

Evaluating the Complexity of Time Series Based on Distributional Orderliness



Wen-Jing Yu¹, Jie Yu¹, Ling-Yu Xu^{1,2*}, Gao-Wei Zhang¹, Shi-Wei Guo³

¹ Department of Computer Engineering and Science, Shanghai University,
No. 99 Shangda Road, Shanghai, China
1522474161@shu.edu.cn

² Shanghai Institute for Advanced Communication and Data Science, Shanghai University,
No. 99 Shangda Road, Shanghai, China
xly@shu.edu.cn

³ Changhai Hospital of Shanghai
No. 168 Changhai Road, Yangpu District, Shanghai, China
1421286162@qq.com

Received 5 December 2017; Revised 31 March 2018; Accepted 2 June 2018

Abstract. It is significant to study the complexity of financial time series, since the financial market is a complex dynamic system. Sample entropy (SampEn) is a widely used method to quantify the complexity of time series. However, studies showed that an increase in the SampEn may not always be associated with an increase in dynamical complexity. To deal with the problem, we proposed a modified method of SampEn to measure the complexity of complex dynamical systems. The method based on a time decay function and presented a different way of the complexity of time series. Simulations were conducted over artificial and stock time series for providing the comparative studies. Results showed that the modified method can distinguish time series from different complex systems and time series with different distributions. Furthermore, compared with SampEn, the results were more consistent with the real complexity of time series. Finally, the modified method was applied to financial time series, and get some interesting results.

Keywords: complexity, financial time series, sample entropy, time decay function

1 Introduction

Since the global economic crisis, studying the complexity of the financial systems has attracted extensive attention. The stock markets are often considered to be complex systems with many complexity factors [1]. Studying the dynamic complexity of the stock market can reveal the rule of financial market hidden behind financial data and prevent risk. Various nonlinear techniques and theoretical methods have been developed to characterize the dynamic of the stock markets, such as Lyapunov exponent method [2], detrended fluctuation analysis [3], correlation dimension method [4], and entropy.

Compared with other methods, entropy is a prevailing approach to measure the complexity of time series. And a range of entropy-based algorithms have been proposed, such as approximate entropy (ApEn) [5], SampEn [6], multiscale entropy (MSE) [7], permutation entropy (PE) [8] etc. SampEn, developed by Richman and Moorman (2000), is a classical entropy approach to quantify the regularity or complexity of time series and it is a modified algorithm of ApEn. The SampEn has been successfully applied to different research fields in the past decades. These applications include the analyses of the human heart rate variability [9], traffic time series [10], financial time series [11], electroencephalography (EEG) [12],

* Corresponding Author

etc.

Despite its advantages mentioned above, studies showed that an increase in the entropy may not always be associated with an increase in dynamical complexity [7]. For instance, some diseased systems, showing more regular behavior, should have produced reduced entropy values compared to the dynamics of free-running healthy systems. However, SampEn gave a higher entropy value to the diseased systems than healthy systems. As for this problem, Costa thinks time series derived from complex systems, such as healthy physiologic are likely to present structures on multiple spatiotemporal scales. But SampEn failed to account for the multiple time scales inherent in healthy physiologic dynamics. Therefore, Costa proposed the MSE method [7], which aims to measure the complexity of time series over different scales. And this method has been accepted by many scholars.

However, the coarse-grained process in MSE actually loses a lot of original information in time series. And the problem of sample entropy algorithm is not actually solved. Since, the SampEn depends on a function's one step difference and reflect the uncertainty of the next new point, given the past history of the series. Therefore, this way does not account for features related to structure other than the shortest one. To deal with the problem, in this paper, we proposed a modified SampEn named two-dimensional entropy (TD_entropy). In TD_entropy, a time-decay function is used in pattern matching to measure the similarity degree of two vectors based on their time distance. Therefore, the influence of the distribution of similar vectors in time series on the complexity of time series is taken into consideration. Moreover, the influence of the degree of the self-similarity of vectors on the complexity of time series is also taken into account in TD_entropy.

The remainder of the paper is organized as follows: In section 2, we introduce some related works. In section 3, we introduce the relationship between similarity and attenuation as well as the TD_entropy algorithm. In section 4, we apply several groups of studies to show the effectiveness and consistency of TD_entropy and the differences between TD_entropy and SampEn. In section 5, the TD_entropy is used in analysis of the financial time series. Finally, conclusions are given in section 6.

2 Related Works

SampEn is a very effective method to measure the complexity of time series. The SampEn algorithm, and modified SampEn algorithm have been applied to many field to study the complexity of different dynamical systems. Therefore, many scholars have done a lot of research on the SampEn.

In SampEn, there are two important parameters parameter: m and r . The m is called embedding dimension and determines the length of the sequences to be compared. And another parameter r is the tolerance threshold, which determines the threshold value for accepting the similarity patterns. On the selection of entropy parameters, there have been several reports about the selection of parameters (e.g., [13-15]). The results of these studies are in agreement on a couple of parameters, but there are different results published for some others. The threshold parameter r usually be set as 0.2 time of the data's standard deviation, and m is 2. Considering the threshold parameter r is based on long term SD of the original time series, Marwaha and Sunkaria [16] proposed an improved sample entropy (ISampEn), in which the r is updated by considering the period-to-period variations of a time series.

As for the problem, that SampEn will be not defined if no template and forward match occurs in the case of small r and datasets length N , some fuzzy entropy methods have been proposed [17-19]. The function in SampEn used for determining the vectors' similarity is replaced with a continuous function. Hence, the fuzzy entropy statistics change smoothly when there is a slight change in the tolerance r .

For the problem, when data series is oversampled, the sampling frequency is much higher than the frequency of the studied time series, SampEn may give misleading results. Therefore, L [20] modified the definition of SampEn by including a lag between successive data points of the vectors to be compared to address the oversampled issue.

To deal with the problem that an increase in the entropy may not always be associated with an increase in dynamical complexity, Costa proposed the MSE method. And MSE is a prevailing and useful method used to quantify the complexity of a time series over a range of scales. Since the coarse-grained process reduces the length of the time series quickly, imprecise estimation of entropy and undefined entropy values will be produced. To address this problem, some modified MSE methods have been proposed [21-23].

Considering that the MSE method and modified MSE methods actually do not solve the inherent problem of SampEn, we propose the TD_entropy algorithm. In TD_entropy, the time attribute of the vectors is considered in pattern matching process. Hence, the similarity of the vectors in the time series is not only depends on their pattern distance but also depends on their time distance. On the other hand, the entropy value of the TD_entropy not only related to the information production, but also related to the similarity degree of the vectors in the time series.

3 Methodology

In this section, we first introduce the time decay function in TD_entropy, and then the TD_entropy algorithm is introduced in detail.

3.1 Similarity and time Attenuation

As we know in some field, the similarity of things will change with time, like hobbies. When recommending according to the preference of the users, we need to give a lower weight to what they used to like, and give a higher weight for that they like recently. And this idea of the effect of time on similarity is applied in many fields, such as interest recommend systems [24], microblog forwarding prediction [25], patent novelty [26], temporal data clustering [27], etc.

However, in SampEn the similarity of vectors is only based on their pattern distance, which is defined as the maximum difference between the two vectors. Therefore, this way actually ignores the time attribute of the vectors in the time series. Thus, in SampEn, so long as the pattern distance of any two vectors is in the similar tolerance, they are similar, no matter what position they are in the time series. Hence, this way ignores the effect of vectors distributions on the complexity of time series.

In general, the time decay model is a method that reduces the number of historical transactions to support the number weight over time [28]. And the time decay function in TD_entropy, is defined as formula (1), which is a widely used model [24-27].

$$f(i, j) = e^{-k \times |i-j|}. \quad (1)$$

The k is the time decay coefficient, i and j are the positions of the vectors in the time series. Fig. 1 illustrates the decay curves of five time decay coefficients with the time distance from 0 to 500.

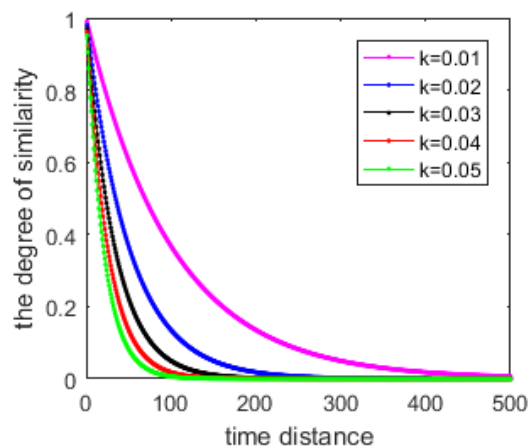


Fig. 1. The time decay curves

As seen in Fig. 1, the smaller the time distance between two vectors, the greater the similarity degree between the two vectors. And when the time distance between two vectors is far away, the similarity degree between the two vectors can be ignored. Moreover, the degree of attenuation increases with the increase of k value.

Then we define the critical distance as the distance, when the similarity of the two vectors under an attenuation coefficient decays to a very small probability. And if the time distance of two vectors is beyond the critical distance, means their similarity is negligible, since the degree of the two vectors'

similarity is too small.

Table 1 shows the critical distance n under different k when the similarity of two vectors is 0.001. From Table 1 we can see the n is rapidly decreasing with the increase of k . Therefore, in the calculation of the similarity between vectors, if we need to consider the whole vectors similarity in the time series, then we should choose the k whose critical distance n is greater than the length of time series; If we calculate the similarity of vectors of long term time series, we can choose the appropriate k according to the critical distance when we don't want to calculate the whole time series.

Table 1. The critical distance of the different coefficients

k	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
n	0	690	345	230	172	138	115	98	86	76
k	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
n	69	34	23	17	13	11	9	8	7	6

3.2 TD_entropy Algorithm

According to the new similarity discussed above, the TD_entropy is defined as follows: For a time series with the data length N , $\{u(i):1 \leq i \leq N\}$. Form template vectors: $X_i^m, X_j^m, \dots, X_{N-m+1}^m$ where m is the embedding dimension and it determines the length of sequences to be compared. The template vectors are given as:

$$X_i^m = \{u(i), u(i+1), \dots, u(i+m-1)\} \quad (2)$$

$$(i = 1, 2, \dots, N - m + 1)$$

X_i^m represents m consecutive u values. The distance $d_{i,j}^m$ between X_i^m and X_j^m is defined as:

$$d_{i,j}^m = \max |u(i+k) - u(j+k)| \quad (3)$$

$$(k \in [0, m-1], i \neq j)$$

Given the tolerance threshold r , where r stands for the tolerance for the accepting matches. The k is the time decay coefficient. The similarity $\mu_{i,j}^m$ between X_i^m and X_j^m is defined as:

$$\mu_{i,j}^m = \begin{cases} 1 \times e^{-k \times |i-j|} & d_{i,j}^m \leq r \\ 0 & d_{i,j}^m > r \end{cases} \quad (4)$$

Therefore, the similarity of two vectors X_i^m and X_j^m depends on their pattern distance and time distance. And the pattern distances determine whether they are similar, the time distance determines the degree of their similarity. Then define

$$B_i^m(r, k) = \frac{\sum_{j=1}^{N-m} \mu_{i,j}^m}{N - m - 1} \quad (5)$$

$$B^m(r, k) = \frac{\sum_{i=1}^{N-m} B_i^m(r, k)}{N - m} \quad (6)$$

$B^m(r, k)$ is the probability that any two sequences will match for m points. Whereas, $B^{m+1}(r, k)$ is the probability that any two sequences will match for $m+1$ points.

Finally, TD_entropy is defined as:

$$TD_entropy(m,r,k) = \lim_{N \rightarrow \infty} -\ln\left[\frac{B^{m+1}(r,k)}{B^m(r,k)} \times B^{m+1}(r,k)\right]. \tag{7}$$

Which is estimated by the statistic:

$$TD_entropy(N,m,r,k) = -\ln\left[\frac{B^{m+1}(r,k)}{B^m(r,k)} \times B^{m+1}(r,k)\right]. \tag{8}$$

As suggested by Richman and Moorman [6], the embedding dimension m is set to be 2, the tolerance r is set to be $0.2 \times SD$ in both $TD_entropy$ and $SampEn$, where SD is the standard deviation of the time series. Besides, the time decay coefficient k is set to 0.01 in $TD_entropy$.

4 Empirical Results

In this section, we use the Logistic data sets, the Hennon map data sets, real stock time series and simulated time series to verify the validity and the consistency of $TD_entropy$, as well as the differences between $TD_entropy$ and $SampEn$.

4.1 The Validity and Consistency of $TD_entropy$

In this subsection, we use the Logistic data sets and the Hennon map data sets to verify the performance of $TD_entropy$. The Logistic data sets is obtained by:

$$x(i+1) = ax(i)(1-x(i)). \tag{9}$$

Time series for x_i are obtained for $a = 3.5, 3.6, 3.7, 3.8, 3.9$ and 4.0 . And when $a = 3.5$ produces periodic (period four) dynamics, and $a = 3.6, 3.7, 3.8, 3.9$ and 4.0 produce chaotic dynamics with increasing complexity [29].

And the Hennon map is given by:

$$\begin{aligned} x_{i+1} &= Ry_i + 1 - 1.4x_i^2 \\ y_{i+1} &= 0.3Rx_i \end{aligned} \tag{10}$$

Time series for x_i is obtained for $R=0.8, 0.9,$ and 1.0 , with increasing complexity [30].

We firstly simulate Logistic time series at different a , and the length of each Logistic time series is 1000. Then we use the $TD_entropy$ to measure the complexity of the logistic time series, when the similar tolerance r varies from 0.1 to 0.25 with the step of 0.01. And the result is shown in Fig. 2.

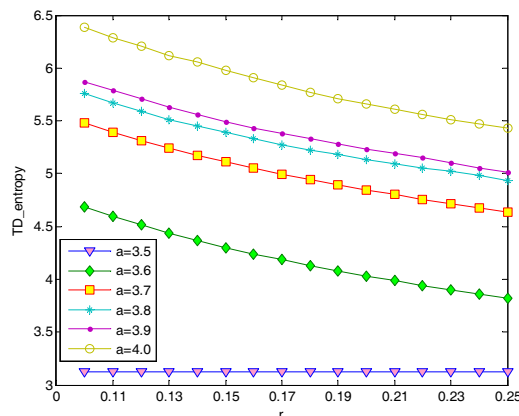


Fig. 2. $TD_entropy$ curves of different complex logistic time series with r change from 0.1 to 0.25 with step of 0.01

Through the Fig. 2, we can see the value of TD_entropy is consistent with the complexity of the Logistic time series, the bigger a has bigger TD_entropy. And TD_entropy can keep consistency when the r changing. In addition, when a is 3.5, the TD_entropy of the Logistic time series remains the same value, and the value is much smaller than the other time series. Since the Logistic time series produced by a is 3.5, the Logistic time series is periodic. Thus, the time series is more regular.

Next, we use the Hennon map time series to evaluate the consistency of the TD_entropy changing with the time attenuation coefficient k . We simulate three Hennon time series with different R , and each time series is 1000. The TD_entropy curve on the three Hennon map when the k changes from 0.01 with the step of 0.01 to 0.25. And the result is shown in Fig. 3.

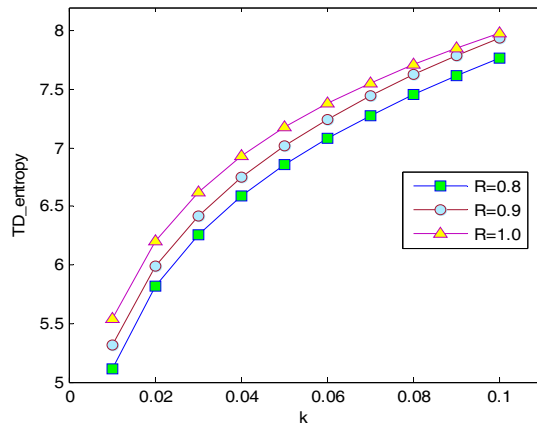


Fig. 3. The TD_entropy curve on different complexity of Hennon map when k varies from 0.01 to 0.1 with the step of 0.01

From the Fig. 3, we can see, the TD_entropy value is consistent with the real complex of Hennon map time series, and the value stays consistent when the k changing.

4.2 Comparative Study

In this subsection, we use the periodic time series, real stock time series and artificial time series to illustrate the differences between TD_entropy and SampEn.

Firstly, we use the Logistic time series obtained by a is 3.5 to display the differences on periodic time series. We simulate several Logistic time series with the length of them vary from 100 with the step of 100 to 2000. Since the Logistic time series are all periodic, so the Logistic time series have the same structural complexity, the only difference among them is the length. The TD_entropy and SampEn value on the different lengths of periodic Logistic time series is shown in Fig. 4.

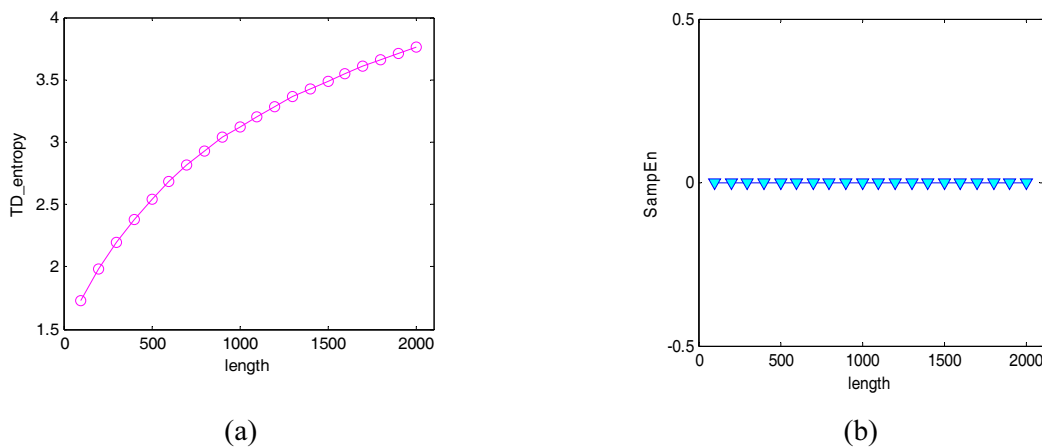


Fig. 4. The TD_entropy and SampEn curves on different length of logistic time series when a is 3.5

As is shown in Fig. 4, we can see the TD_entropy value increases with the length of Logistic time series. However, the SampEn of the different length of Logistic time series is the same, the value is 0. The results of the two methods are different.

In fact, the periodic time series are considered to be the regular ones in SampEn. And in this case the SampEn will always be zero, which means the time series are not complex. Therefore, the SampEn cannot give the complexity of the regular time series according to their structural complexity. However, through the Fig. 4(a), we can see the TD_entropy results are more reasonable, since the TD_entropy of the periodic Logistic time series is not zero, and the value becomes bigger with the length of the time series increasing. Because, the longer time series means the more data and more vectors in the time series, and the time series is more complex.

Next, we use four groups of stock time series to verify the differences between TD_entropy and SampEn. The first two groups are the closing time series shown in Fig. 5(a), and the last two groups are the volume time series shown in Fig. 5(b).

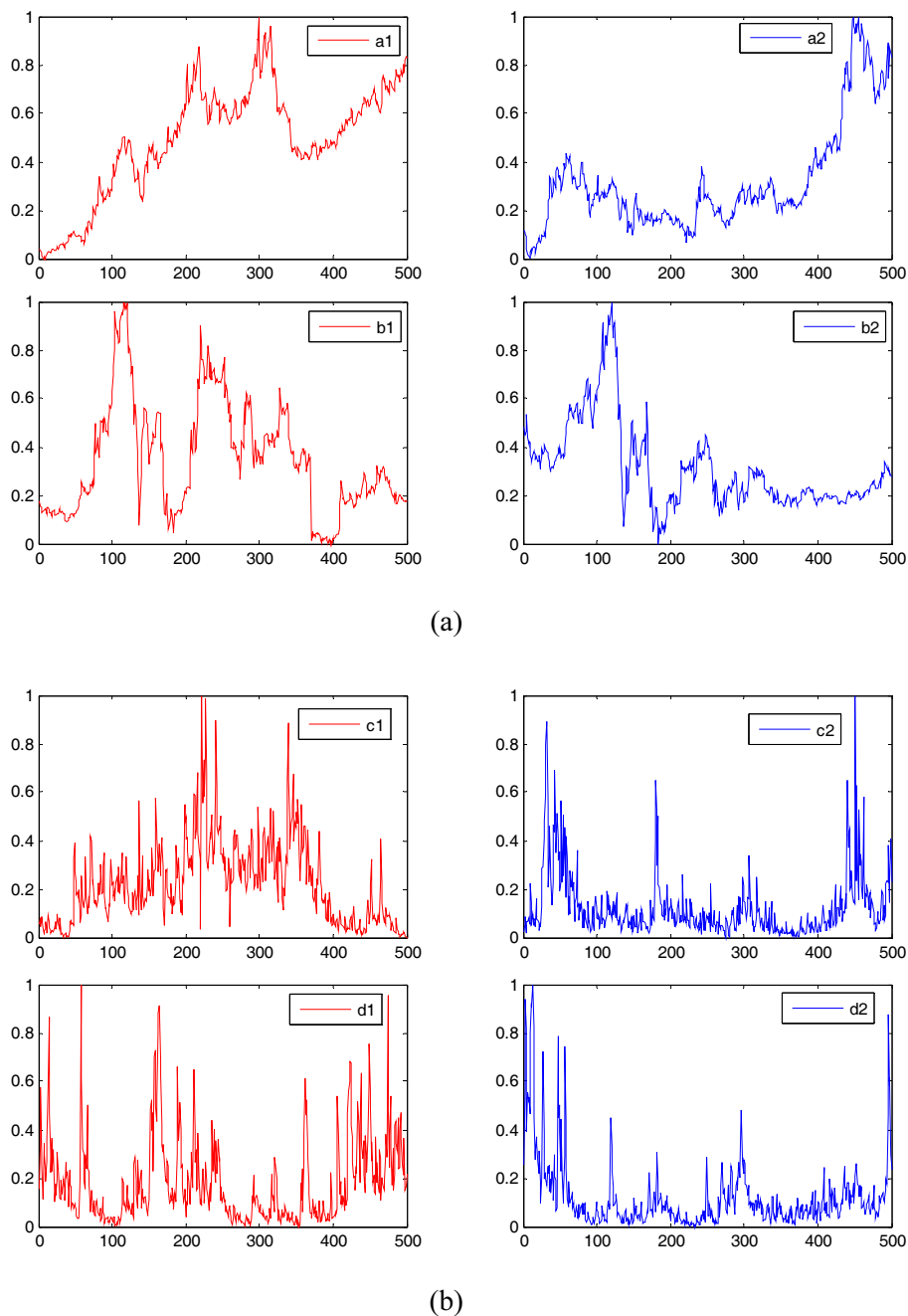


Fig. 5. Four groups of stock time series

To show the differences between TD_entropy and SampEn in detail. We calculate not only the TD_entropy value and the SampEn value, but also the self-similarity of the vectors in the time series, that are the $B^m(r, k)$ and $B^{m+1}(r, k)$ when the embedding dimension m is 2.

Table 2. The SampEn and TD_entropy results of the two groups of time series in Fig. 5(a)

	a1	a2	b1	b2
SampEn_P2	0.1009	0.1348	0.0838	0.1263
SampEn_P3	0.0779	0.0986	0.0617	0.0919
SampEn	0.2592	0.3133	0.3049	0.3183
TD_entropy_P2	0.0525	0.0670	0.0433	0.0647
TD_entropy_P3	0.0414	0.0511	0.0339	0.0504
TD_entropy	3.4198	3.2441	3.6302	3.2387

Table 2 shows the SampEn and TD_entropy results of the two groups of time series in Fig. 5(a). Through the Table 2, we can see the SampEn results indicate that the time series of a2 is more complex than that of a1, the time series of b2 is more complex than that of b1. However, the SampEn_P2 and SampEn_P3 of a2 are all bigger than that of a1, the SampEn_P2 and SampEn_P3 of b2 are all bigger than that of b1. But the TD_entropy results are different from those of SampEn.

Table 3, shows the SampEn and TD_entropy results of the two groups of time series in Fig. 5(b). Through the Table 3, we can see that from SampEn results, the time series of c2 is more complex than c1, and d2 is more complex than d1. However, like the results in Table 2, the SampEn_P2 and SampEn_P3 of the second time series in the two groups are bigger than the first one. But the complexity results of TD_entropy are different from the SampEn results.

Table 3. The SampEn and TD_entropy results of the two groups of time series in Fig. 5(b)

	c1	c2	d1	d2
SampEn_P2	0.0500	0.0843	0.0778	0.1189
SampEn_P3	0.0250	0.0334	0.0433	0.0588
SampEn	0.6927	0.9272	0.5866	0.7048
TD_entropy_P2	0.0211	0.0332	0.0313	0.0464
TD_entropy_P3	0.0106	0.0143	0.0182	0.0247
TD_entropy	5.2358	5.0896	4.5442	4.3355

Intuitively, from the Fig. 5, we can see the time series of the first one in each group with higher self-similarity probability seems more regular than that of the second in each group. In fact, the higher degree self-similarity of vectors in the time series, means the less pattern types contained in the time series, which represents less variation in the time series. In this case, the time series should be not complex. Thus, the TD_entropy is more reasonable.

Finally, through several artificial simulated time series, we illustrate another difference between TD_entropy and SampEn. As is shown in Fig. 6, are the six artificial simulated time series. To show them clearly, the length of the time series is not so long. The six time series have the same data composition, and the type and number of the vectors of every two points and every three points in the six time series are the same.

Fig. 7 shows the TD_entropy and SampEn curve of the six time series. From Fig. 7, we can see the TD_entropy of the six time series is different. However, the SampEn of them is the same.

So, what is the real complexity of the six time series? And which complexity result is more reasonable, the TD_entropy, or the SampEn? As we know, the data composition of the six time series is the same, and through the Fig. 6, we can see the the type and number of the vectors of every two points and every three points in the six time series are also the same. In this case, can we see the complexity of the six time series is the same?

Through the Fig. 6, we can see the difference between the six time series is the distribution of the vectors in the time series. And for time series, the distribution of the vectors surely can affect the complexity of the vectors, since the data or the vectors in the time series has the time attribute. The time attribute also means the sequence of each data generated on the time series. Therefore, the time attribute of the data is significant in a time series.

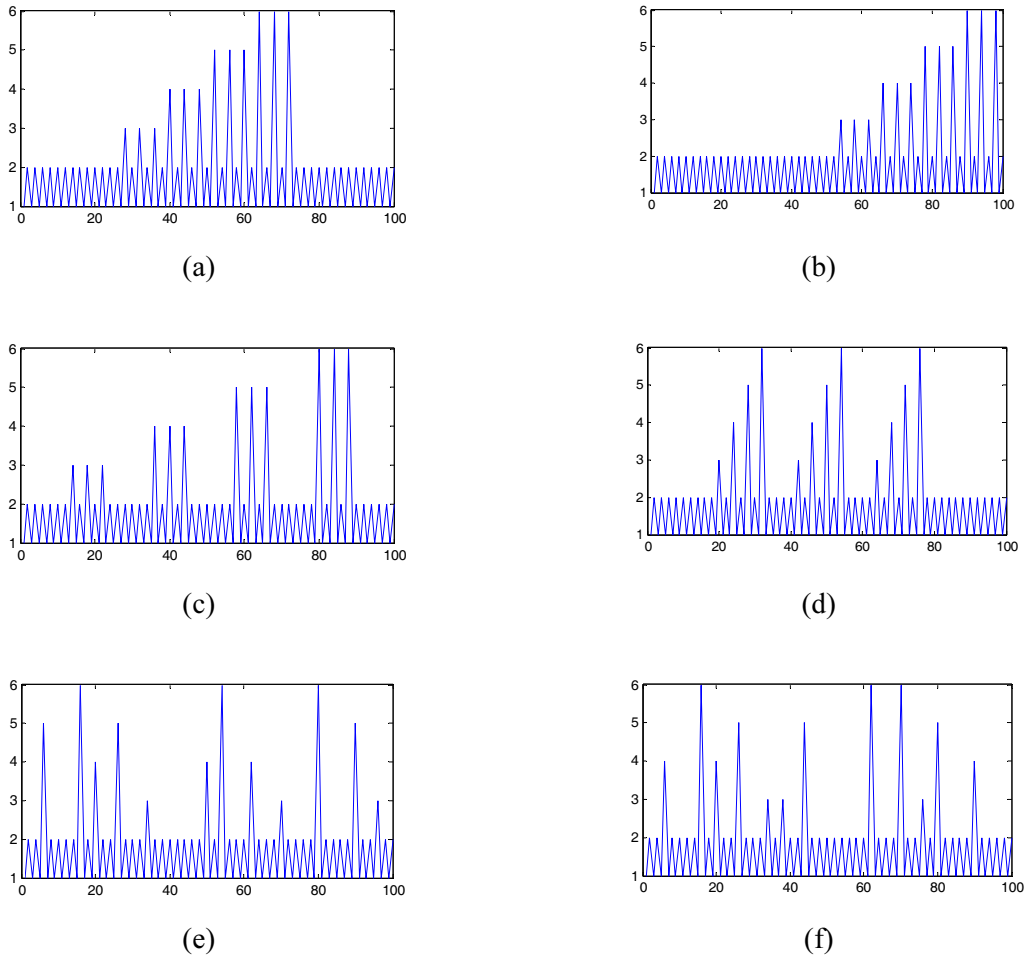


Fig. 6. Six artificial simulated time series

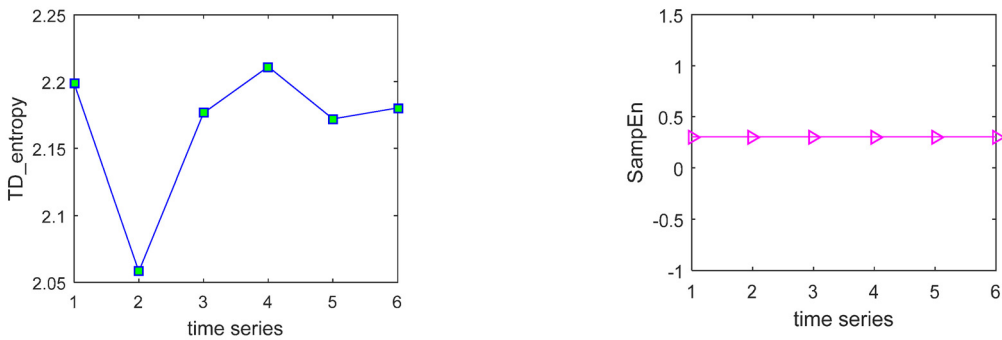


Fig. 7. The TD_entropy and SampEn curve of the six artificial simulated time series

However, though the SampEn takes the sequence of the data in the time series into account when calculating the complexity of the time series, the SampEn only takes one step difference into consideration. Because, SampEn is defined as the negative natural logarithm of the conditional probability that any two vectors of successive data points are within a tolerance r for m points remain within the tolerance at the next point [15]. Hence, sometimes, time series with the same data composition but different distributions can't be distinguished by SampEn.

However, in TD_entropy, a time decay function is used to measure the similarity degree according to the time distance between the vectors. In this case, though two vectors have the same number of similar vectors in a time series, if their distribution of the similar vectors is different, the average similarity between these two types of vectors is different. Thus, the TD_entropy of the two time series is different. Therefore, the TD_entropy is more reasonable than SampEn.

5 TD_entropy Analysis on Financial Time Series

In this section, we apply TD_entropy algorithm to financial time series to test the applicability in real world data. The daily adjusted closing values of seven stock indices from Jan 1st, 2006 to Dec 31st, 2010 are obtained from the Yahoo Finance Website. The datasets consist of DJI (Dow Jones Industrial Average, America), GSPC (S&P 500 Index, America), DAXGDAXI(Germany DAX Index, Europe), CACFCHI(France, CAC Index, Europe), SSEC(SSE Composite Index, Asia), HSI (Hang Seng Index, Asia), N225 (Nikkei 225 Index, Asia).

Since the data obtained from different markets are different in length, we choose the first 1200 data to study the complexity of the closing price time series of different stock markets during the financial crisis. Fig. 8 shows the closing price data of the seven stock indices in the period of time. From the Fig. 8, we can see the closing price in 2008 is decreasing. Then, the closing price is slowly recovery.

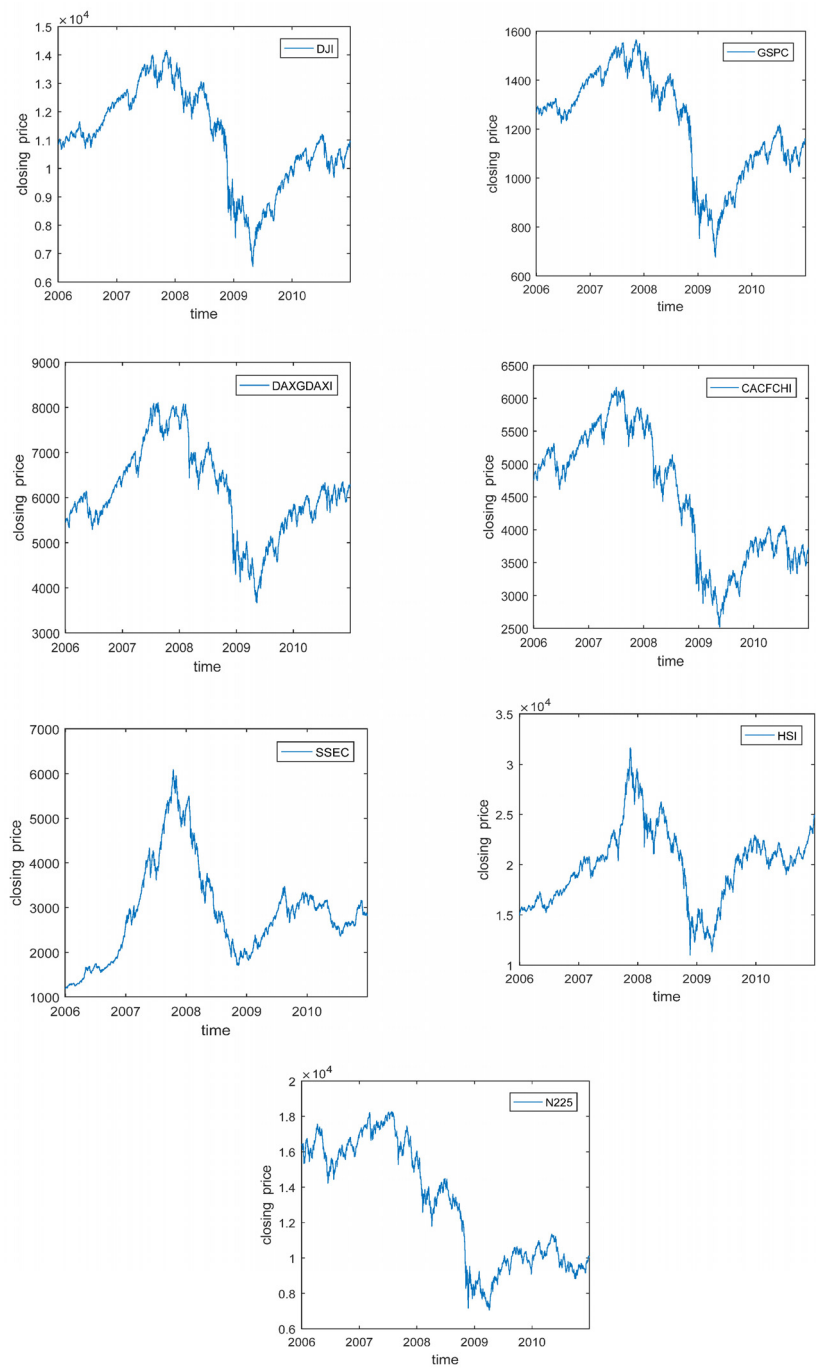


Fig. 8. The closing price of DJI and SSEC from Feb 1st, 2006 to Dec, 31st, 2010

In order to show the changes in the complexity of the closing price of the stock index, we calculate the complexity of each year of the seven stock indices, and there are 240 data contained in each year. The complexity results are shown in Fig. 9.

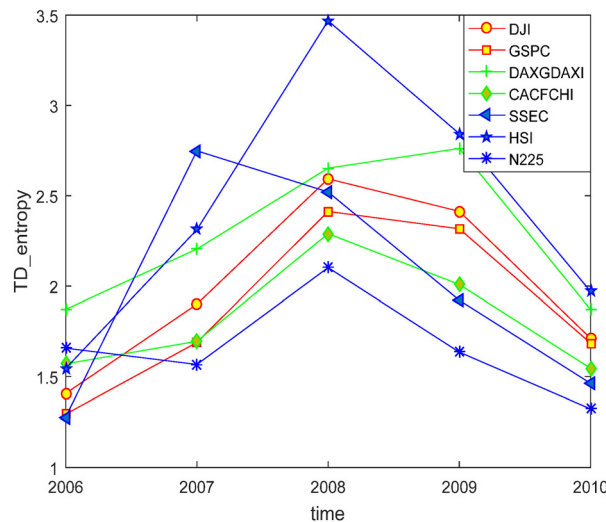


Fig. 9. The TD_entropy curve of the stock index on different period

In Fig. 9, the red ones represent the American market, the green ones represent the European markets and the blue ones represent the Asian markets. Through the Fig. 9, we can see the complexity of the financial market is relatively small before and after the financial crisis. However, in the case of the financial crisis, the stock market is unstable, the stock price fluctuates violently, and the complexity of the financial market is the biggest.

As is shown in Fig. 9, the complexity curve of American markets and European is consistent with each other. However, the stock complexity in Asian markets is different from those in American and European markets. And the complexity curve of the three stock indexes in the Asian market is also not consistent with each other.

In addition, through Fig. 8, we can see the closing price curve of DJI is different from that of SSEC. Since, the country in America and Europe are developed countries, and the country in Asia are developing countries. The environment of the financial markets in the two types of countries is also different. Therefore, the impact of financial crisis on the stability of financial markets in different type of countries is also different. Hence, at the time of the financial crisis, the complexity of their stock index changes is not the same. Moreover, the TD_entropy of HSI is the same as the American and European financial time series. This is because the stock market of Hongkong not only influenced by the US economy, but also influenced by the Chinese market. Therefore, the complexity by TD_entropy is consistent with the real complexity of the financial markets.

6 Conclusion

This paper, we proposed a modified SampEn (TD_entropy) to deal with the problem that an increase of the entropy may not always be associated with an increase in dynamical complexity. In TD_entropy, a time-decay function is used to measure the similarity degree of the similar vectors. Thus the similarity of the vectors in TD_entropy not only depends on the pattern distance of the vectors but also depend on the time distance between them. The time distance is the position distance in the time series. In addition, the self-similarity probability of the vectors is included in the model of the TD_entropy.

Through several groups of test, the TD_entropy is verified to be more reasonable and more consistent with the real dynamical complexity. In general, there are two differences between TD_entropy and SampEn. Firstly, the influence of the distribution of the similar vectors in the time series on the complexity of the time series is considered in TD_entropy. In this case, TD_entropy can distinguish more time series than SampEn, and the results are more reasonable. Secondly, for regular time series, TD_entropy can give the complexity results of different time series according to their structural

complexity; for irregular time series, the time series with higher self-similarity of vectors will have the lower TD_entropy.

Finally, the TD_entropy is applied to financial market, and the TD_entropy can get some useful information, and the complexity results are consistent with the real complexity of dynamical financial systems.

Acknowledgements

This work is supported by National Key R&D Program of China, Grant No. is 2016YFC1401900.

References

- [1] M. Xu, P. Shang, J. Huang, Modified generalized sample entropy and surrogate data analysis for stock markets, *Communications in Nonlinear Science & Numerical Simulation* 35(2016) 17-24.
- [2] A.-I. Korda, P.-A. Asvestas, G.-K. Matsopoulos, Automatic identification of eye movements using the largest lyapunov exponent, *Biomedical Signal Processing & Control* 41(2018) 10-20.
- [3] L. Telesca, C. Haro-Pérez, L.R. Moreno-Torres, A. Ramirez-Rojas, Multifractal detrended fluctuation analysis of intensity time series of photons scattered by tracer particles within a polymeric gel, *Physica A: Statistical Mechanics and Its Applications* 490(C)(2018) 994-1003.
- [4] H. Peng, W. Yang, Selection method of small current grounding fault line based on EEMD and correlation dimension. <http://en.cnki.com.cn/Article_en/CJFDTotal-IKJS201701030.htm>, 2017.
- [5] H. Zhang, Z. Feng, J. Zou, Research on feature extraction and pattern recognition of acoustic signals based on MEMD and approximate entropy, in: *Proc. 29th Chinese Control and Decision Conference*, 2017.
- [6] J.-S. Richman, J.-R. Moorman, Physiological time-series analysis using approximate entropy and sample entropy, *American Journal of Physiology Heart & Circulatory Physiology* 278(6)(2000) 2039.
- [7] M. Costa, A.-L. Goldberger, C.-K. Peng, Multiscale entropy analysis of complex physiologic time series, *Physical Review Letters* 89(6)(2002) 068102.
- [8] Y. Zhang, P. Shang, Permutation entropy analysis of financial time series based on Hill's diversity number, *Communications in Nonlinear Science & Numerical Simulation* 6(1)(2017) 1659-1671.
- [9] R.-K. Udhayakumar, C. Karmakar, M. Palaniswami, Understanding irregularity characteristics of short-term HRV signals using sample entropy profile, in: *Proc. IEEE Transactions on Biomedical Engineering*, 2018.
- [10] D. Shang, M. Xu, P. Shang, Generalized sample entropy analysis for traffic signals based on similarity measure, *Physica A Statistical Mechanics & Its Applications* 474(2017) 1-7.
- [11] Y. Wu, P. Shang, Y. Li, Modified generalized multiscale sample entropy and surrogate data analysis for financial time series, *Nonlinear Dynamics* 92(3)(2018) 1335-1350.
- [12] Y. Song, J. Zhang, Discriminating preictal and interictal brain states in intracranial EEG by sample entropy and extreme learning machine, *Journal of Neuroscience Methods* 257(2016) 45-54.
- [13] C. Mayer, M. Bachler, M. Hörtenhuber, C. Stocker, A. Holzinger, S. Wassertheurer, Selection of entropy-measure parameters for knowledge discovery in heart rate variability data. <<https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-15-S6-S2>>, 2014.

- [14] L. Zhao, S. Wei, C. Zhang, Y. Zhang, X. Jiang, F. Liu, C. Liu, Determination of Sample entropy and fuzzy measure entropy parameters for distinguishing congestive heart failure from normal sinus rhythm subjects, *Entropy* (17)(2015) 6270-6288.
- [15] C. Mayer, M. Bachler, A. Holzinger, P.K. Stein, S. Wassertheurer, The effect of threshold values and weighting factors on the association between entropy measures and mortality after myocardial infarction in the cardiac arrhythmia suppression trial (CAST), *Entropy* 18(4)(2016) 129.
- [16] P. Marwaha, R.-K. Sunkaria, Complexity quantification of cardiac variability time series using improved sample entropy (I-SampEn), *Australasian Physical & Engineering Sciences in Medicine* 39(3)(2016) 755-763.
- [17] F.-A. Pujol, M. Pujol, A. Jimeno-Morenilla, M.J. Pujol, Face detection based on skin color segmentation using fuzzy entropy, *Entropy* 19(1)(2017) 1-22.
- [18] Z. Cao, C.-T. Lin, Inherent fuzzy entropy for the improvement of EEG complexity evaluation, *IEEE Transactions on Fuzzy Systems* 26(2)(2018) 1032-1035.
- [19] S. Koppel, S.-I. Chang, A process capability analysis method using adjusted modified sample entropy, *Procedia Manufacturing* 5(2016) 122-131.
- [20] F. Liao, Y.-K. Jan, Using modified sample entropy to characterize aging-associated Microvascular dysfunction. <<https://www.frontiersin.org/articles/10.3389/fphys.2016.00126/full>>, 2016.
- [21] Y. Liu, J. Wang, J. Wang, P. Shang, Refined generalized multiscale entropy analysis for physiological signals, *Physica A Statistical Mechanics & Its Applications* 490(2018) 975-985.
- [22] Y. Wu, P. Shang, Y. Li, Multiscale sample entropy and cross-sample entropy based on symbolic representation and similarity of stock markets, *Communications in Nonlinear Science & Numerical Simulations* 56(2018) 49-61.
- [23] W. Shi, P. Shang, Y. Ma, S. Sen, C.H. Yeh, A comparison study on stages of sleep: quantifying multiscale complexity using higher moments on coarse-graining, *Communications in Nonlinear Science & Numerical Simulation* (44)(2016) 292-303.
- [24] Z.-M. Zhong, Y. Hu, C.-H. Li, Z.T. Liu, Discovering similar users for specific user on Microblog, *Chinese Journal of Computers* 39(4)(2016) 65-778.
- [25] W. Li, M. He, L.-H. Wang, Y. Liu, H.-W. Shen, X.-Q. Chen, Research on Microblog Retweeting prediction based on user behavior features, *Chinese Journal of Computers* 39(10)(2016) 1992-2006.
- [26] L. Feng, Z. Y. Peng, B. Liu, D. Chen., A latent-citation-network based patent value evaluation method, *Journal of Computer Research and Development* 52(3)(2015) 649-660.
- [27] Z. Liu, Y. Yang, J.-P. Zhang, J. Yang, Y. Chu, Z. Zhang, An adaptive grid-density based data stream clustering algorithm based on uncertainty model, *Journal of Computer Research and Development* 51(11)(2014) 2518-2527.
- [28] G.-H. Li, H. Chen, Mining the frequent patterns in an arbitrary sliding window over online data streams, *Journal of Software* 19(19)(2008) 2585-2596.
- [29] W. Chen, J. Zhuang, W. Yu, Z. Wang, Measuring complexity using FuzzyEn, ApEn, and SampEn. *Medical Engineering & Physics* 31(1)(2009) 61-68.
- [30] H.-B. Xie, W. He, X. Liu, Measuring time series regularity using nonlinear similarity-based sample entropy, *Physics Letters A* 372(48)(2008) 7140-7146.