Bombay Stock Exchange Sensitive	0.01633	-0.18087	9.35773	0.713399367	0.07133	0.996100933
Indonesia	JCI	Jakarta Composite	0.01486	-0.67816	8.87694	0.678410191	0.11022	0.997225449
Ireland	ISEQ	Irish Overall	0.01451	-0.56634	10.53043	0.664173043	0.04400	0.998153752
Korea	KOSPI	Korea Composite	0.01711	-0.52719	8.20985	0.701709654	0.01953	0.996907849
Malaysia	KLCI	Kuala Lumpur Composite	0.01119	-0.24248	9.411541	0.700792033	-0.09151	0.997169286
Mexico	MEXBOL	Mexico Bolsa	0.01415	0.02337	7.49024	0.692089705	0.09156	0.997898983
Philippines	PCOMP	Philippines Composite	0.01390	0.32669	17.55425	0.667679565	0.12300	0.998395004
Taiwan	TWSE	Taiwan Taiex	0.01515	-0.23434	5.68430	0.730935015	0.04490	0.996032884
Thailand	SET	Stock Exchange of Thailand	0.01461	-0.70709	11.66831	0.688520396	0.03502	0.998100887
Turkey	XU100	Istambul Stock Exchange National 100	0.02340	0.01436	9.95065	0.724199712	0.00479	0.998452833
Frontier								
Argentina	MERVAL	Argentina Merval	0.02121	-0.11700	7.89989	0.722627547	0.05484	0.998193693
Venezuela	IBVC	Venezuela Stock MKT	0.01477	-0.27208	26.97341	0.61503886	0.12370	0.997946076

Table 2: Main statistical indicators for the Financial Time Series for Emergent and Frontier countries. Normalized Amplitude Entropy $\mathcal{H}^{\mathcal{ML}}$ depends on the shape of the PDF of the data, represented by the skewness and kurtosis and in a lesser extent, the standard deviation, and Normalized Permutation Entropy $\mathcal{H}^{\mathcal{BP}}$ depends on the structure of the autocorrelation, represented in this case by $\rho(r_t, r_{t+1})$.

Country	Code	Index	Std. Deviation	Skewness	Kurtosis	$\mathcal{H}^{\mathcal{ML}}$	$\rho(r_t, r_{t+1})$	$\mathcal{H}^{\mathcal{BP}}$
Developed								
Austria	ATX	Australian Traded	0.01509	-0.30355	10.02285	0.664091001	0.06345	0.998279483
Belgium	BEL20	BEL20	0.01330	0.04558	8.78691	0.655916946	0.06578	0.998183036
Canada	SPTSX	S&P/TSX Composite	0.01204	-0.63524	11.61914	0.74109337	-0.01982	0.998409756
Denmark	KFX	OMX Copenhagen 20	0.01326	-0.25255	8.53230	0.669285595	0.03166	0.999236637
France	CAC	CAC 40	0.01546	0.02853	7.59435	0.679019783	-0.03262	0.997102279
Germany	DAX	DAX	0.01590	0.02741	7.07142	0.712247375	-0.01668	0.998039359
Hong Kong	HSI	Hang Seng	0.01590	-0.06732	10.72049	0.682334856	-0.01464	0.997462628
Italy	MIB30	Milan MIB30	0.01465	-0.04972	9.21087	0.667288604	-0.02176	0.997329488
Japan	TPX	Tokio Stock Price	0.01437	-0.35025	8.93408	0.683068356	0.01708	0.99675351
Singapore	STI	Straits Times	0.01235	-0.42140	9.03488	0.655154607	0.00960	0.998320456
Spain	IBEX	IBEX 35	0.01557	0.11338	7.79378	0.693421573	0.00264	0.998957353
Sweden	OMX	OMX Stockholm 30	0.01605	0.06962	6.15044	0.70933964	-0.02431	0.998385254
Switzerland	SMI	Swiss Market	0.01232	0.01696	9.08348	0.664941954	0.02465	0.99867319
United Kingdom	UKX	FTSE 100	0.01272	-0.12137	8.84082	0.657877244	-0.05049	0.999176443
USA	SPX	S&P 500	0.01324	-0.17111	10.59726	0.654697844	-0.08382	0.996275065

Table 3: Main statistical indicators for the Financial Time Series for Developed countries. Normalized Amplitude Entropy $\mathcal{H}^{\mathcal{ML}}$ depends on the shape of the PDF of the data, represented by the skewness and kurtosis and in a lesser extent, the standard deviation, and Normalized Permutation Entropy $\mathcal{H}^{\mathcal{BP}}$ depends on the structure of the autocorrelation, represented in this case by $\rho(r_t, r_{t+1})$.

6. Conclusion

In summary, in this paper, we introduce a new informational plane, $\mathcal{H}^{BP} \times \mathcal{H}^{ML}$, which is very simple and fast to compute. This plane complements the $\mathcal{H}^{BP} \times \mathcal{C}$ giving global information about the family distribution of the data generator process. This informational plane allows to discriminate between probability distribution as well as the autocorrelation structure presented in the time series *leading a complete description of the data generation process.* In all the informational planes presented in the literature [5, 7] while the dynamics of the processes are reflected in their location in an informational plane, this not necessarily means that given a location of a time series in an informational plane one can determine for sure what kind of process is the generator of that data, and the same problematic persists in this $\mathcal{H}^{BP} \times \mathcal{H}^{\mathcal{ML}}$ plane. One aspect to consider is that any entropy measure is a measure of global character, and is not too sensitive to strong changes on the distribution taking place on small sized region. Once the empirical PDF is calculated for a given time series, the order of the elements to which the probability is calculated (for example in this contribution for the Shannon Amplitude Entropy are the bins, and for the Shannon Permutation Entropy are the symbols π_i) does not matter so, given a distribution with similar probabilities for the bins, the Shannon Amplitude Entropy will not differentiate between unimodal or bimodal distributions. For detection of differences in local aspects of the PDF, other measures can be considered as the Fisher Information Measure [31]. The potential application of this plane is presented with two examples from actives research areas: climate change model and efficiency in financial markets.In the former, evidence about extreme evens change during the Holocene is obtained and in the last, a new dimension of financial risk is grasp, the presence of very asymmetric distribution in the emergent countries. We conclude by encouraging researchers to use this informational plane along the well established $\mathcal{H}^{BP} \times \mathcal{C}$ plane to characterize autoregressive process and stochastic processes in general.

References

[1] C. E. Shannon, A mathematical theory of communication, ACM SIGMOBILE Mobile Computing and Communications Review 5 (1) (2001) 3–55.

[2] C. Bandt, B. Pompe, Permutation entropy: a natural complexity measure for time series, Physical review letters 88 (17) (2002) 174102.

[3] O. Rosso, H. Larrondo, M. Martin, A. Plastino, M. Fuentes, Distinguishing noise from chaos, Physical review letters 99 (15) (2007) 154102.

[4] O. A. Rosso, L. C. Carpi, P. M. Saco, M. G. Ravetti, A. Plastino, H. A. Larrondo, Causality and the entropy–complexity plane: Robustness and missing ordinal patterns, Physica A: Statistical Mechanics and its Applications 391 (1) (2012) 42–55.

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

[6] L. Zunino, M. C. Soriano, O. A. Rosso, Distinguishing chaotic and stochastic dynamics from time series by using a multiscale symbolic approach, Physical Review E 86 (4) (2012) 046210.

[7] O. A. Rosso, F. Olivares, L. Zunino, L. De Micco, A. L. Aquino, A. Plastino, H. A. Larrondo, Characterization of chaotic maps using the permutation bandt-pompe probability distribution, The European Physical Journal B 86 (4) (2013) 1–13.

[8] O. A. Rosso, L. De Micco, H. A. Larrondo, M. T. Martín, A. Plastino, Generalized statistical complexity measure, International Journal of Bifurcation and Chaos 20 (03) (2010) 775–785.

[9] L. Zunino, D. Pérez, M. Martín, M. Garavaglia, A. Plastino, O. Rosso, Permutation entropy of fractional brownian motion and fractional gaussian noise, Physics Letters A 372 (27) (2008) 4768–4774.

[10] G. E. Box, G. M. Jenkins, G. C. Reinsel, G. M. Ljung, Time series analysis: forecasting and control, John Wiley & Sons, 2015.

[11] A. Lawrance, Uniformly distributed first-order autoregressive time series models and multiplicative congruential random number generators, Journal of Applied Probability (1992) 896–903.

[12] R. F. Engle, J. R. Russell, Autoregressive conditional duration: a new model for irregularly spaced transaction data, Econometrica (1998) 1127–1162.

[13] C. Hafner, Nonlinear time series analysis with applications to foreign exchange rate volatility, Springer Science & Business Media, 2013.

[14] S. Farashi, Spike sorting method using exponential autoregressive modeling of action potentials, World Academy of Science, Engineering and Technology, International Journal of Medical, Health, Biomedical, Bioengineering and Pharmaceutical Engineering 8 (12) (2015) 864–870.

[15] K. Ishizuka, H. K. Solvang, T. Nakatani, Speech signal analysis with exponential autoregressive model., in: ICASSP (1), 2005, pp. 225–228.

[16] A. Lawrance, P. Lewis, A new autoregressive time series model in exponential variables (near (1)), Advances in Applied Probability (1981) 826–845.

[17] M. R. Chernick, A limit theorem for the maximum of autoregressive processes with uniform marginal distributions, The Annals of Probability (1981) 145–149.

[18] M. A. Cane, The evolution of el niño, past and future, Earth and Planetary Science Letters 230 (3) (2005) 227–240.

[19] C. Moy, G. Seltzer, D. Rodbell, D. Anderson, Laguna pallcacocha sediment color intensity data, IGBP PAGES/World Data Center for Paleoclimatology Data Contribution Series 76.

[20] G. Ogg, International stratigraphic chart, International Commission on Stratigraphy.

[21] M. Walker, S. Johnsen, S. O. Rasmussen, T. Popp, J.-P. Steffensen, P. Gibbard, W. Hoek, J. Lowe, J. Andrews, S. Björck, et al., Formal definition and dating of the gssp (global stratotype section and point) for the base of the holocene using the greenland ngrip ice core, and selected auxiliary records, Journal of Quaternary Science 24 (1) (2009) 3–17.

[22] C. M. Moy, G. O. Seltzer, D. T. Rodbell, D. M. Anderson, Variability of el niño/southern oscillation activity at millennial timescales during the holocene epoch, Nature 420 (6912) (2002) 162–165.

[23] G. L. Ferri, A. Figliola, O. A. Rosso, Tsallis statistics in the variability of el niño/southern oscillation during the holocene epoch, Physica A: Statistical Mechanics and its Applications 391 (5) (2012) 2154–2162.

[24] L. C. Carpi, P. M. Saco, A. Figliola, E. Serrano, O. A. Rosso, Analysis of an el nino-southern oscillation proxy record using information theory quantifiers, Analyzing Fish CF Index Through Skew-Gaussian Information Theory JE Contreras-Reyes.

[25] A. A. Tsonis, Dynamical changes in the enso system in the last 11,000 years, Climate dynamics 33 (7-8) (2009) 1069–1074.

[26] B. B. Mandelbrot, A multifractalwalkdown, Scientific American (1999) 71.

[27] M. Pasquini, M. Serva, Multiscale behaviour of volatility autocorrelations in a financial market, Economics Letters 65 (3) (1999) 275–279.

[28] L. Zunino, B. M. Tabak, A. Figliola, D. Pérez, M. Garavaglia, O. Rosso, A multifractal approach for stock market inefficiency, Physica A: Statistical Mechanics and its Applications 387 (26) (2008) 6558–6566.

[29] L. Zunino, M. Zanin, B. M. Tabak, D. G. Pérez, O. A. Rosso, Forbidden patterns, permutation entropy and stock market inefficiency, Physica A: Statistical Mechanics and its Applications 388 (14) (2009) 2854–2864.

[30] L. Zunino, A. Figliola, B. M. Tabak, D. G. Pérez, M. Garavaglia, O. A. Rosso, Multifractal structure in latin-american market indices, Chaos, Solitons & Fractals 41 (5) (2009) 2331–2340.

[31] B. R. Frieden, Science from Fisher information: a unification, Cambridge University Press, 2004.