Physica A 390 (2011) 876–890

Contents lists available at ScienceDirect

Physica A

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

Commodity predictability analysis with a permutation information theory approach

Luciano Zunino [a,b,c,*], Benjamin M. Tabak [d,e], Francesco Serinaldi [f], Massimiliano Zanin [g], Darío G. Pérez [h], Osvaldo A. Rosso [i,j]

[a] Instituto de Física Interdisciplinar y Sistemas Complejos (IFISC) CSIC-UIB, Campus Universitat de les Illes Balears, E-07122 Palma de Mallorca, Spain
[b] Centro de Investigaciones Ópticas (CONICET La Plata - CIC), C.C. 3, 1897 Gonnet, Argentina
[c] Departamento de Ciencias Básicas, Facultad de Ingeniería, Universidad Nacional de La Plata (UNLP), 1900 La Plata, Argentina
[d] Banco Central do Brasil - SBS Quadra 3, Bloco B, 9 andar, DF 70074-900, Brazil
[e] Universidade Catolica de Brasilia, Brasilia, DF, Brazil
[f] Dipartimento GEMINI, Università della Tuscia, Via S. Camillo de Lellis snc, 01100 Viterbo, Italy
[g] Universidad Autónoma de Madrid, 28049 Madrid, Spain
[h] Instituto de Física, Pontificia Universidad Católica de Valparaíso (PUCV), 23-40025 Valparaíso, Chile
[i] Instituto de Ciências Exatas (Fisica), Universidade Federal de Minas Gerais, Av. Antônio Carlos, 6627 - Campus Pampulha, 31270-901 Belo Horizonte - MG, Brazil
[j] Chaos & Biology Group, Instituto de Cálculo, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Pabellón II, Ciudad Universitaria, 1428 Ciudad de Buenos Aires, Argentina

ARTICLE INFO

Article history:
Received 10 August 2010
Received in revised form 8 October 2010
Available online 28 November 2010

Keywords:
Commodity efficiency
Complexity–entropy causality plane
Permutation entropy
Permutation statistical complexity
Bandt and Pompe method
Ordinal time series analysis

ABSTRACT

It is widely known that commodity markets are not totally efficient. Long-range dependence is present, and thus the celebrated Brownian motion of prices can be considered only as a first approximation. In this work we analyzed the predictability in commodity markets by using a novel approach derived from Information Theory. The complexity–entropy causality plane has been recently shown to be a useful statistical tool to distinguish the stage of stock market development because differences between emergent and developed stock markets can be easily discriminated and visualized with this representation space [L. Zunino, M. Zanin, B.M. Tabak, D.G. Pérez, O.A. Rosso, Complexity–entropy causality plane: a useful approach to quantify the stock market inefficiency, Physica A 389 (2010) 1891–1901]. By estimating the permutation entropy and permutation statistical complexity of twenty basic commodity future markets over a period of around 20 years (1991.01.02–2009.09.01), we can define an associated ranking of efficiency. This ranking is quantifying the presence of patterns and hidden structures in these prime markets. Moreover, the temporal evolution of the commodities in the complexity–entropy causality plane allows us to identify periods of time where the underlying dynamics is more or less predictable.

1. Introduction

In the past, before the existence of money, commodities were used to buy and to trade. Nowadays, they are the primary raw materials in all production stages and several developing countries are highly dependent on them. For example, the

* Corresponding author at: Centro de Investigaciones Ópticas (CONICET La Plata - CIC), C.C. 3, 1897 Gonnet, Argentina.
E-mail addresses: lucianoz@ciop.unlp.edu.ar (L. Zunino), benjamin.tabak@bcb.gov.br (B.M. Tabak), f.serinaldi@unitus.it (F. Serinaldi), massimiliano.zanin@hotmail.com (M. Zanin), dario.perez@ucv.cl (D.G. Pérez), oarosso@fibertel.com.ar (O.A. Rosso).

0378-4371/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.physa.2010.11.020

crude oil's price has reached historical values around 140 dollars in 2008. For exporter countries this increase in price is positive for the balance of payments; however, for other countries this fact can derive in an increasing overall inflation. It is also worth noting that, especially in crisis periods, commodities can be seen as a measure of value. Despite these facts, these markets have attracted much less attention within the Econophysics community than stock and currency markets, and a few previous works directly related to the commodity analysis by using physical concepts and tools can be mentioned [1–10].

Similar to stocks and currencies, commodities were initially modeled as a geometric Brownian motion. Changes of their prices would be, therefore, random and unpredictable. However, deviations from this model have been found in many empirical studies since the revolutionary papers of Benoit Mandelbrot on the evolution of cotton and wheat prices [11,12]. The existence of autocorrelation between distant observations breaks the market efficiency because past prices can help to predict future prices, i.e., correlated markets allow for arbitrage opportunities. The Hurst exponent has been widely estimated to determine whether stock prices, stock indices and currency exchange rates exhibit long-range correlations. Without being exhaustive we can mention Refs. [13–23] related to the use of Hurst exponent to measure the strength of long-range dependence and, consequently, to the stock market inefficiency quantification. It is also interesting to point out that a link between the local time-dependent Hurst exponent and the appearance of crashes on the financial markets has been shown [24,25]. Power and Turvey [6] have recently estimated this long-memory parameter in the daily volatility of future prices for 14 agricultural and energy commodities. By using a wavelet-based analysis they have found that the geometric Brownian motion should be rejected for all commodities in favor of long-range dependence with $H > 0.5$. In addition they reject, for most commodities, the null hypothesis of a stationary Hurst exponent. However, Bassler et al. [26] have shown that the estimation of this parameter alone cannot be used to determine either the existence of long-term memory or the efficiency of markets, finding that Hurst exponents $H \neq 1/2$ are perfectly consistent with Markov processes. Therefore, it is concluded that the Hurst exponent, taken alone, may be misleading regarding long time correlations. Taking also into account that Hurst exponent estimations are strongly affected by the presence of heavy tails [27], we conclude that the use of this parameter to quantify the efficiency in financial time series should be considered with caution.

It is clear that commodity markets have some particular features. Most of them represent physical products needed for some purpose that require storage and transportation. What is more important, commodities can exhibit a slower response to change in demand because their prices depend on the supply. Matia et al. [2] have conjectured that this latter feature is the main reason behind the broader multifractal spectrum of the price fluctuations of commodities compared to stock markets. Commodities respond slower than stocks to demand changes, and higher-order correlations are introduced.[1] In the same work it is also shown that the Hurst exponent is not able to provide information regarding the clustering observed in commodity returns.

To the best of our knowledge, little is known concerning the efficiency of commodity markets. In this paper we try to fill this gap by using a novel permutation information theory approach. The complexity-entropy causality plane has been recently shown to be a practical way to discriminate linear and nonlinear correlations present in financial time series [29]. The location in the complexity–entropy causality plane allows us to quantify the efficiency of each one of the commodity markets under study because the presence of temporal patterns derives in deviations from the ideal position associated to a totally random process. The null hypothesis is that commodity prices are not predictable from their past values. Large entropy and low complexity values are associated with this hypothesis. Consequently, the distance to this random ideal location can be used to define a ranking of efficiency. In addition, by analyzing the locations of these estimated permutation quantifiers it is possible to extract very useful information about the underlying (stochastic or chaotic) nature of the financial prices under analysis [30]. This is of great importance for modeling and forecasting purposes.

The remainder of the paper is organized as follows. In the following section, in order to keep our description as self-contained as possible, we describe the permutation information theory quantifiers employed to analyze the commodity data. The data sets considered are detailed in Section 3. In Section 4 we present the empirical results obtained for the different commodity markets under consideration. Finally, we summarize the findings of this paper in Section 5.

2. Permutation information theory quantifiers

2.1. Shannon entropy and statistical complexity

Tools derived from Information Theory can be very useful for the analysis of financial data. For example, the concept of entropy is able to capture the uncertainty and disorder of the time series regardless of the empirical probability distribution evidenced by the data [31]; moreover, it is a function of many moments of the probability distribution, and thus is considered a consistent alternative to the standard deviation for assessing stock market volatility [32]. Shannon entropy is very often used as the first natural entropy measure. Given any arbitrary probability distribution $P = \{p_i : i = 1, \ldots, M\}$, the widely known Shannon's logarithmic information measure, $S[P] = -\sum_{i=1}^{M} p_i \ln p_i$, is related to the uncertainty associated with the physical process described by P. If $S[P] = 0$ we are in position to predict with complete certainty which of the possible outcomes i whose probabilities are given by p_i will actually take place. In this case, our knowledge of the underlying process described by the probability distribution is maximal; on the contrary, our knowledge is minimal for a uniform distribution.

[1] Supply and demand with different elasticities can be the source of these correlations [28].

It is widely known that an entropy measure does not quantify the degree of structure or patterns present in a process [33]. This is why we have proposed to consider also the statistical complexity for the analysis of financial time series [29]. The statistical complexity of a system is defined as zero in the opposite extreme situations of perfect order and maximal randomness (a periodic sequence and a fair coin toss, for example). The former situation is fully predictable and the latter one has a very simple statistical description. At a given distance from these extremes, a wide range of possible degrees of physical structure exists, that should be quantified by the statistical complexity measure. Lamberti et al. [34] introduced an effective statistical complexity measure (SCM) that is able to detect essential details of the dynamics and differentiate different degrees of periodicity and chaos. This statistical complexity measure is defined, following the intuitive notion advanced by López-Ruiz et al. [35], through the product

$$\mathcal{C}_{JS}[P] = \mathcal{Q}_J[P, P_e]\, \mathcal{H}_S[P] \tag{1}$$

of the normalized Shannon entropy

$$\mathcal{H}_S[P] = S[P]/S_{\max} \tag{2}$$

with $S_{\max} = S[P_e] = \ln M$, $(0 \leq \mathcal{H}_S \leq 1)$ and $P_e = \{1/M, \ldots, 1/M\}$ the uniform distribution, and the disequilibrium $\mathcal{Q}_J$ defined in terms of the extensive (in the thermodynamical sense) Jensen–Shannon divergence. That is, $\mathcal{Q}_J[P, P_e] = \mathcal{Q}_0 \mathcal{J}[P, P_e]$ with $\mathcal{J}[P, P_e] = \{S[(P + P_e)/2] - S[P]/2 - S[P_e]/2\}$ the above-mentioned Jensen–Shannon divergence and $\mathcal{Q}_0$ a normalization constant, equal to the inverse of the maximum possible value of $\mathcal{J}[P, P_e]$. This value is obtained when one of the components of P, say p_m, is equal to one and the remaining p_i are equal to zero. The Jensen–Shannon divergence, that quantifies the difference between two (or more) probability distributions, is especially useful to compare the symbol composition between different sequences [36]. We stress the fact that the above SCM is not a trivial function of the entropy because it depends on two different probability distributions, the one associated to the system under analysis, P, and the uniform distribution, P_e. Furthermore, it was shown that for a given $\mathcal{H}_S$ value, there exists a range of possible SCM values [37]. Thus, it is clear that important additional information related to the correlational structure between the components of the system and the emergence of nontrivial collective behavior is provided by evaluating the statistical complexity [38,39]. Of course there exist many other complexity measures. For a comparison among them see the paper by Wackerbauer et al. [40].

2.2. Bandt and Pompe symbolization method

In order to evaluate the two above-mentioned quantifiers, $\mathcal{H}_S$ and $\mathcal{C}_{JS}$, an associated probability distribution should be constructed in advance. The adequate choice of the probability distribution associated to a time series is a crucial step for obtaining a successful characterization of the system. Bandt and Pompe [41] introduced a simple and robust method to evaluate the probability distribution taking into account the time causality of the system dynamics. They suggested that the symbol sequence should arise naturally from the time series, without any model assumptions. Thus, they took partitions by comparing the order of neighboring values rather than partitioning the amplitude into different levels. That is, given a time series $\{x_t, t = 1, \ldots, N\}$, an embedding dimension $D > 1$ ($D \in \mathbb{N}$), and an embedding delay time τ ($\tau \in \mathbb{N}$), the ordinal pattern of order D generated by

$$s \mapsto \left(x_{s-(D-1)\tau}, x_{s-(D-2)\tau}, \ldots, x_{s-\tau}, x_s\right) \tag{3}$$

has to be considered. To each time s we assign a D-dimensional vector that results from the evaluation of the time series at times $s - (D - 1)\tau, \ldots, s - \tau, s$. Clearly, the higher the value of D, the more information about the past is incorporated into the ensuing vectors. By the ordinal pattern of order D related to the time s we mean the permutation $\pi = (r_0, r_1, \ldots, r_{D-1})$ of $(0, 1, \ldots, D - 1)$ defined by

$$x_{s-r_0\tau} \geq x_{s-r_1\tau} \geq \cdots \geq x_{s-r_{D-2}\tau} \geq x_{s-r_{D-1}\tau}. \tag{4}$$

In this way the vector defined by Eq. (3) is converted into a unique symbol π. The procedure can be better illustrated with a simple example; let us assume that we start with the fictional time series depicted in Fig. 1 and we set the embedding dimension $D = 4$, $s = 20$ and the embedding delay $\tau = 3$. In this case the state space is divided into 4! partitions and 24 mutually exclusive permutation symbols are considered. The first 4-dimensional vector is $(0.1, 0.4, 0.5, 0.35)$. According to Eq. (3) this vector corresponds with $(x_{s-3\tau}, x_{s-2\tau}, x_{s-\tau}, x_s)$, and following Eq. (4) we find that $x_{s-\tau} \geq x_{s-2\tau} \geq x_s \geq x_{s-3\tau}$. Then, the ordinal pattern which allows us to fulfill Eq. (4) will be $(1, 2, 0, 3)$. The next 4-dimensional vector is $(0.25, 0.6, 0.7, 0.2)$, and $(1, 2, 3, 0)$ will be its associated permutation, and so on. For all the $D!$ possible orderings (permutations) π_i of order D, their associated relative frequencies can be naturally computed by the number of times this particular order sequence is found in the time series divided by the total number of sequences. Thus, an ordinal pattern probability distribution $P = \{p(\pi_i), i = 1, \ldots, D!\}$ is obtained from the time series. It is clear that with this ordinal time series analysis details of the original amplitude information are lost. However, a meaningful reduction of the complex systems to their basic intrinsic structure is provided. This way of symbolizing time series, based on a comparison of consecutive points, allows a more accurate empirical reconstruction of the underlying phase space of chaotic time series affected by weak (observational and dynamical) noise [41]. Furthermore, ordinal pattern distribution is invariant with respect to nonlinear monotonous transformations. Thus, nonlinear drifts or scalings artificially introduced by a measurement

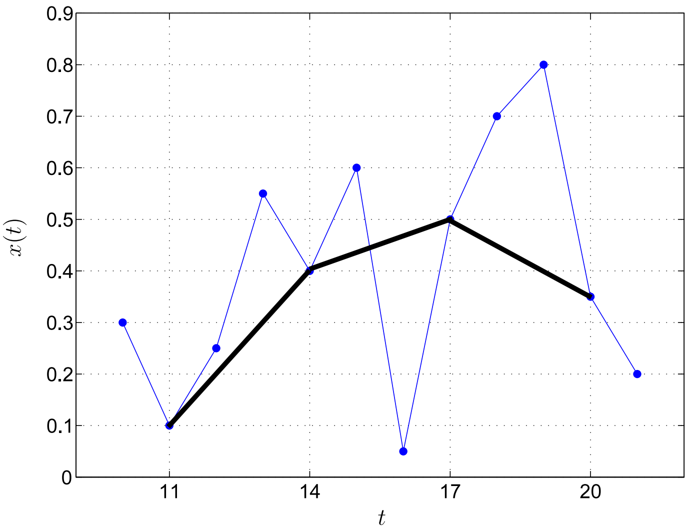

Fig. 1. (Color online) Procedure to identify ordinal patterns from a fictional time series. In this particular example embedding dimension $D = 4$, embedding delay $\tau = 3$ and time $s = 20$ are considered.

device do not modify the quantifiers' estimations, a property highly desired for the analysis of experimental data. These are the main advantages with respect to more conventional methods based on range partitioning. The probability distribution P is obtained once we fix the embedding dimension D and the embedding delay time τ. The former parameter plays an important role for the evaluation of the appropriate probability distribution, since D determines the number of accessible states, given by $D!$. Moreover, it was established that the length N of the time series must satisfy the condition $N \gg D!$ in order to obtain a reliable statistics [42]. With respect to the selection of the other parameter, Bandt and Pompe specifically considered an embedding delay $\tau = 1$ in their cornerstone paper [41]. Nevertheless, it is clear that other values of τ could provide additional information. It has been recently shown that this parameter is strongly related, if it is relevant, with the intrinsic time scales of the system under analysis [43,44].

In this work we evaluate the normalized Shannon entropy, $\mathcal{H}_S$ (Eq. (2)), and the SCM, $\mathcal{C}_{JS}$ (Eq. (1)), using the permutation probability distribution, $P = \{p(\pi_i), i = 1, \ldots, D!\}$. Defined in this way, the former quantifier is called *permutation entropy* and the latter *permutation statistical complexity*. These symbolic quantifiers were shown to be particularly useful for different purposes like characterizing stochastic processes [45,46], detecting noise-induced temporal correlations in stochastic resonance phenomena [47], measuring the stock market inefficiency [48], quantifying the randomness of chaotic pseudo-random number generators [49], and characterizing the complexity of low-frequency fluctuations in semiconductor lasers with optical feedback [50].

2.3. Complexity–entropy causality plane

In several situations it is important to analyze the time evolution of the SCM. The second law of thermodynamics states that the entropy of an isolated system grows monotonically with time until it reaches its equilibrium state.[2] Thus, $\mathcal{H}_S$ can be regarded as an arrow of time and a diagram of $\mathcal{C}_{JS}$ versus $\mathcal{H}_S$ can be employed for that purpose. It has been successfully used to study changes in a system dynamics originated by modifications of some characteristic parameters [34,35,51,52]. The *complexity–entropy causality plane* is the plane obtained with the permutation entropy of the system in the horizontal axis and the permutation statistical complexity in the vertical one. The term *causality* takes into consideration that the temporal correlation between successive samples is included in the permutation probability distribution used to estimate both information theory quantifiers. This representation space is particularly useful to discriminate between chaotic systems and stochastic processes, locating them at different planar positions [30]. More importantly within the econophysics framework, it has been recently shown that this statistical approach is an effective tool for distinguishing the stage of stock market development, allowing a more refined classification of their dynamics [29]. Emergent and developed stock markets can be discriminated with this statistical tool because it is shown that the former have lower entropy and higher complexity values revealing the presence of significant time correlations and some degree of order. Besides, the influence of linear and nonlinear correlations can be unveiled by employing surrogate tests (time and phase-randomized data). See Ref. [29] for further details. We conjecture that this permutation information tool can be also useful for detecting and quantifying the presence of correlations and hidden structures in the temporal evolution of commodity markets.

[2] It should be stressed that the premise of an isolated system can hardly be accepted in the case of financial systems.

Table 1
Commodities for which we analyzed future prices.

Name	Code	Sector
Aluminum	AL	Metal
Cocoa	CC	Agriculture
Coffee	CF	Agriculture
Copper	CO	Metal
Corn	CN	Agriculture
Cotton	CT	Agriculture
Crude oil	CR	Energy
Gold	GO	Metal
Heating oil	HO	Energy
Lean hogs	LH	Agriculture
Live cattle	LC	Agriculture
Natural gas	NG	Energy
Nickel	NI	Metal
Silver	SI	Metal
Soyabean oil	SO	Agriculture
Soyabean	SY	Agriculture
Sugar	SU	Agriculture
Unleaded gas	UG	Energy
Wheat	WH	Agriculture
Zinc	ZI	Metal

3. Data

In this paper we have employed the commodity Dow Jones UBS subindexes (http://www.djindexes.com/commodity), which are composed of commodities traded on US exchanges, with the exception of aluminum, nickel and zinc, which are traded on the London Metal Exchange (LME). These are benchmark indexes, composed of future contracts on physical commodities, which provide a good approximation for the behavior of commodity prices worldwide. We investigate 20 different commodities using a daily recorded database from January 2, 1991 to September 1, 2009. Thus, 4673 observations for future prices denominated in US dollars were considered. Table 1 details the code and sector of commodities analyzed in this paper. A similar database was recently used to characterize the topology and taxonomy of the commodity network [7]. In what follows we will evaluate the complexity–entropy causality plane location of these commodity futures daily prices.

Most research on financial markets focus on price returns, i.e. the forward change of the logarithm of the price at successive times separated by a fixed time interval, because they are stationary. However, it is clear that prices and returns contain the same information about long-range correlations [53]. In this work we have analyzed daily prices; clearly, these time series are non-stationary. Permutation quantifiers involved in this work can be applied to processes with stationary increments.[3] Indeed, the Bandt and Pompe scheme is preferred when the observed data are not completely stationary or when changes in time are more significant than absolute values [55]. They have been used to characterize the fractional Brownian motion, a widely known non-stationary process—see Refs. [30,45,46,56] for further details. Moreover, we have previously shown that in the particular case of the normalized permutation entropy and in order to discriminate time series, better results are obtained for prices than log-returns [48].

4. Empirical results

In order to estimate the permutation entropy, $\mathcal{H}_S$, and permutation statistical complexity, $\mathcal{C}_{JS}$, it is necessary to fix previously the embedding dimension D and embedding delay τ. The condition $N - (D-1)\tau \gg D!$, with N the length of the time series under analysis, should be satisfied for reliable statistics. Therefore, taking into account that $N = 4673$ for the commodity time series under study in this paper, the embedding dimension should be, at most, equal to 6. In the following sections we perform the analysis for different embedding dimensions ($D = 4$, 5 and 6) and for different embedding delays ($1 \le \tau \le 500$, $\tau \in \mathbb{N}$). Finally, in Section 4.3, we analyze how the permutation quantifiers evolve with time by implementing a rolling sample approach [16,57]. For the computation of the Bandt and Pompe probability distribution we follow the very efficient algorithm described by Keller and Sinn in Ref. [58].

4.1. Analysis for different embedding dimensions

In Fig. 2 we have plotted the location of the commodity markets in the complexity–entropy causality plane for different embedding dimension $D = 4$, $D = 5$ and $D = 6$, and time delay $\tau = 1$. Commodities are labeled by the corresponding codes listed in Table 1. It is worth noting that the position in this representation space is not directly dependent on the

[3] The distribution of ordinal patterns is time-invariant for processes with stationary increments. Consequently, unbiased estimators of the ordinal pattern probabilities are obtained by their corresponding relative frequencies [54].

Table 2
Ranking of efficiency for the commodity markets with different embedding dimensions and embedding delay $\tau = 1$.

Position	$D = 4$	$D = 5$	$D = 6$
1.	Silver	Silver	Silver
2.	Copper	Cotton	Cotton
3.	Cotton	Copper	Sugar
4.	Soybeans	Sugar	Copper
5.	Cocoa	Cocoa	Cocoa
6.	Sugar	Live cattle	Live cattle
7.	Lean hogs	Soybeans	Soybeans
8.	Live cattle	Lean hogs	Coffee
9.	Coffee	Coffee	Lean hogs
10.	Gold	Gold	Gold
11.	Aluminum	Heating oil	Heating oil
12.	Heating oil	Unleaded gas	Unleaded gas
13.	Unleaded gas	Aluminum	Zinc
14.	Zinc	Zinc	Nickel
15.	Crude oil	Crude oil	Aluminum
16.	Natural gas	Natural gas	Natural gas
17.	Soybean oil	Nickel	Crude oil
18.	Nickel	Soybean oil	Soybean oil
19.	Wheat	Wheat	Wheat
20.	Corn	Corn	Corn

commodity sector. However, we have found that energy commodities are located together in an intermediate position, with intermediate entropy and complexity values. The Euclidean distance to the random ideal vertex, i.e. $\mathcal{H}_S = 1$ and $\mathcal{C}_{JS} = 0$, quantifies the inefficiency of these markets. By inefficiency we mean that the time series are not completely random in the Fama's sense [59]. The goal of any market is to approach the ideal $(1, 0)$-point in this plane, as closely as possible, since there randomization is optimal. Based on this fact we can define a ranking of efficiency for the commodity markets, which are detailed in Table 2. It can be concluded that there are only small changes in the order derived for different values of D and, afterward, our approach is practically independent of the embedding dimension. Silver and corn are, respectively, the most and the least efficient ones, regardless the embedding dimension value. Taking into account that by increasing the embedding dimension we are increasing the length and the number of symbols, it is reasonable to assume that with the largest possible value, in our case $D = 6$, a better characterization can be achieved. Thus, we fix $D = 6$ for the next estimations.

We have also analyzed the location in the complexity–entropy causality plane for the shuffled commodity prices. In the shuffling procedure the data are put into random order and all non-trivial temporal correlations are destroyed. As it can be seen in Fig. 3 the estimated values of $\mathcal{H}_S$ and $\mathcal{C}_{JS}$ for the shuffled data are very close to the random ideal ones ($\mathcal{H}_S \approx 0.988$ and $\mathcal{C}_{JS} \approx 0.029$). We verify in this way that the positions obtained from original data are not obtained by chance. Moreover, the underlying correlations are significant and play a starring role in the commodity price formation. A similar result was found in the analysis of stock markets [29].

It is important to note that we have only considered four, five and six consecutive days to build the ordinal patterns. However, the information related with the long-range correlations is provided by the permutation probability distribution associated to each commodity market. Differences in the long-range correlations of time series translate to differences in their associated probability distribution. Moreover, it was shown theoretically and through numerical simulations that $\mathcal{H}_S$ and $\mathcal{C}_{JS}$ are able to distinguish stochastic processes with different long-range correlations, like the fractional Gaussian noise (fGn) and fractional Brownian motion (fBm), for embedding dimensions $D = 3$, 4, 5 and 6 [30,45,46].

Taking into account that fBm is widely used in the Econophysics community to model the dynamics of financial systems, we have compared the locations of numerical realizations of this stochastic process with the positions obtained for the original commodity markets. One hundred independent numerical realizations of length $N = 4673$ for Hurst exponent $H \in \{0.4, 0.45, 0.5, 0.55, 0.6\}$ were simulated, each series starting at a different initial condition. The method of Wood and Chan, which is both exact and fast [60], was adopted for the numerical simulations. This algorithm simulates fGn that are cumulated to obtain fBm. From Fig. 4 it can be concluded that commodity and fBm locations are very close in the complexity–entropy causality plane. Indeed, the commodities under analysis in this work are located in the same position than fBm numerical simulations with the Hurst exponent in the range $[0.45, 0.55]$. According to these results we can conclude that commodity markets and fBm share similar dynamical properties. This finding is relevant for modeling and forecasting purposes. Moreover, comparing with Fig. 1 of Ref. [30], it is possible to affirm that the proposed deterministic chaotic nature of commodities [61] should be rejected because chaotic systems have entropies that are seen to be in the entropy region lying between 0.45 and 0.7, and located near the maximum $\mathcal{C}_{JS}$. These high complexity values are due to nonlinear structures immersed in chaotic time series.

4.2. Analysis for different embedding delays

The embedding delay τ is directly related to the sampling frequency of the system under analysis [62]. By increasing this parameter the original time series is subsampled in a very efficient way. It has been recently shown that characteristic

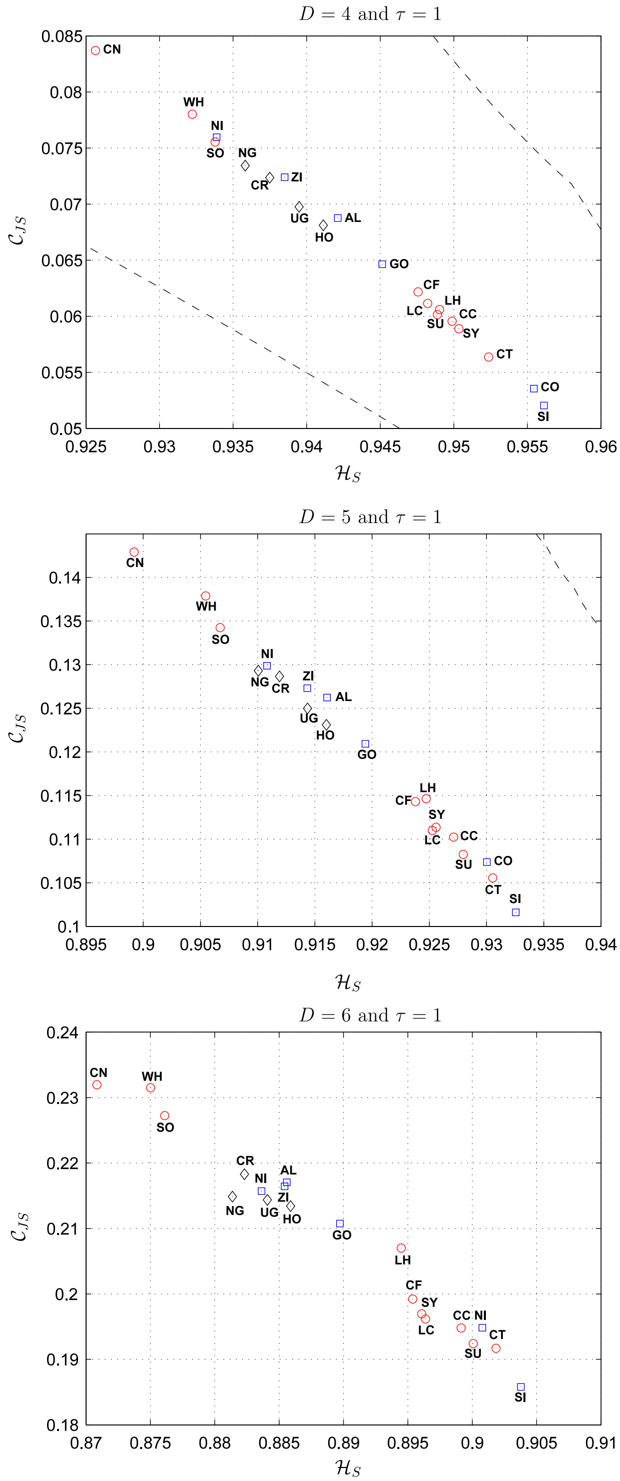

Fig. 2. (Color online) Position of the commodity markets in the complexity–entropy causality plane with embedding dimensions $D = 4$ (upper plot), $D = 5$ (central plot) and $D = 6$ (lower plot), and time delay $\tau = 1$. The shape and color is based on the commodity sector: blue squares for metals, red circles for agriculture and black diamonds for energies. The codes used in these plots are detailed in Table 1. We also display the maximum and minimum possible values of the permutation statistical complexity (segmented curves). For further details about the range of possible SCM values see Ref. [37].

time scales present in the system dynamics are detected through the presence of clear extrema of permutation entropy and permutation complexity when they are calculated as a function of the embedding delay [43,44]. Also periodicities present in the system can be identified by analyzing the behavior of these quantifiers as a function of τ [63].

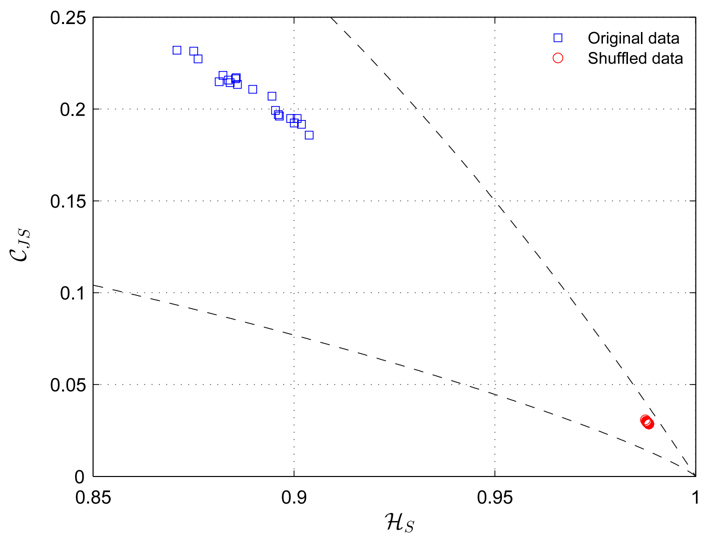

Fig. 3. (Color online) Location of the shuffled commodity markets in the complexity–entropy causality plane with embedding dimension $D = 6$ and time delay $\tau = 1$. The shuffled positions are estimated averaging over ten different realizations. The maximum and minimum possible values of the permutation statistical complexity are also shown (segmented curves).

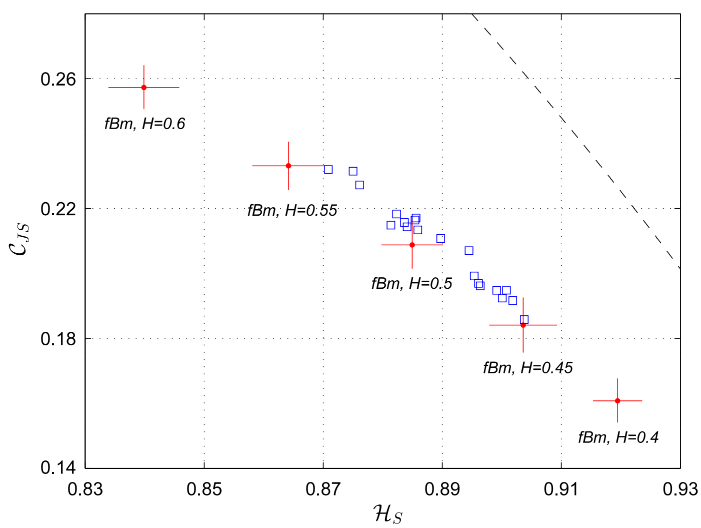

Fig. 4. (Color online) Comparison of the positions of the original commodity markets (blue squares) and numerical realizations of fBm in the complexity–entropy causality plane for embedding dimension $D = 6$ and time delay $\tau = 1$. Mean and standard deviation of both permutation quantifiers for 100 independent realizations of length $N = 4673$ for each value of $H \in \{0.4, 0.45, 0.5, 0.55, 0.6\}$ were considered. The maximum possible values of the permutation statistical complexity is also shown (segmented curves).

Fig. 5 shows the permutation quantifiers estimated from the original commodity time series for different embedding delays. More precisely, time scales associated to the interval $1 \leq \tau \leq 500$ are considered. Although the quantifiers are only estimated for discrete values, we use continuous lines in the plot because the visualization is improved in this way. $\mathcal{H}_S$ is a decreasing function of τ for all commodity markets under analysis. Thus, the randomness decreases for larger time scales. $\mathcal{C}_{JS}$ has a different behavior. This quantifier increases as a function of the embedding delay for the interval $1 \leq \tau \leq 200$. It can be seen that for larger values of τ the permutation statistical complexity saturates for most of the commodity markets. In some particular cases (corn, cotton, lean hogs and wheat) a pronounced maximum of $\mathcal{C}_{JS}$ is found for $\tau \approx 250$, close to a business year cycle. This extremum is due to the presence of patterns and hidden structures for this time scale, resulting in the probability distribution of the ordinal patterns being different from the uniform probability distribution. It is worth mentioning that this behavior is only observed in the permutation complexity analysis. Permutation entropy has a monotone behavior with τ and, consequently, this measure is not able to identify this particular time scale. The presence of this peak can be attributed to the seasonal production of these agricultural commodities. It seems to be reasonable that the harvest annual cycles play a relevant role in the price dynamics, especially if the commodities are produced mostly in specific regions of the world. The US is the major exporter of corn, cotton and wheat. On the other hand, agricultural commodities with a

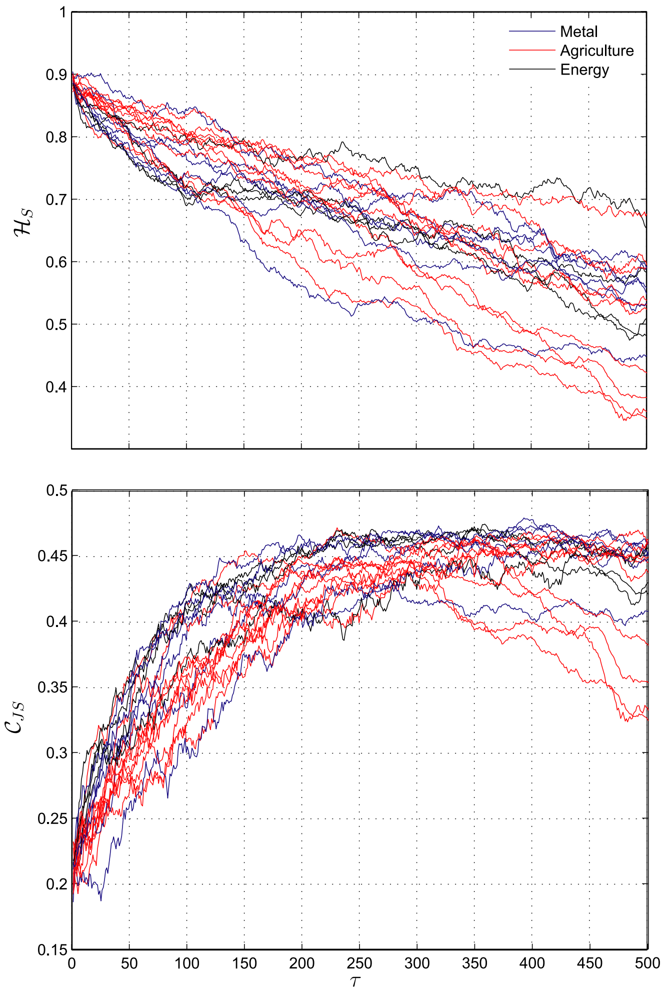

Fig. 5. (Color online) Permutation entropy $\mathcal{H}_S$ and permutation statistical complexity $\mathcal{C}_{JS}$ of commodity stock markets as a function of the embedding delay τ ($1 \leq \tau \leq 500$) for embedding dimension $D = 6$. The color is based on the commodity sector: blue for metals, red for agriculture and black for energies.

global production should not be affected. For instance, in the case of soybean, which is mainly produced and exported by the US, Brazil and Argentina, we do not find the signature of an annual cycle.

It is well known that fBm are self-similar processes.[4] As a consequence of this property it is found that the relative frequencies of the ordinal patterns do not depend on the value of the embedding delay [54,56]. Therefore, quantifiers derived from the permutation probability distribution, like permutation entropy and permutation statistical complexity, are independent of the time scale considered. From the behavior of the permutation quantifiers depicted in Fig. 5, we firstly conclude that commodity markets are not self-similar and the fBm, widely considered to model financial time series, is not suitable for commodity markets.

With the intention to numerically check this fact we have estimated both permutation quantifiers as a function of the embedding delay for numerical simulations of geometric Brownian motions (gBm), i.e. fBm with $H = 0.5$. For simulating these time series we have summed up the Gaussian white noise generated by using the function *randn* of Matlab. One hundred independent numerical realizations of length $N = 4673$ were considered in order to take into account the finite-size effect present in commodity time series and embedding dimension $D = 6$ was chosen for the evaluation of the corresponding associated permutation probability distribution. The concomitant mean values plus the corresponding standard deviations of both quantifiers, $\mathcal{H}_S$ and $\mathcal{C}_{JS}$, are plotted in Fig. 6 (upper plot) as functions of the embedding delay τ. Contrary to what was expected, permutation quantifiers are not constant functions of τ. We conjecture that this numerical

[4] Self-similar stochastic processes are invariant in distribution under suitable scaling of time and space. Formally, a (stochastic) process $X(t)$ is self-similar with index H (also named the Hurst exponent) if, for any positive stretching factor c, $X(t) \stackrel{d}{=} c^H X(c^{-1}t)$, where $\stackrel{d}{=}$ means equality in distribution [64]. These processes are of considerable importance because they appear in a natural way from limit theorems for sums of random variables [65].

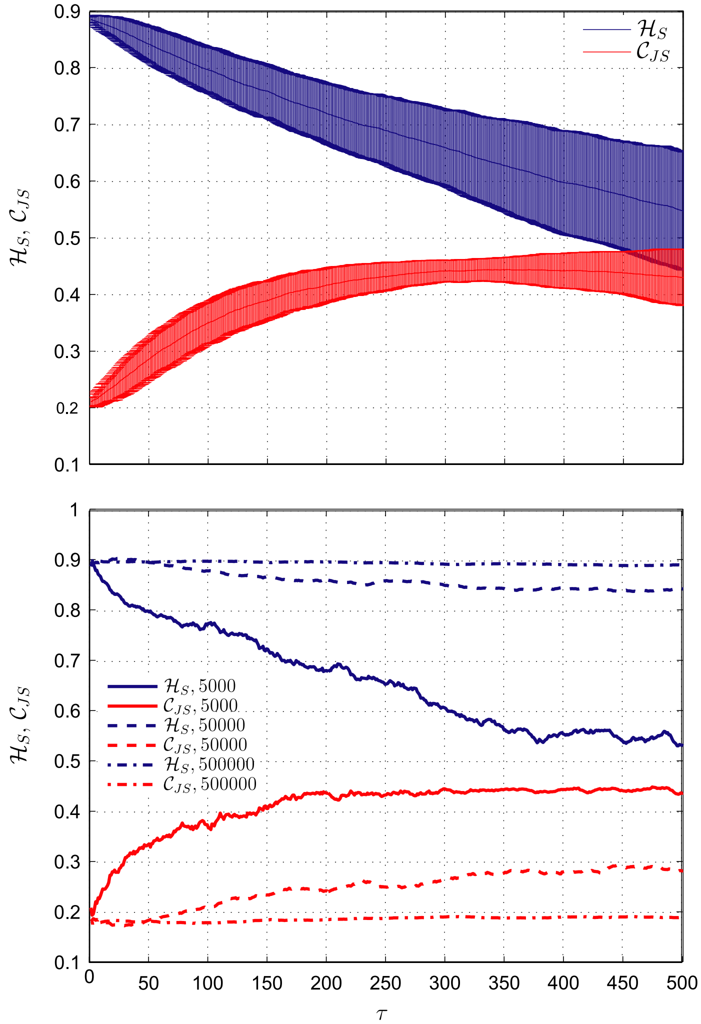

Fig. 6. (Color online) Permutation entropy $\mathcal{H}_S$ and permutation statistical complexity $\mathcal{C}_{JS}$ of gBm simulations as a function of the embedding delay τ ($1 \leq \tau \leq 500$) for embedding dimension $D = 6$. Top: Mean and standard deviation of both permutation quantifiers for 100 independent realizations of length $N = 4673$. Bottom: Estimated quantifiers for simulations with $N = 5000$, 50 000 and 500 000 data points.

behavior is due to the finite-size of the realizations. By increasing the length of Bm simulations the theoretical finding is verified as it can be concluded from the lower plot in Fig. 6. Consequently, self-similarity cannot be discarded in the commodity time series because very long traces are necessary to verify the theoretical expected behavior.

We have estimated the Hurst exponent associated to the commodity prices in order to compare it with the results obtained by employing our permutation information theory approach. Three well known and widely used techniques were implemented in R [66,67] to estimate the Hurst exponent: the Higuchi method, the detrended fluctuation analysis (DFA) and the detrended moving average (DMA) analysis. It has been recently shown that these methods are reliable for high non-stationary time series [53,65]. Since a description of these graphical estimators is beyond the scope of the present work, the reader is referred to Refs. [57,68–70] for a more comprehensive survey. The results obtained for the 20 different commodity daily prices are summarized in Table 3. All methods were applied between 10 and 300 time step windows, in order to have a sufficient number of points in small windows and an appropriate number of windows for large scales. Hurst exponent estimations are performed on the range of scales where the log–log plots are approximately linear. As the three methods are applied to the same range of scales, they are eventually affected by the same problem. It is worth mentioning that according to these Hurst exponent estimations only cocoa, copper, crude oil, gold, heating oil, nickel, silver and zinc can be considered outside the confidence intervals $(0.45, 0.55)$ usually associated to a geometric Brownian motion [3]. The rest of commodities appear to be strongly or marginally consistent with this totally random stochastic process. In Fig. 7 we plot the Pearson's correlation coefficient ρ between the Hurst exponent and our inefficiency measure, i.e. the distance to the random ideal vertex in the complexity–entropy causality plane, as a function of the embedding delay τ.[5] The correlation

[5] Similar results are obtained for the Kendall and Spearman rank correlation coefficients.

Table 3
Hurst exponents estimated for the commodity daily prices.

Commodity	Higuchi	DFA	DMA
Aluminum	0.545	0.503	0.576
Cocoa	0.412	0.458	0.457
Coffee	0.476	0.455	0.494
Copper	0.576	0.570	0.608
Corn	0.541	0.517	0.576
Cotton	0.525	0.474	0.562
Crude oil	0.574	0.549	0.655
Gold	0.443	0.443	0.442
Heating oil	0.572	0.536	0.631
Lean hogs	0.507	0.483	0.544
Live cattle	0.491	0.487	0.518
Natural gas	0.551	0.538	0.534
Nickel	0.580	0.547	0.623
Silver	0.423	0.461	0.448
Soyabean oil	0.503	0.489	0.549
Soyabean	0.488	0.501	0.525
Sugar	0.490	0.494	0.548
Unleaded gas	0.523	0.545	0.567
Wheat	0.509	0.512	0.524
Zinc	0.544	0.469	0.578

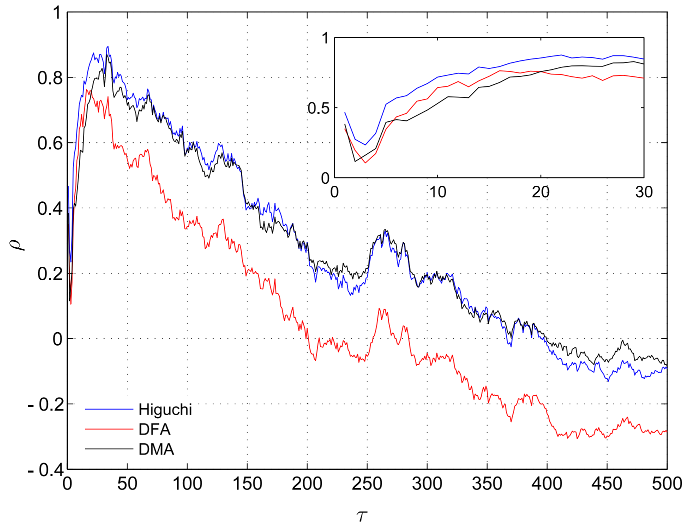

Fig. 7. (Color online) Pearson's correlation coefficient ρ between the Hurst exponent estimations and the distance to the random ideal location in the complexity–entropy causality plane as a function of the embedding delay τ. Embedding dimension $D = 6$ was used to estimate the permutation quantifiers. Three different graphical estimators were considered for the Hurst exponent estimation: the Higuchi method, DFA and DMA analysis. Inset: Enlargement of the correlation coefficient in the interval $1 \leq \tau \leq 30$.

coefficients for the original time series, i.e. with $\tau = 1$, are equal to 0.4673, 0.3503 and 0.3848 for the Higuchi, DFA and DMA methodologies, respectively. Then, these positive correlations decrease for the three estimators, reaching a minimum value for $\tau = 2$ in the case of DMA and for $\tau = 3$ in the Higuchi and DFA cases (see the inset in Fig. 7). An absolute maximum of the correlation is found for an embedding delay $\tau \approx 25$ in the case of the Higuchi and DMA methods, and around $\tau \approx 15$ for DFA. After that the correlations decrease. These high correlations found between the Hurst exponent estimations and our approach are particularly curious and their reasons will be the scope of a future study.

4.3. Time evolution of the quantifiers

We have analyzed the time evolution of the quantifiers in order to see how the inefficiency of commodities is changing in time. Both permutation quantifiers are estimated on the ensemble of points obtained from the intersection of the time series and a sliding window of size $N_s = 1000$ (around four business years). Then, the time window is rolled $\delta_s = 20$ points forward (close to a business month) eliminating the first δ_s observations and including the next ones, and the quantifiers are re-estimated. This procedure is repeated until the end of the time series. Fig. 8 depicts the evolution of the locations of silver and corn in the complexity–entropy causality plane with $D = 5$ and $\tau = 1$. We can observe that the position of these commodities is changing with time. In particular, silver is moving to lower entropy values and higher complexity

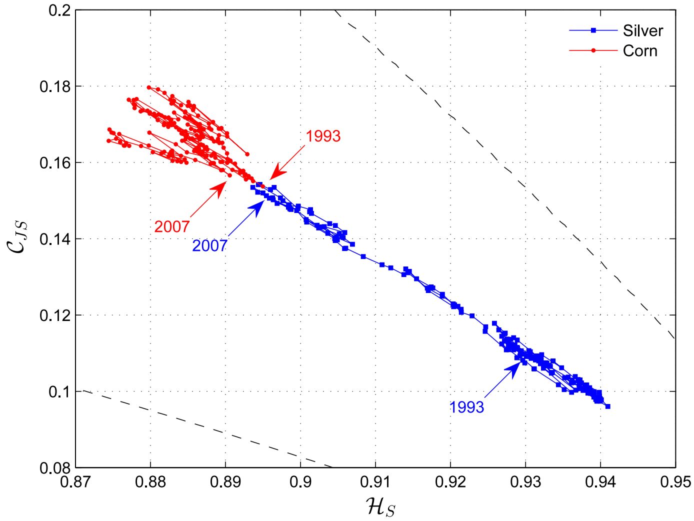

Fig. 8. (Color online) Time evolution of positions of silver (blue squares curve) and corn (red circles curve) in the complexity–entropy causality plane by employing a rolling sample approach with a sliding window of size $N_s = 1000$, step $\delta_s = 20$, embedding dimension $D = 5$ and embedding delay $\tau = 1$. The maximum and minimum possible values of the permutation statistical complexity are also shown (segmented curves).

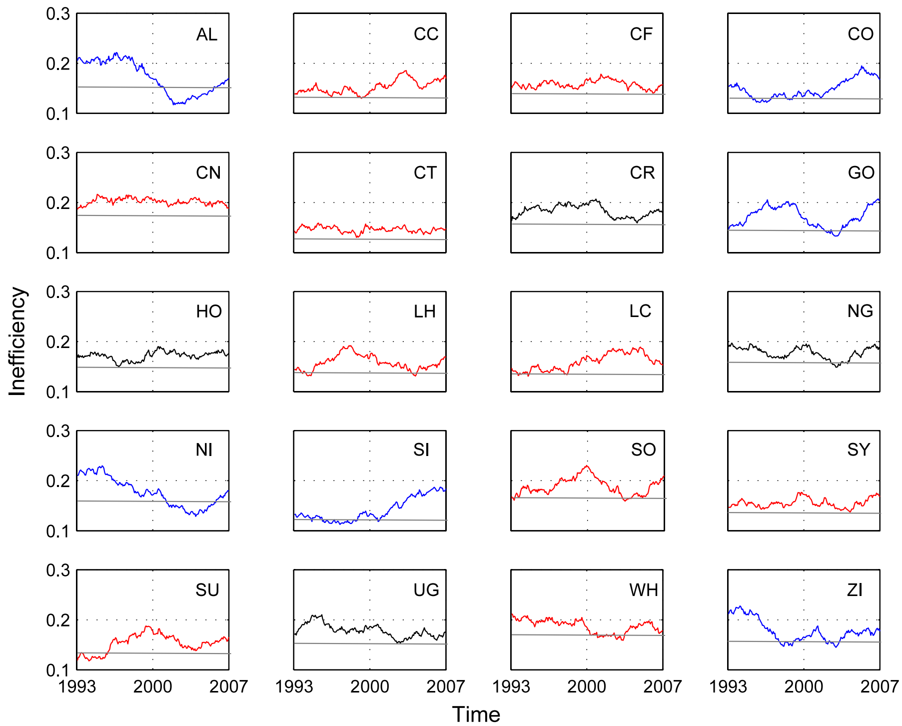

Fig. 9. (Color online) Inefficiency measure for the commodity daily prices as a function of time. A rolling sample approach with a sliding window of size $N_s = 1000$, step $\delta_s = 20$, embedding dimension $D = 5$ and embedding delay $\tau = 1$ was implemented. The codes used in these plots are detailed in Table 1 and the color is based on commodity sector: blue for metals, red for agriculture and black for energies. The continuous gray line corresponds to the global inefficiency measure considering the time series as a whole.

ones. Consequently, it is more inefficient with time. In Fig. 9 we have plotted the inefficiency measure, i.e. the Euclidean distance to the random ideal point (1, 0) of the representation space, for the twenty commodities as a function of time. This number is an indicator of the predictability degree of the system. When this number decreases the evolution of the time series is more random and less predictable. Contrarily, when this number increases, randomness is decreasing and the underlying dynamics would be more predictable. We observe that the inefficiency of commodity markets is changing in time with periods of increasing and decreasing randomness.

5. Conclusions

The study of commodity predictability is of great importance and interest. In this paper we address this relevant issue by using the complexity–entropy causality plane. This diagnostic tool measures the presence of patterns, and consequently the long-range dependence of a temporal trace, in a parameter-free way by combining the information provided by two complementary quantifiers: permutation entropy and permutation statistical complexity. Linear and nonlinear correlations are considered by applying this novel approach. Moreover, the difference-based symbolization methodology implemented increases the efficiency of unveiling hidden structures in financial time series and reduces notably the sensitivity to noise. Consequently, our approach can be considered of more general applicability than other widely used alternatives like the Hurst exponent. In fact, a multifractal behavior was found for commodities in Ref. [2]. Different moments of the variable's distribution are associated with different scaling laws, and the 2-point correlations is not sufficient to uncover the clustering observed in commodity returns. Furthermore, a time-varying Hurst coefficient has been recently accounted for 9 of 14 commodities in Ref. [6].

The complexity–entropy causality plane location allows us to derive an associated ranking of efficiency. The presence of patterns in the temporal evolution translates into deviations from the ideal position associated with a totally random process. Thus, the distance to this random ideal location ($\mathcal{H}_S = 1$ and $\mathcal{C}_{JS} = 0$) is a way to quantify the inefficiency of the market under analysis. It is worth mentioning that only small changes are observed in the commodity rankings obtained with different embedding dimensions. According to the empirical results obtained in this work the efficiency is not related to the commodity sector. We have found that highly demanded commodities (corn, crude oil, natural gas and wheat) are less efficient. Thus, we conjecture that significant correlations introduced by the increasing demand can be the main source of this inefficiency.

We have also studied the temporal evolution of commodity prices in the complexity–entropy causality plane. This analysis allows us to identify periods of increasing and decreasing randomness (lower and higher predictability, respectively) in the commodity dynamics. The information extracted can be very valuable for policy makers and regulatory authorities for investigating market bubbles or manipulation.

Matia et al. [1] have shown that commodity spot and future prices have different scaling behaviors. More specifically, power-law probability distributions with larger exponent values are found for commodity future prices, indicating the presence of smaller fluctuations. Taking into account this relevant difference, it will be interesting to compare the complexity–entropy causality plane locations of commodity spot and future prices in a further research. More recently, Romero et al. [5] have provided empirical evidence of very different temporal organization of ancient Babylon and medieval English commodity prices compared to commodities traded on contemporary markets. Strong persistent correlations are found in both Babylon and England commodity prices. Our approach can be useful to follow the evolution of commodity dynamics in different historical periods.

It is clear that the Bandt and Pompe symbolization methodology plays a key role in our approach. In order to better understand how relevant it is, a comparison with the complexity–entropy plane locations obtained by using other ways to estimate the probability distribution associated to financial time series will be considered in the future. We conjecture that the probability density functions related to the length, duration and area of the clusters delimited by two consecutive intersections between the time series x_t and the average of the signal over a number of points [71,72], could be particularly helpful for distinguishing long-range correlated time series.

Finally, it should be remarked that the results obtained in this work confirm the usefulness of our permutation information theory approach for detecting and quantifying the presence of correlations and hidden structures, highlighting its relevance and encouraging its application to real time series of other scientific fields.

Acknowledgements

The authors would like to thank C.G. Turvey (Cornell University, US) and three anonymous reviewers for their useful comments and suggestions. Luciano Zunino and Osvaldo A. Rosso were supported by Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina. Benjamin M. Tabak gratefully acknowledges financial support from CNPq foundation. The opinions expressed in the paper do not necessarily reflect those of the Banco Central do Brasil. Darío G. Pérez was supported by Comisión Nacional de Investigación Científica y Tecnológica (CONICYT, FONDECYT project No. 1100753), Chile, and partially by Pontificia Universidad Católica de Valparaíso (PUCV, Project No. 123.704/2010), Chile. Osvaldo A. Rosso is PVE fellowship, CAPES, Brazil.

References

[1] K. Matia, L.A. Nunes Amaral, S.P. Goodwin, H.E. Stanley, Different scaling behaviors of commodity spot and future prices, Phys. Rev. E 66 (2002) 045103.
[2] K. Matia, Y. Ashkenazy, H.E. Stanley, Mulifractal properties of price fluctuations of stocks and commodities, Europhys. Lett. 61 (2003) 422–428.
[3] C.G. Turvey, A note on scaled variance ratio estimation of the Hurst exponent with application to agricultural commodity prices, Physica A 377 (2007) 155–165.
[4] P. Sieczka, J.A. Hołyst, Correlations in commodity markets, Physica A 388 (2009) 1621–1630.
[5] N.E. Romero, Q.D.Y. Ma, L.S. Liebovitch, C.T. Brown, P.Ch. Ivanov, Correlated walks down the Babylonian markets, Europhys. Lett. 90 (2010) 18004.
[6] G.J. Power, C.G. Turvey, Long-range dependence in the volatility of commodity futures prices: wavelet-based evidence, Physica A 389 (2010) 79–90.

[7] B.M. Tabak, T.R. Serra, D.O. Cajueiro, Topological properties of commodities networks, Eur. Phys. J. B 74 (2010) 243–249.
[8] E. Leite Siqueira Jr., T. Stošić, L. Bejan, B. Stošić, Correlations and cross-correlations in the Brazilian agrarian commodities and stocks, Physica A 389 (2010) 2739–2743.
[9] S.-P. Chen, L.-Y. He, Multifractal spectrum analysis of nonlinear dynamical mechanisms in China's agricultural futures markets, Physica A 389 (2010) 1434–1444.
[10] L.-Y. He, S.-P. Chen, Are developed and emerging agricultural futures markets multifractal? A comparative perspective, Physica A 389 (2010) 3828–3836.
[11] B. Mandelbrot, The variation of certain speculative prices, J. Business 36 (1963) 394–419.
[12] B.B. Mandelbrot, The variation of the prices of cotton, wheat, and railroad stocks, and of some financial rates, J. Business 40 (1967) 393–413.
[13] N. Vandewalle, M. Ausloos, Coherent and random sequences in financial fluctuations, Physica A 246 (1997) 454–459.
[14] P. Grau-Carles, Empirical evidence of long-range correlations in stock returns, Physica A 287 (2000) 396–404.
[15] T. Di Matteo, T. Aste, M.M. Dacorogna, Scaling behaviors in differently developed markets, Physica A 324 (2003) 183–188.
[16] D.O. Cajueiro, B.M. Tabak, The Hurst exponent over time: testing the assertion that emerging markets are becoming more efficient, Physica A 336 (2004) 521–537.
[17] D.O. Cajueiro, B.M. Tabak, Ranking efficiency for emerging markets, Chaos Solitons Fractals 22 (2004) 349–352.
[18] D.O. Cajueiro, B.M. Tabak, Ranking efficiency for emerging markets II, Chaos Solitons Fractals 23 (2005) 671–675.
[19] T. Di Matteo, T. Aste, M.M. Dacorogna, Long-term memories of developed and emerging markets: using the scaling analysis to characterize their stage of development, J. Bank. 29 (2005) 827–851.
[20] L. Zunino, B.M. Tabak, D.G. Pérez, M. Garavaglia, O.A. Rosso, Inefficiency in Latin-American market indices, Eur. Phys. J. B 60 (2007) 111–121.
[21] C. Eom, S. Choi, G. Oh, W.-S. Jung, Hurst exponent and prediction based on weak-form efficient market hypothesis of stock markets, Physica A 387 (2008) 4630–4636.
[22] C. Eom, G. Oh, W.-S. Jung, Relationship between efficiency and predictability in stock price change, Physica A 387 (2008) 5511–5517.
[23] Ł Czarnecki, D. Grech, Multifractal dynamics of stock markets, Acta Phys. Polon. A 117 (2010) 623–629.
[24] D. Grech, Z. Mazur, Can one make any crash prediction in finance using the local Hurst exponent idea? Physica A 336 (2004) 133–145.
[25] D. Grech, G. Pamuła, The local Hurst exponent of the financial time series in the vicinity of crashes on the Polish stock exchange market, Physica A 387 (2008) 4299–4308.
[26] K.E. Bassler, G.H. Gunaratne, J.L. McCauley, Markov processes, Hurst exponents, and nonlinear diffusion equations: with application to finance, Physica A 369 (2006) 343–353.
[27] J. Barunik, L. Kristoufek, On Hurst exponent estimation under heavy-tailed distributions, Physica A 389 (2010) 3844–3855.
[28] C.G. Turvey, J.R. Stokes, Market structure and the values of agricultural contingent claims, Canad. J. Agricultural Econom. 56 (2008) 79–94.
[29] L. Zunino, M. Zanin, B.M. Tabak, D.G. Pérez, O.A. Rosso, Complexity–entropy causality plane: a useful approach to quantify the stock market inefficiency, Physica A 389 (2010) 1891–1901.
[30] O.A. Rosso, H.A. Larrondo, M.T. Martín, A. Plastino, M.A. Fuentes, Distinguishing noise from chaos, Phys. Rev. Lett. 99 (2007) 154102.
[31] S.R. Bentes, R. Menezes, D.A. Mendes, Long memory and volatility clustering: is the empirical evidence consistent across stock markets? Physica A 387 (2008) 3826–3830.
[32] S.R. Bentes, Econophysics: a new discipline, 2010, arXiv:1006.3956.
[33] D.P. Feldman, J.P. Crutchfield, Measures of statistical complexity: why? Phys. Lett. A 238 (1998) 244–252.
[34] P.W. Lamberti, M.T. Martín, A. Plastino, O.A. Rosso, Intensive entropic non-triviality measure, Physica A 334 (2004) 119–131.
[35] R. López-Ruiz, H.L. Mancini, X. Calbet, A statistical measure of complexity, Phys. Lett. A 209 (1995) 321–326.
[36] I. Grosse, P. Bernaola-Galván, P. Carpena, R. Román-Roldán, J. Oliver, H.E. Stanley, Analysis of symbolic sequences using the Jensen–Shannon divergence, Phys. Rev. E 65 (2002) 041905.
[37] M.T. Martín, A. Plastino, O.A. Rosso, Generalized statistical complexity measures: geometrical and analytical properties, Physica A 369 (2006) 439–462.
[38] J.R. Sánchez, R. López-Ruiz, A method to discern complexity in two-dimensional patterns generated by coupled map lattices, Physica A 355 (2005) 633–640.
[39] M. Escalona-Morán, M.G. Cosenza, R. López-Ruiz, P. García, Statistical complexity and nontrivial collective behavior in electroencephalografic signals, Internat. J. Bifur. Chaos 20 (2010) 1723–1729.
[40] R. Wackerbauer, A. Witt, H. Atmanspacher, J. Kurths, H. Scheingraber, A comparative classification of complexity measures, Chaos Solitons Fractals 4 (1994) 133–173.
[41] C. Bandt, B. Pompe, Permutation entropy: a natural complexity measure for time series, Phys. Rev. Lett. 88 (2002) 174102.
[42] M. Staniek, K. Lehnertz, Parameter selection for permutation entropy measurements, Internat. J. Bifur. Chaos 17 (2007) 3729–3733.
[43] L. Zunino, M.C. Soriano, I. Fischer, O.A. Rosso, C.R. Mirasso, Permutation-information-theory approach to unveil delay dynamics from time-series analysis, Phys. Rev. E 82 (2010) 046212.
[44] M.C. Soriano, L. Zunino, O.A. Rosso, I. Fischer, C.R. Mirasso, Time scales of a chaotic semiconductor laser with optical feedback under the lens of a permutation information analysis, IEEE J. Quantum Electron. 47 (1) (2011), doi:10.1109/JQE.2010.2078799.
[45] O.A. Rosso, L. Zunino, D.G. Pérez, A. Figliola, H.A. Larrondo, M. Garavaglia, M.T. Martín, A. Plastino, Extracting features of Gaussian self-similar stochastic processes via the Bandt & Pompe approach, Phys. Rev. E 76 (2007) 061114.
[46] L. Zunino, D.G. Pérez, M.T. Martín, M. Garavaglia, A. Plastino, O.A. Rosso, Permutation entropy of fractional Brownian motion and fractional Gaussian noise, Phys. Lett. A 372 (2008) 4768–4774.
[47] O.A. Rosso, C. Masoller, Detecting and quantifying stochastic and coherence resonances via information-theory complexity measurements, Phys. Rev. E 79 (2009) 040106(R).
[48] L. Zunino, M. Zanin, B.M. Tabak, D.G. Pérez, O.A. Rosso, Forbidden patterns, permutation entropy and stock market inefficiency, Physica A 388 (2009) 2854–2864.
[49] L.D. Micco, H.A. Larrondo, A. Plastino, O.A. Rosso, Quantifiers for randomness of chaotic pseudo-random number generators, Phil. Trans. R. Soc. A 367 (2009) 3281–3296.
[50] J. Tiana-Alsina, M.C. Torrent, O.A. Rosso, C. Masoller, J. García-Ojalvo, Quantifying the statistical complexity of low-frequency fluctuations in semiconductor lasers with optical feedback, Phys. Rev. A 82 (2010) 013819.
[51] L. Zunino, D.G. Pérez, M.T. Martín, A. Plastino, M. Garavaglia, O.A. Rosso, Characterization of Gaussian self-similar stochastic processes using wavelet-based informational tools, Phys. Rev. E 75 (2007) 021115.
[52] O.A. Rosso, H. Craig, P. Moscato, Shakespeare and other English Renaissance authors as characterized by Information Theory complexity quantifiers, Physica A 388 (2009) 916–926.
[53] F. Serinaldi, Use and misuse of some Hurst parameter estimators applied to stationary and non-stationary financial time series, Physica A 389 (2010) 2770–2781.
[54] M. Sinn, K. Keller, Estimation of ordinal pattern probabilities in Gaussian processes with stationary increments, 2009 http://www.math.uni-luebeck.de/keller/wwwpapers/csda.pdf.
[55] C.S. Daw, C.E.A. Finney, E.R. Tracy, A review of symbolic analysis of experimental data, Rev. Sci. Instrum. 74 (2003) 915–930.
[56] C. Bandt, F. Shiha, Order patterns in time series, J. Time Ser. Anal. 28 (2007) 646–665.
[57] A. Carbone, G. Castelli, H.E. Stanley, Time-dependent Hurst exponent in financial time series, Physica A 344 (2004) 267–271.
[58] K. Keller, M. Sinn, Ordinal analysis of time series, Physica A 356 (2005) 114–120.
[59] E.F. Fama, Efficient capital markets: a review of theory and empirical work, J. Finance 25 (1970) 383–417.

[60] J.-F. Coeurjolly, Simulation and identification of the fractional Brownian motion: a bibliographical and comparative study, J. Statist. Software 5 (2000) 1–53.
[61] G.P. DeCoster, W.C. Labys, D.W. Mitchell, Evidence of chaos in commodity futures prices, J. of Future Markets 12 (1992) 291–305.
[62] X. Li, G. Ouyang, Estimating coupling direction between neuronal populations with permutation conditional mutual information, Neuroimage 52 (2010) 497–507.
[63] C. Bandt, Ordinal time series analysis, Ecol. Modelling 182 (2005) 229–238.
[64] J. Beran, Statistics for Long-Memory Processes, Chapman & Hall, 1994.
[65] F. Esposti, M. Ferrario, M.G. Signorini, A blind method for the estimation of the Hurst exponent in time series: theory and application, Chaos 18 (2008) 033126.
[66] R Development Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0 (2009). URL: http://www.R-project.org.
[67] William Constantine (Insightful Corporation) and Donald Percival (Applied Physics Laboratory and University of Washington), fractal: Insightful Fractal Time Series Modeling and Analysis, R package version 1.0-2 (2007).
[68] T. Higuchi, Approach to an irregular time series on the basis of the fractal theory, Physica D 31 (1988) 277–283.
[69] C.-K. Peng, S.V. Buldyrev, S. Havlin, M. Simons, H.E. Stanley, A.L. Goldberger, Mosaic organization of DNA nucleotides, Phys. Rev. E 49 (1994) 1685–1689.
[70] S. Arianos, A. Carbone, Detrending moving average algorithm: a closed-form approximation of the scaling law, Physica A 382 (2007) 9–15.
[71] A. Carbone, G. Castelli, H.E. Stanley, Analysis of clusters formed by the moving average of a long-range correlated time series, Phys. Rev. E 69 (2004) 026105.
[72] A. Carbone, H.E. Stanley, Scaling properties and entropy of long-range correlated time series, Physica A 384 (2007) 21–24.