Volatility, irregularity, and predictable degree of accumulative return series

Wen-Qi Duan[1,2] and H. Eugene Stanley[2]
[1]*School of Economics and Management, Zhejiang Normal University, Jinhua 321004, People's Republic of China*
[2]*Center for Polymer Studies and Department of Physics, Boston University, Boston, Massachusetts 02215, USA*
(Received 24 March 2010; published 22 June 2010)

Recently it was shown that financial time series are not completely random process but exhibit long-term or short-term dependences, which offer promises for predictability. However, we do not clearly understand the potential relationship between serial structure and predictability. This paper proposed a framework to magnify the correlations and regularities contained in financial time series through constructing accumulative return series. This method can help us distinguish the real world financial time series from random-walk process effectively by examining the change patterns of volatility, Hurst exponent, and approximate entropy. Furthermore, we have found that the predictable degree increases continually with the increasing length of accumulative return. Our results suggest that financial time series are predictable to some extent and approximate entropy is a good indicator to characterize the predictable degree of financial time series if we take the influence of their volatility into account.

DOI: 10.1103/PhysRevE.81.066116 PACS number(s): 89.65.Gh, 05.40.−a, 89.75.−k

I. INTRODUCTION

Can we obtain excess profits from financial market by forecasting financial time series? This question is directly related to whether the financial market prices are predictable. After running several tests, many financial economists have accepted the random-walk hypothesis, which states that financial market prices are completely random because of the efficiency of the market and thus the prices of the financial market cannot be predicted [1]. However, other economists and investors believe that prices may move in trends and past prices can be used to forecast future price changes to some degree [2]. Empirical studies have provided some evidences that there are long-term or short-term dependent relationships in financial time series, which offer promises for predictability [3,4]. As a model-independent measurement, approximate entropy (ApEn) has been proposed to characterize the irregular degree of financial time series. ApEn assigns a nonnegative number to a given sequential data, with larger values corresponding to greater apparent serial randomness or lower predictability and with smaller values corresponding to more regularity or higher predictability [5]. As shown in Ref. [5], ApEn values of Dow Jones index and Hang Seng (Hong Kong) index are significantly different from that of random series. From the viewpoint of investment, large number of models has been proposed to provide investors with more accurate forecasts. Time-series models, based on conventionally statistical methods, were very popular in constructing various kinds of financial market prediction models in the past [6]. Without relying too much on specific assumptions and error distributions, artificial neural networks (ANNs) have been demonstrated to be successful research models to forecast, detect, and summarize the structure of financial variables [7]. Recent years, results of some forecast exercises suggested that support vector machine (SVM) might be the most promising method for predicting financial time series [8,9]. Though there are a great number of papers investigating the predictability of financial market, we still do not fully understand the relationship between the predictable degree and the information structure of financial time series. Different from those studies emphasized on detecting the existence of correlated relationships, characterizing degree of randomness, and developing better forecast models, this paper proposed a simple but effective method to magnify the correlations and regularities contained in financial time series through constructing accumulative return series. By examining the change patterns of volatility, long-term memory property, and irregularity, we found that (1) the standard deviation increases convexly, (2) the Hurst (H) exponent increases monotonically and, (3) the approximate entropy decreases continually with the length of accumulative return, which are fundamentally different from random-walk process. Then, we investigated the predictability of accumulative return series and showed that the predictable degree increases with the increasing length of accumulative return. Our results suggest that we can obtain a deeper and clearer understanding of the predictability of financial market by magnifying the correlations and regularities contained in time-series data.

II. DATA DESCRIPTIONS

We analyzed 22 financial time series, which covered bond market, commodity market, stock market, and exchange market. Specifically, the data sets are comprised of (1) daily opening, highest, lowest, and closing interest rates of 10- and 30-yr treasury bond; (2) daily Reuters Jefferies CRB index; (3) daily opening, highest, lowest, and closing prices of Dow Jones index; (4) daily U.S. Dollar index; (5) daily opening, highest, lowest, and closing prices of Nasdaq Composite index; and (6) daily opening, highest, lowest, and closing prices of S&P 500 index. We obtained these historical data from Board of Governors of the Federal Reserve System, www.jefferies.com, and the finance section of Yahoo, respectively. The whole data sets stretched from 1994 to 2009. Each return series was calculated by the logarithmic change of the corresponding price series, $r(t)=\ln P(t)-\ln P(t-1)$, in which $P(t)$ denotes the index price on day t. Table I presents the name, simple description of all analyzed financial time series, and some basic statistical properties of their return series. As we can see from the third column in Table I, all

TABLE I. Name and description of the analyzed financial time series. The mean, standard deviation (SD), Hurst (H) exponent, and ApEn of each corresponding return series are also reported.

Serial name	Simple description	Mean	SD	H	ApEn
BOND10_O	Opening prices of 10-yr treasury bond	−0.00026905	0.012646	0.48803	1.8586
BOND10_H	Highest prices of 10-yr treasury bond	−0.00027001	0.011399	0.48919	1.8267
BOND10_L	Lowest prices of 10-yr treasury bond	−0.00027115	0.012029	0.49105	1.8103
BOND10_C	Closing prices of 10-yr treasury bond	−0.00026719	0.011807	0.48887	1.8202
BOND30_O	Opening prices of 30-yr treasury bond	−0.00019896	0.020444	0.35742	1.2778
BOND30_H	Highest prices of 30-yr treasury bond	−0.0001982	0.02005	0.35629	1.1636
BOND30_L	Lowest prices of 30-yr treasury bond	−0.00019899	0.020162	0.35965	1.1746
BOND30_C	Closing prices of 30-yr treasury bond	−0.00019574	0.020267	0.35707	1.2413
CRB	Reuters Jefferies CRB Index	0.00014146	0.008834	0.60794	1.9704
DJI_O	Opening prices of Dow Jones index	0.00020161	0.010608	0.4018	1.7824
DJI_H	Highest prices of Dow Jones index	0.00020566	0.0086197	0.42424	1.8152
DJI_L	Lowest prices of Dow Jones index	0.00020177	0.0098965	0.4003	1.6998
DJI_C	Closing prices of Dow Jones index	0.00020454	0.0107	0.40057	1.8041
USDI	U.S. Dollar index	−0.00002044	0.0039965	0.50658	1.8604
NCI_O	Opening prices of Nasdaq Composite index	0.00022315	0.01805	0.50863	1.5354
NCI _H	Highest prices of Nasdaq Composite index	0.00022751	0.014429	0.53292	1.5924
NCI _L	Lowest prices of Nasdaq Composite index	0.00022185	0.017282	0.49775	1.4945
NCI _C	Closing prices of Nasdaq Composite index	0.00022649	0.01725	0.50789	1.5785
S&P _O	Opening prices of S&P 500 index	0.0001716	0.010916	0.40832	1.7474
S&P _H	Highest prices of S&P 500 index	0.00017609	0.0088251	0.43118	1.7413
S&P _L	Lowest prices of S&P 500 index	0.00017215	0.010427	0.4034	1.6915
S&P _C	Closing prices of S&P 500 index	0.0001754	0.010933	0.407	1.7532

mean returns are in the range of ±0.0003, which is very close to zero. Standard deviation is often taken as a measurement to characterize the volatility of time series [5], which varies a lot with the least value less than 0.004 and the largest value larger than 0.02. It is worth noting that our data sets are very representative because their Hurst exponents (it is often referred to as the "index of dependence" and is the relative tendency of a time series to either strongly regress to the mean or "cluster" in a direction [10]) can be divided into three parts: (1) Reuters Jefferies CRB Index has a Hurst exponent larger than 0.5 significantly; (2) Hurst exponents of U.S. Dollar index, interest rates of 10-yr treasury bond, and Nasdaq Composite index fluctuate around 0.5; (3) other remained Hurst exponents are far less than 0.5. Furthermore, the correlation coefficients of standard deviation and Hurst exponent, standard deviation and approximate entropy, Hurst exponent and approximate entropy are −0.3398, −0.8909, and 0.5713, respectively. It was reported that the Hurst exponent and the ApEn value are negatively correlated [11,12], but our data sets exhibit positive correlation. This is owing to most Hurst exponents being less than 0.5 in our work, but most of them larger than 0.5 in Refs. [11,12]. Therefore, the correlated relationship between Hurst exponent and ApEn depends on the specific data sets being analyzed.

III. CHANGE PATTERNS OF ACCUMULATIVE RETURN SERIES

In this section, we propose a simple but effective method to magnify structural information contained in financial time series. Given return series $r(t)$, its accumulative return series is defined as

$$R_l(t) = \sum_{j=1}^{l} r(t+1-j) = \ln[P(t)] - \ln[P(t-l)], \qquad (1)$$

where l denotes the length of the accumulative return. Evidently, the accumulative return series is simply logarithmic-return time series constructed by using overlapping time windows, and it will be the same as the original return series if $l=1$. It is reasonable to anticipate that $R_l(t)$ contains more information than $r(t)$ does, at least the same, provided that there are some long-term or short-term dependent relationships in real world financial time series. In another word, we believe $R_l(t)$ would enlarge the correlated relationships and regularities which already exist in $r(t)$. In this paper, we use three methods to characterize the quantitative information contained in a series data. The first one is volatility, which is an implicit measurement of risk and grades the extent of deviation from centrality. Large swings exhibit high standard deviation and are often conceived as highly volatile. The second one is Hurst exponent, which characterizes the degree of interdependence and measures the relative tendency of a time series to either strongly regress to the mean or cluster in a direction [10]. The last one is ApEn, which reflects the likelihood that similar patterns of observations will not be followed by additional similar observations. Less repetitive patterns contained in a time series corresponds to higher

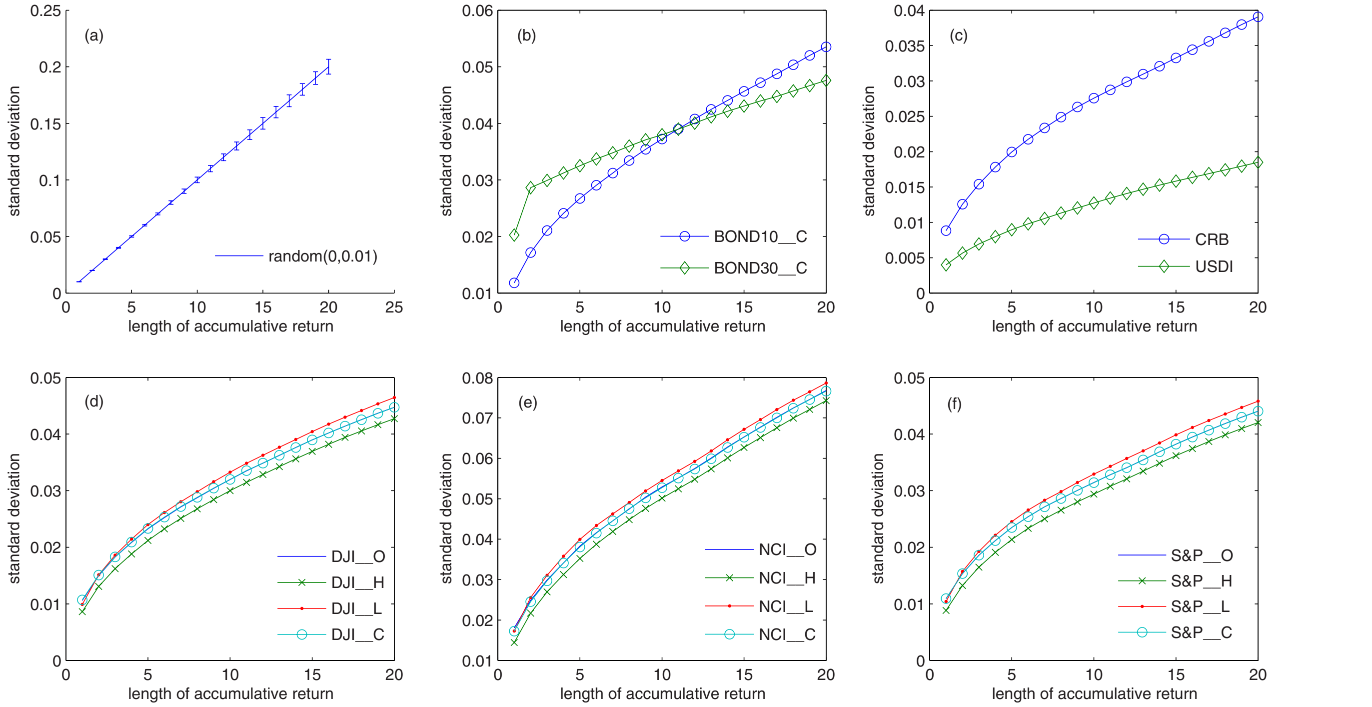

FIG. 1. (Color online) Measured the standard deviation in a function of the length of accumulative return. (a) Accumulative return series of random-walk process. The random return series is constructed with zero mean and standard deviation 0.01. (b) Accumulative return series of 10- and 30-yr treasury bond. Considering that the standard deviations of the daily opening, highest, lowest, and closing return series are roughly collapsed onto a single curve, only the results on closing return series are presented. (c) Accumulative return series of Reuters Jefferies CRB index and U.S. Dollar index. (d) Accumulative return series of Dow Jones index. (e) Accumulative return series of Nasdaq Composite index. (f) Accumulative return series of S&P 500 index. As we can see from this figure, the standard deviation of real world accumulative return series increases convexly with the length of accumulative return, which is completely different from that of random accumulative return series.

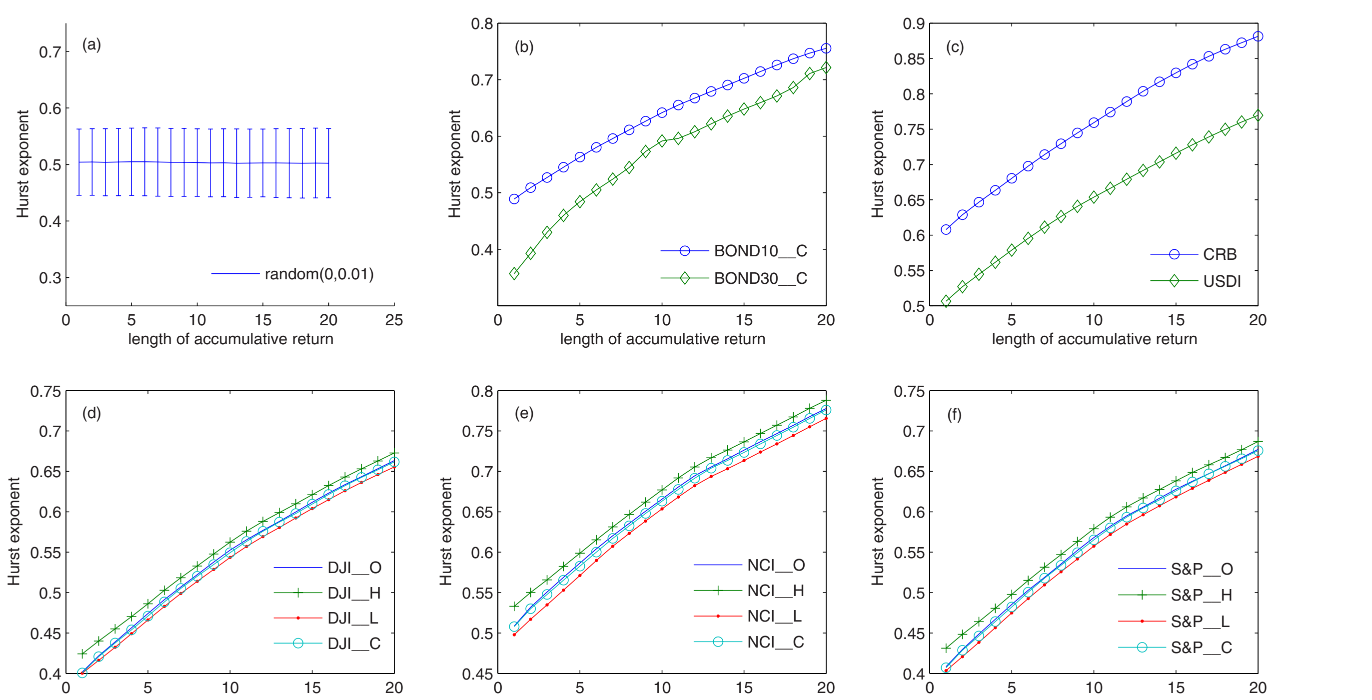

FIG. 2. (Color online) Measured the Hurst exponent in a function of the length of accumulative return. (a) Accumulative return series of random-walk process. (b) Accumulative return series of 10- and 30-yr treasury bond. (c) Accumulative return series of Reuters Jefferies CRB index and U.S. Dollar index. (d) Accumulative return series of Dow Jones index. (e) Accumulative return series of Nasdaq Composite index. (f) Accumulative return series of S&P 500 index. Hurst exponent increases monotonically with the increasing length of accumulative return in all 22 financial time series. This pattern does not appear in random accumulative return series.

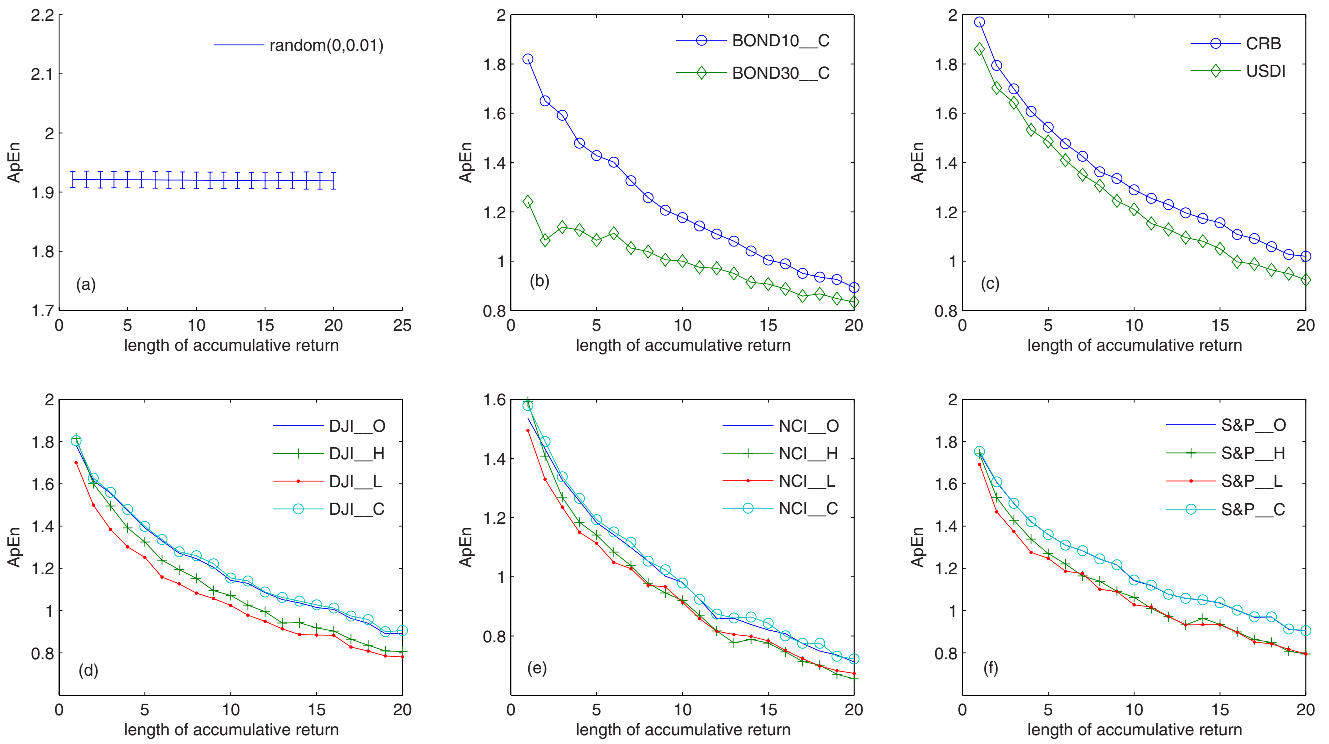

FIG. 3. (Color online) Measured the approximate entropy in a function of the length of accumulative return. (a) Relationship between ApEn and the length of accumulative return in random-walk process. (b) Relationship between ApEn and the length of accumulative return in 10- and 30-yr treasury bond. (c) Relationship between ApEn and the length of accumulative return in Reuters Jefferies CRB index and U.S. Dollar index. (d) Relationship between ApEn and the length of accumulative return in Dow Jones index. (e) Relationship between ApEn and the length of accumulative return in Nasdaq Composite index. (f) Relationship between ApEn and the length of accumulative return in S&P 500 index. The ApEn values of random accumulative return series are irrelevant to the length of accumulative return. However, the irregular degree decreases continually with the increasing length of accumulative return in all 22 real world financial time series. This results indicate that the constructing procedure of accumulative return can effectively magnify all kinds of regularities existed in financial time series, regardless of the specified financial market.

ApEn value and lower predictable degree. Therefore, ApEn can be an effective technology to measure the degree of irregular or unpredictable [13]. Considering it is very easy to compute volatility (i.e., standard deviation), we only describe the methods to compute Hurst exponent and ApEn in the latter.

There are a variety of techniques that exist for estimating Hurst exponent. Considering that the detrended fluctuation analysis (DFA) is one of the most efficient methods [14–16], we utilize the DFA method to compute Hurst exponent in this work. Given time series $X(i)$ with length N, we first compute the accumulative sum $y(k)=\sum_{i=1}^{k}[X(i)-\bar{X}]$, where $\bar{X}$ is the average value of $X(i)$. Next, we divide $y(k)$ into time windows of length n samples and estimate the trend lines $y_n(k)$ by using the ordinary least-squares principle in each window. Then, the root-mean-square deviation from the trend, the fluctuation, is calculated over every window at every time scale:

$$F(n) = \sqrt{\frac{1}{N}\sum_{k=1}^{N}[y(k) - y_n(k)]^2}. \qquad (2)$$

Typically, a linear relationship $F(n)$ increasing with n on a double logarithmic graph indicates the presence of scaling relationship $F(n) \sim n^H$, where H is the Hurst exponent. $H = 0.5$ corresponds to random-walk process which means no memory in the time series. $0 < H < 0.5$ indicates large and small values of the time series are more likely to alternate. If $0.5 < H < 1$, there are persistent long-range power-law correlations in the time series.

Now we introduce the algorithm for computing ApEn. Given a time series with N observations, X_h, $h=1,\ldots,N$, creates embedding vector $v(i)$, each made up of m consecutive values of X, $v(i)=[X_i,\ldots,X_{i+k}]$ with $k=1,\ldots,m-1$. The distance of vector $v(i)$ and $v(j)$ is computed as $d[v(i),v(j)] = \max\{|X_{i+k}-X_{j+k}|, k=1,\ldots,m\}$. To quantify the regularity of a particular pattern, count the relative frequency of distance between the template vector $v(i)$ to all the vectors $v(j)$, which lie within the neighborhood r [13]:

$$C_i^m(r) = (N - m + 1)^{-1} \sum_{j=0}^{N-m-1} \Theta(r - d[v(i),v(j)]), \qquad (3)$$

where $\Theta(\cdot)$ represents the binary Heaviside function. Let $\phi^m(r)=(N-m+1)^{-1}\sum_{i=1}^{N-m+1}\ln C_i^m(r)$, we can obtain the approximate entropy of the time series: $\mathrm{ApEn}(m,r)=\phi^m(r) - \phi^{m+1}(r)$. As shown in the computation process, we need to choose values for two input parameters, m and r. In this paper, we apply the embedding dimension $m=2$ and the tolerance of similarity $r=20\%$ of standard deviation of the

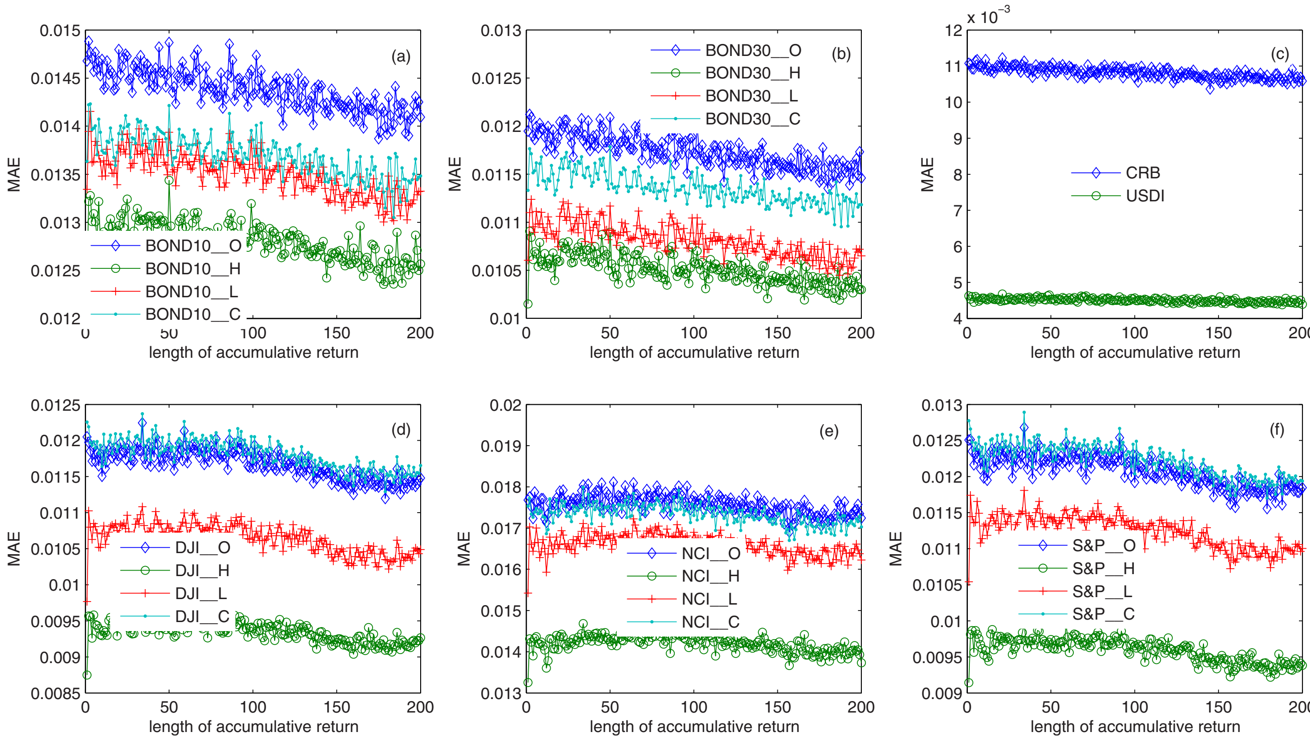

FIG. 4. (Color online) Measured the MAE in a function of the length of accumulative return. (a) Relationship between MAE and the length of accumulative return in 10-yr treasury bond. (b) Relationship between MAE and the length of accumulative return in 30-yr treasury bond. (c) Relationship between MAE and the length of accumulative return in Reuters Jefferies CRB index and U.S. Dollar index. (d) Relationship between MAE and the length of accumulative return in Dow Jones index. (e) Relationship between MAE and the length of accumulative return in Nasdaq Composite index. (f) Relationship between MAE and the length of accumulative return in S&P 500 index. For each financial time series, the mean and standard deviation of 200 MAEs are reported in the sixth and seventh columns of Table II. The correlation coefficient between the mean value of MAEs and the standard deviation of original return series is 0.5113.

specified time series, similar to what has been used in previous works [5]. We have also applied $m=1$ to compute approximate entropy and found that the following analysis does not change.

Let us now illustrate how the accumulative return method can magnify the correlations and regularities in financial time series. For each return series, we first construct 20 accumulative return series and then apply the above three measurements to them. Considering that there is a common belief of daily return being a random-walk process, we also construct a simulated return series based on random-walk model with zero mean and standard deviation 0.01. Therefore, we can examine whether a systematically different change patterns of accumulative return series exists between real world financial time series and random return series. To test the consistency of the results, we have run 100 trials for each random-walk process and compute the corresponding averages and deviations, which are plotted as the curves with error bars in Figs. 1(a), 2(a), and 3(a).

As shown in Fig. 1(a), the standard deviation of random return series increases linearly with the length of accumulative return. However, we observed a completely different function relationship in the real world accumulative return series. As we can see from Figs. 1(b)–1(f), the standard deviation increases convexly with the length of accumulative return, which makes the magnitude of accumulative return variations enlarge much slower than the one generated by random-walk hypothesis. Therefore, our accumulative return method can help us distinguish a real world financial time series from a random series effectively even though these two return series are very similar. This effect is achieved by magnifying the existed correlations and regularities, which is explained in Figs. 2 and 3. Our results confirm that the random-walk hypothesis is not acceptable because it ignores some structural information contained in real world financial time series. Furthermore, our results also put forward a useful standard to evaluate the performance of time-series model. A good model should reproduce the above empirical change pattern of standard deviation of real world financial time series.

As we know, Hurst exponent is often used to characterize the long-term dependent relationships in a time series. For a random series, its Hurst exponent should be 0.5 and does not change with the increasing length of accumulative return, which is confirmed by our results [see Fig. 2(a)]. However, the Hurst exponent of real world financial time series increases monotonically with the increasing length of accumulative return. As shown in Figs. 2(b)–2(f), all correlated relationships will be magnified continually by increasing the length of accumulative return. Furthermore, we have noted that this trend is regardless of the Hurst exponent of the original return series. For example, the Hurst exponent values of daily opening, highest, lowest, and closing interest rates of 10-yr treasury bond and Nasdaq Composite index are

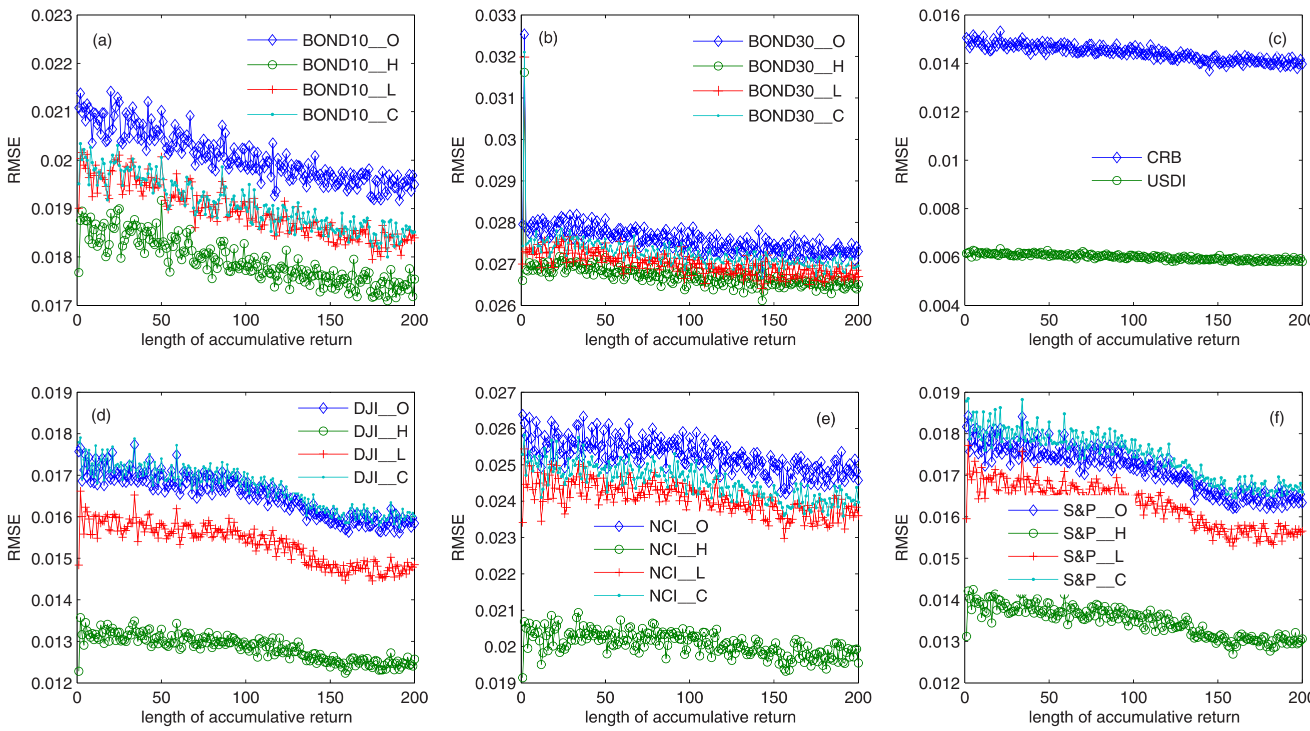

FIG. 5. (Color online) Measured the RMSE in a function of the length of accumulative return. (a) Relationship between RMSE and the length of accumulative return in 10-yr treasury bond. (b) Relationship between RMSE and the length of accumulative return in 30-yr treasury bond. (c) Relationship between RMSE and the length of accumulative return in Reuters Jefferies CRB index and U.S. Dollar index. (d) Relationship between RMSE and the length of accumulative return in Dow Jones index. (e) Relationship between RMSE and the length of accumulative return in Nasdaq Composite index. (f) Relationship between RMSE and the length of accumulative return in S&P 500 index. For each financial time series, the mean and standard deviation of 200 RMSEs are reported in the fourth and fifth columns of Table II. The correlation coefficient between the averaged RMSEs and the standard deviation of original return series is 0.9878.

very close to 0.5, but their accumulative return series exhibit similar behaviors like other time series with Hurst exponent larger or less than 0.5 significantly.

Figures 3(b)–3(f) show that the irregular degree of real world accumulative return series decreases continually with the increasing length of accumulative return. However, as a comparative base, the approximate entropy of the accumulative random return series swings up and down and exhibits no systematic change pattern [see Fig. 3(a)]. The constructing procedure of accumulative return does not change the random return series into a nonrandom series, thus the increased regularities in the real world accumulative return series only come from the original return series. In another word, the accumulative return method can enlarge the existed regularities effectively but not generate any new regularity. As shown in Fig. 3, the empirical change patterns of approximate entropy provide some evidences that financial time series are predictable to some extent, because there is some structural information contained in serial data.

As analyzed in the above, we have found that (1) the standard deviation increases convexly, (2) the Hurst exponent increases monotonically, and (3) the approximate entropy decreases continually with the length of accumulative return in real world financial time series. These patterns might be observed in intraday return series because they exhibit scaling and memory behaviors similar to daily return series [17,18]. However, these empirical change patterns of volatility, correlated dependence, and irregularity cannot be observed in random-walk process. Therefore, we can conclude that (1) we can magnify the correlations and regularities contained in return series effectively by employing the accumulative return method, (2) the random-walk hypothesis is not acceptable for financial time series, and (3) we can predict financial time series to some degree by exploiting those correlations and regularities contained in serial data. Furthermore, the accumulative return method can also be used to evaluate time-series model. A good model should reproduce the previously empirical change patterns observed in real world financial time series.

IV. PREDICTABLE DEGREE OF ACCUMULATIVE RETURN SERIES

In this section, we examine how the predictable degree of accumulative return series changes with the length of accumulative return. This paper mainly focuses on the predictable degree of time series rather than on the predictability of forecast model. Considering that prediction accuracy is the most concerned issue for forecasters, we use prediction accuracy measurements to characterize the predictable degree of specified serial data. In order to exclude the influences of parameters used in forecast model and the model itself on prediction accuracy, we use a model-independent forecast method to perform our forecast exercises. Given accumula-

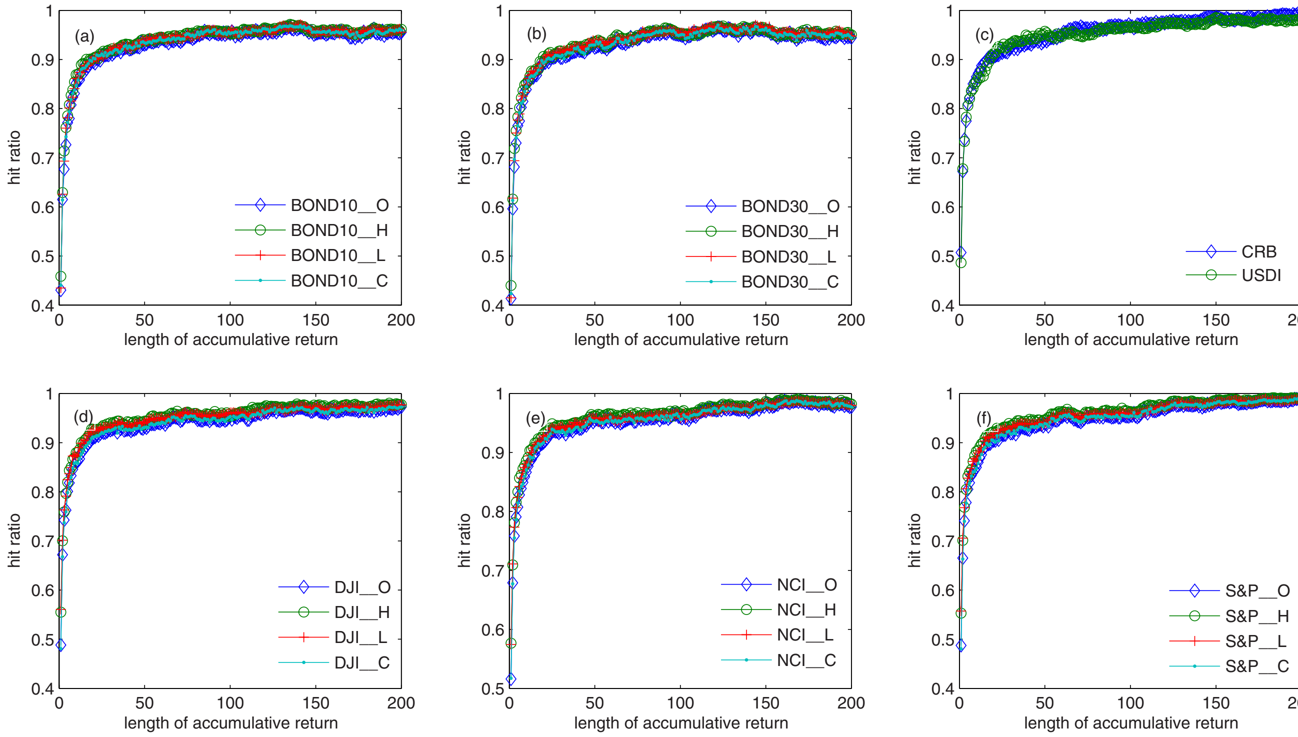

FIG. 6. (Color online) Measured the hit ratio in a function of the length of accumulative return. (a) Relationship between the hit ratio and the length of accumulative return in 10-yr treasury bond. (b) Relationship between the hit ratio and the length of accumulative return in 30-yr treasury bond. (c) Relationship between the hit ratio and the length of accumulative return in Reuters Jefferies CRB index and U.S. Dollar index. (d) Relationship between the hit ratio and the length of accumulative return in Dow Jones index. (e) Relationship between the hit ratio and the length of accumulative return in Nasdaq Composite index. (f) Relationship between the hit ratio and the length of accumulative return in S&P 500 index. As shown in this figure, the directional accuracy of the predicted accumulative returns increases continually with the length of accumulative return, though the directional accuracy of the predicted original returns fluctuates around 0.5.

tive return series $R_l(t)$, let the correspondingly predicted accumulative return series $\widetilde{R_l(t)}=R_l(t-1)$. This method is very easy to implement and can be applied to any time series. Another advantage is that we can compare the predictable degree of different serial data. Following we introduce several measurements to characterize the prediction accuracy.

The mean absolute error (MAE) and the root-mean-squared error (RMSE) are two most common measurements [19], which are defined as follows:

$$\mathrm{MAE}=\sum_{t=2}^{N}|R_l(t)-\widetilde{R_l(t)}| \Big/ (N-1), \quad (4)$$

$$\mathrm{RMSE}=\sqrt{\sum_{t=2}^{N}(R_l(t)-\widetilde{R_l(t)})^2 \Big/ (N-1)}, \quad (5)$$

where N is the number of data points in $R_l(t)$. Furthermore, financial forecasts should be tied to the profitability of virtual investors' decisions rather than to simple statistical measurements such as the root-mean-squared error (RMSE). Such measurement aims at minimizing an unrelated to profitability loss function rather than at getting a significant outcome from the viewpoint of profit maximization. Therefore, we further consider the third measurement which is related to profitability: hit ratio, corresponding to the rate of consistency between the direction of the actual price change and that of the predicted one [20].

We now apply the above forecast procedure and measurements of predictable degree to all financial time series and their accumulative return series. As shown in Fig. 4, 200 MAEs swing up and down with a downward trend in each curve. Even though there are some differences, all curves exhibit the same trend. We also observed similar fluctuating patterns of RMSEs in Fig. 5. These results show that the prediction errors do not increase with the length of accumulative return though the standard deviation of accumulative return series increase monotonically with that. Conversely, we have observed a general trend of decreasing prediction errors for all financial time series. Furthermore, as we can see from Fig. 6, the hit ratio increases continually with the length of accumulative return. Therefore, we can conclude that the predictable degree will be improved to some degree if we increase the length of accumulative return. Considering that ApEn characterizes the irregular degree of serial data, the previous results are not difficult to understand if we take the decreasing ApEn of accumulative return series into account.

Let us now examine how the predictable degree of financial time series is related to its standard deviation, Hurst exponent, and approximate entropy. We first compute the correlation coefficients between the mean values of MAEs and RMSEs (see the fourth and sixth columns in Table II)

TABLE II. Prediction accuracy and factors related to the predictable degree of financial time series. The first column presents the name of each financial time series. Standard deviation, Hurst exponent, and approximate entropy are three measurements tightly connected to the predictable degree of time series. For each return series, the correlated coefficients of (standard deviation, Hurst exponent) and (standard deviation, approximate entropy) of corresponding accumulative return series are reported in the second and third columns. For every analyzed financial time series, its mean value, standard deviation of 200 RMSEs and 200 MAEs are presented in the fourth, fifth, sixth, and seventh columns, respectively.

Serial name	ρ (H, SD)	ρ (ApEn, SD)	Mean_RMSE	STD_RMSE	Mean_MAE	STD_MAE
BOND10_O	0.99833	−0.99484	0.020083	0.00053543	0.01438	0.0002129
BOND10_H	0.99596	−0.99313	0.017915	0.00046531	0.012802	0.00019835
BOND10_L	0.99757	−0.99442	0.018954	0.00053564	0.013498	0.00021569
BOND10_C	0.99711	−0.99483	0.019087	0.00052976	0.013697	0.00021297
BOND30_O	0.98728	−0.97754	0.027584	0.00044388	0.011738	0.00017163
BOND30_H	0.9866	−0.97847	0.026726	0.00040157	0.010508	0.00014965
BOND30_L	0.98657	−0.96512	0.027024	0.00043798	0.010815	0.00016612
BOND30_C	0.98711	−0.98008	0.027243	0.00042155	0.011357	0.00015916
CRB	0.99379	−0.99401	0.014427	0.00031852	0.010816	0.00015213
DJI_O	0.99645	−0.99478	0.016529	0.00051123	0.011675	0.00020141
DJI_H	0.99428	−0.98803	0.012841	0.00030186	0.0093479	0.00014557
DJI_L	0.99324	−0.98868	0.015352	0.00048237	0.010641	0.00020425
DJI_C	0.99592	−0.99399	0.016702	0.00052078	0.011808	0.00020747
USDI	0.99801	−0.99379	0.0060032	0.0001164	0.0045062	5.5955e-005
NCI_O	0.99821	−0.99049	0.025265	0.00044943	0.01756	0.00023434
NCI _H	0.99624	−0.98262	0.020069	0.00033173	0.014205	0.00021111
NCI _L	0.99593	−0.98941	0.024073	0.00042899	0.016596	0.00025268
NCI _C	0.99734	−0.98903	0.024534	0.00047081	0.017331	0.00023697
S&P _O	0.99595	−0.99322	0.017134	0.00054322	0.012083	0.000219
S&P _H	0.99411	−0.98778	0.013474	0.00036353	0.009593	0.00016475
S&P _L	0.99303	−0.98963	0.016286	0.00052837	0.011242	0.00022255
S&P _C	0.99564	−0.99293	0.017467	0.00059761	0.012228	0.00023043

and the corresponding standard deviations of 22 original return series (see the fourth column in Table I) and find that they are 0.5113 and 0.9878, respectively. These results suggest that standard deviation is also related to predictability though it is often conceived as measuring the extent of deviation from centrality. Next, for each financial time series, we compute the correlation coefficients of (standard deviation, Hurst exponent) and (standard deviation, approximate entropy) of its accumulative return series (see the second and third columns in Table II), and find that all values are very close to 1. Furthermore, Hurst exponent characterizes correlated relationship, which is only one kind of regularities contained in serial data. Therefore, there might be some equivalent or encompassing relationship between Hurst exponent and approximate entropy. We further compute the correlation coefficient between the mean values of MAEs and RMSEs and ApEn (Hurst exponent), and then find that they are 0.8306 (0.7860) and 0.9351 (0.8844). In order to exclude the influences of standard deviation on predictability, all the mean values of MAEs and RMSEs in this computation process are divided by their corresponding standard deviations of the original return series. Combining the previous analysis, we can conclude that (1) approximate entropy is a better indicator of predictable degree than Hurst exponent; (2) when we compare the predictable degrees of several different financial time series based on ApEn, we need to exclude the influences of their standard deviations.

V. DISCUSSIONS AND CONCLUSIONS

This paper proposed a framework to explore the relationship between the predictable degree of financial time series and its statistical properties of standard deviation, Hurst exponent, and approximate entropy. By constructing accumulative return series, the correlations and regularities contained in the original return series can be magnified effectively, which can help us distinguish the real world financial time series from completely random series. Given a real world return series, we would observe that (1) the standard deviation increases convexly, (2) the Hurst exponent increases monotonically, and (3) the approximate entropy decreases continually with the length of accumulative return, which are fundamentally different from that of random walk. These founds set some limitations for developing better financial time-series model, which should reproduce the previously empirical change patterns of volatility, correlated dependences, and irregularity. Furthermore, our results have demonstrated that financial time series are predictable to some extent and approximate entropy is a good indicator to characterize the predictable degree of financial time series if we take the influence of their volatility into account. Our method can magnify the regularities contained in financial time series, but how to utilize this method to develop better forecast model is still unknown and needs some further study.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China under Grant No. 70701030 and Zhejiang Provincial Natural Science Foundation of China under Grant No. Y607001.

[1] B. G. Malkiel, *A Random Walk Down Wall Street: The Time-Tested Strategy for Successful Investing*, 6th ed. (W.W. Norton, New York, 1973).

[2] A. W. Lo and A. C. Mackinlay, *A Non-Random Walk Down Wall Street*, 5th ed. (Princeton University Press, Princeton, NJ, 2002).

[3] K. Aydogan and G. G. Booth, South. Econ. J. **55**, 141 (1988).

[4] A. W. Lo, Econometrica **59**, 1279 (1991).

[5] S. Pincus and R. E. Klman, Proc. Natl. Acad. Sci. U.S.A. **101**, 13709 (2004).

[6] P. H. Franses and D. V. Dijk, J. Forecast. **15**, 229 (1996).

[7] J. V. Perez-Rodriguez, S. Torra, and J. Andrada-Felix, J. Empirical Finance **12**, 490 (2005).

[8] V. N. Vapnik, *Statistical Learning Theory* (Wiley, New York, 1998).

[9] K. J. Kim, Neurocomputing **55**, 307 (2003).

[10] B. B. Mandelbrot, Econometrica **39**, 68 (1971).

[11] C. Eom, S. Choi, G. Oh, and W.-S. Jung, Physica A **387**, 4630 (2008).

[12] C. Eom, G. Oh, and W.-S. Jung, Physica A **387**, 5511 (2008).

[13] S. Pincus, Proc. Natl. Acad. Sci. U.S.A. **88**, 2297 (1991).

[14] C. K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger, Phys. Rev. E **49**, 1685 (1994).

[15] B. Podobnik and H. E. Stanley, Phys. Rev. Lett. **100**, 084102 (2008).

[16] R. Weron, Physica A **312**, 285 (2002).

[17] F. Wang, K. Yamasaki, S. Havlin, and H. E. Stanley, Phys. Rev. E **73**, 026117 (2006).

[18] F. Z. Wang, S. J. Shieh, S. Havlin, and H. E. Stanley, Phys. Rev. E **79**, 056109 (2009).

[19] G. S. Atsalakis and K. P. Valavanis, Expert Syst. Applic. **36**, 5932 (2009).

[20] M. H. Pesaran and A. Timmermann, J. Bus. Econ. Stat. **10**, 461 (1992).