Nonlinear Dyn (2021) 104:1737–1751
https://doi.org/10.1007/s11071-021-06340-3

ORIGINAL PAPER

Directed vector visibility graph from multivariate time series: a new method to measure time series irreversibility

Binbin Shang · Pengjian Shang

Received: 2 December 2020 / Accepted: 1 March 2021 / Published online: 13 March 2021

Abstract As a practical tool, visibility graph provides a different perspective to characterize time series. In this paper, we present a new visibility algorithm called directed vector visibility graph and combine it with the Kullback–Leibler divergence to measure the irreversibility of multivariable time series. T directed vector visibility algorithm converts the time series into a directed network. Subsequently, the ingoing and outgoing degree distributions of the directed network can be got to calculate the Kullback–Leibler divergence, which will be applied to assess the level of irreversibility of the time series. This is a simple and effective method without any special symbolic process. The numerical results from various types of systems are used to validate that this method can accurately distinguish reversible time series from those irreversible ones. Finally, we employ this method to estimate the irreversibility of financial time series and the results show that our method is efficient to analyze the financial time series irreversibility.

Keywords Multivariate time series · Directed vector visibility graph · Kullback–Leibler divergence · Multiscale

B. Shang (✉) · P. Shang
School of Science, Beijing Jiaotong University, Beijing 100044, People's Republic of China
e-mail: 202115605@qq.com

1 Introduction

For a time series $X(t)$, if the series $\{X(t_1),\ldots,X(t_N)\}$ and $\{X(t_N),\ldots,X(t_1)\}$ possess the identical joint probability distribution for any N, in other words, if its statistical properties will not change with the reversal of time, this time series will be regarded as reversible time series [1]. Statistically speaking, the reversible time series has the same probability as its reversed time series. From the perspective of physics, the second law of thermodynamics defines the unidirectionality of time for the first time. Time irreversibility means that when time is reversed, the system cannot return to the past state. The stationary transformations of some nonlinear sequences, Gaussian linear processes and Fourier transform substitutions of Gaussian processes all belong to reversible processes. On the contrary, the irreversibility of time series means that the given dynamic system has nonlinear properties, which is related to dissipative chaos and non-Gaussian stochastic processes [2, 3].

In the past few decades, many different irreversible measures have been proposed [4–15]. However, on the one hand, the reversibility test of time series can merely analyze the irreversibility of time series qualitatively rather than quantitatively. On the other hand, most of the studies mainly focus on the irreversibility of time series in the low-dimensional phase space at a single scale. Therefore, researchers have gradually proposed some statistics which can be used to quantitatively measure irreversibility and extended the irreversible measurement index to high-dimensional and multiscale analysis [16–22]. At present, researchers believe that the time series irreversibility can reflect the dynamic characteristics of the system and the directionality of time series; they also find that the irreversibility is related to physical dissipation [23, 24]. An effective method to characterize it is the Kullback–Leibler divergence (KLD)[25]. Later, some methods, which measure time series irreversibility by using the difference of probability distribution between positive and inverse order, were proposed one after another [26–29]. Lacasa et al. mapped the time series into a network and applied *KLD* to estimate the irreversibility of the univariate time series [30]. By comparing the degree distributions of positive and inverse time series, the level of irreversibility of this time series was reflected. Theoretically, the visibility graph provides a new method to characterize time series and this method does not need to set up the symbolic conditions in advance. It has been extended and applied to the financial field [31, 32], fluid dynamics [33, 34] and medical research [35]. However, this method is not suitable for multivariate time series.

Since the algorithm of mapping a time series into a complex network and using graph theory to explore the characteristics of time series, most researches have focused on the analysis of univariate time series. What is exciting is that Ren et al. [36] proposed the vector visibility graph (VVG) for multivariable time series, which provides a way to transform the multivariable time series into a directed complex network. As an effective method to convert a multivariate time series into a graph, *VVG* is a practical tool to analyze multivariate time series from the perspective of graph theory. The multivariate time series are mapped into a directed complex network, while each multidimensional data vector is regarded as one node and the visibility between the corresponding data vectors determines the connection of the network. According to these facts, we propose the directed vector visibility algorithm. The multivariate time series are mapped into a graph, and the properties of the association graph are analyzed. More accurately, we apply the *KLD*, which is calculated by the *ingoing* and *outgoing* degree distributions of the time series, to measure the time series irreversibility. Based on the numerical results, we confirm that it is a convenient and powerful method to measure the irreversibility of time series.

The rest of the paper is arranged as follows. Section 2 shows the methods of the directed vector visibility graph, the *KLD* and provides a simple proof of this method for uncorrelated stochastic series. Later, we introduce the multiscale method and give the definition of KLD_τ. In Sect. 3 , we apply the new proposed method to analyze several different classes of processes and verify its validity. Section 4 first introduces some other statistics and then displays the practical application of financial time series. Finally, the conclusions are given in Sect. 5 .

2 Methodology

2.1 2.1 Directed vector visibility graph

Visibility algorithm family is a set of methods which convert time series into networks on the basis of geometric criteria [37, 38]. The principle of these methods is to map the information contained in time series into another mathematical structure, so that the effective tools of graph theory can be applied to describe time series from the different angle.

Here, we use the algorithms of horizontal visibility graph [38] and vector visibility graph [36] for reference and introduce the new method, which is defined as follows:

For a m-dimensional time series $X_t = \{x_t^i\}_{i=1}^m$ and the length of each dimension is N, map the multivariate time series into a vector space, then we will gain a sequence of vectors $\{\vec{X}_t\}$, where $\vec{X}_t = [x_t^1, x_t^2, \ldots x_t^m]$. For any two vectors ($\vec{X}_a$ and $\vec{X}_b$) in the vector sequence, the projection from $\vec{X}_a$ to $\vec{X}_b$ is defined as follows:

$$\left|\left|\vec{X}_b^a\right|\right| = \frac{\sum_{i=1}^m x_a^i x_b^i}{\sqrt{\sum_{i=1}^m x_a^i x_a^i}} \tag{2.1}$$

and $\left|\left|\vec{X}_a\right|\right| = \sqrt{\sum_{i=1}^m x_a^i x_a^i}$. Each vector in the vector sequence is regarded as one node in the network, and the visibility criteria for vectors can be shown as follows:

Any two vectors $\vec{X}_a$ and $\vec{X}_b$ will he the directed visibility from $\vec{X}_a$ to $\vec{X}_b$, if the arbitrary vector $\vec{X}_c$ situated between them fulfills:

$$\left|\left|\vec{X}_a\right|\right|, \left|\left|\vec{X}_b^a\right|\right| > \left|\left|\vec{X}_c^a\right|\right| \tag{2.2}$$

where $t_a < t_c < t_b$, $\left|\left|\vec{X}_b^a\right|\right|$ and $\left|\left|\vec{X}_c^a\right|\right|$ is the projection from $\vec{X}_b$ and $\vec{X}_c$ to $\vec{X}_a$. Then, we obtain a directed link from the node standing for $\vec{X}_a$ to the node standing for $\vec{X}_b$ in the network. Therefore, the directed complex network named directed vector visibility graph (DVV_g) can be defined. The number of connections of node t linked to other past nodes $t^{'}(t' < t)$ is expressed by the *ingoing* degree $k_{\text{in}}(t)$. On the contrary, the number of connections of node t linked to other future nodes $t'(t'' > t)$ is expressed by the *outgoing* degree $k_{\text{out}}(t)$. The degree $k(t)$ of the node t includes these two parts: the *ingoing* degree $k_{\text{in}}(t)$ and the *outgoing* degree $k_{out}(t)$, that is to say, $k(t) = k_{in}(t) + k_{out}(t)$.

Figure 1 displays the procedures from the multivariate time series to its DVV_g. The degree distribution of the DVV_g refers to the probability of any node with degree k. The *outgoing* and *ingoing* degree distributions of a DVV_g are, respectively, defined as the probability distributions of k_{out} and k_{in}, where $P_{\text{out}}(k) \equiv P(k_{\text{out}} = k)$ and $P_{\text{in}}(k) \equiv P(k_{\text{in}} = k)$.

2.2 Kullback–Leibler divergence via DVV_g

The premise of the work is that the information contained in the *ingoing* and *outgoing* degree distributions can indicate the irreversibility of time series. More accurately, this can be determined by the distance between the *ingoing* and *outgoing* degree distributions in the first-order approximation.

The distance between the *ingoing* and *outgoing* degree distributions can be shown by the Kullback–Leibler divergence (KLD)[25]. In information theory, KLD (or known as relative entropy) is proposed as an asymmetric measure of the difference between two probability distributions. Given the random variable x with two probability distributions $p(x)$ and $q(x)$, KLD with $p(x)$ and $q(x)$ is given as follows:

$$\text{KLD(pq)} = \sum_{x \in X} p(x) \log \frac{p(x)}{q(x)} \tag{2.3}$$

which is equal to zero if and only if the two probability distributions $p(x)$ and $q(x)$ are equal. Otherwise, the KLD is greater than zero.

Theoretically, the information contained in the *outgoing* degree distribution k_{out} is enough to distinguish the reversible time series from the irreversible ones. The probability corresponding to the *outgoing* degree distribution of the time-reversed time series is equal to the probability corresponding to the *ingoing* degree distribution of the actual process, that is to say, $P_{k_{\text{out}}}\left(k\middle|\{X(t)\}_{t=N,\ldots,1}\right) = P_{k_{\text{in}}}\left(k\middle|\{X(t)\}_{t=1,\ldots,N}\right)$. The KLD between $P_{\text{out}}(k)$ and $P_{\text{in}}(k)$ is written as:

$$D[P_{out}(k)P_{in}(k)] = \sum_k P_{out}(k) \log \frac{P_{out}(k)}{P_{in}(k)} \tag{2.4}$$

KLD is equal to zero, if and only if, the probability distribution of both the *ingoing* and the *outgoing* degrees of the series is the same, i.e., $P_{\text{in}}(k) = P_{\text{out}}(k)$. Otherwise, it is positive. The time series is reversible if and only if $P_{k_{\text{in}}}\left(k\middle|\{X(t)\}_{t=1,\ldots,N}\right) = P_{k_{in}}\left(k\middle|\{X(t)\}_{t=N,\ldots,1}\right) = P_{k_{out}}\left(k\middle|\{X(t)\}_{t=1,\ldots,N}\right).$, that is, the distribution of *ingoing* degree is the same as that of *outgoing* degree. In other words, if the KLD between the *ingoing* and *outgoing* degree distributions gradually inclines to zero with the increase in series size, it means that the time series is reversible. However, if the KLD converges to a finite positive value, the time series is considered to be irreversible. Different from other measures applied to evaluate the irreversibility of time series [27, 39–41], the KLD has the statistical significance. Concretely, KLD is a measure of "distinguishability." The more distinguishable $P_{\text{out}}(k)$ and $P_{\text{in}}(k)$ are from each other, the more the KLD deviates from 0, which means that the time series is more irreversible. Therefore, we use the value of KLD to reflect the degree of irreversibility of the time series.

Most of the previous methods for estimating the irreversibility of time series generally started with a local symbolization of the sequence, and the occurrences of word from the forward- and reverse-symbolized series are statistically analyzed [42, 43]. As a result, the irreversibility of time series is related to the difference between the word statistics of the forward- and reverse-symbolized series. If we only take advantage of the information contained in the series $\{k_{\text{out}}(t)\}_{t=1,\ldots,N}$ and $\{k_{\text{in}}(t)\}_{t=1,\ldots,N}$, KLD can also be regarded as a symbolization. Nevertheless, unlike other methods, this method does not need

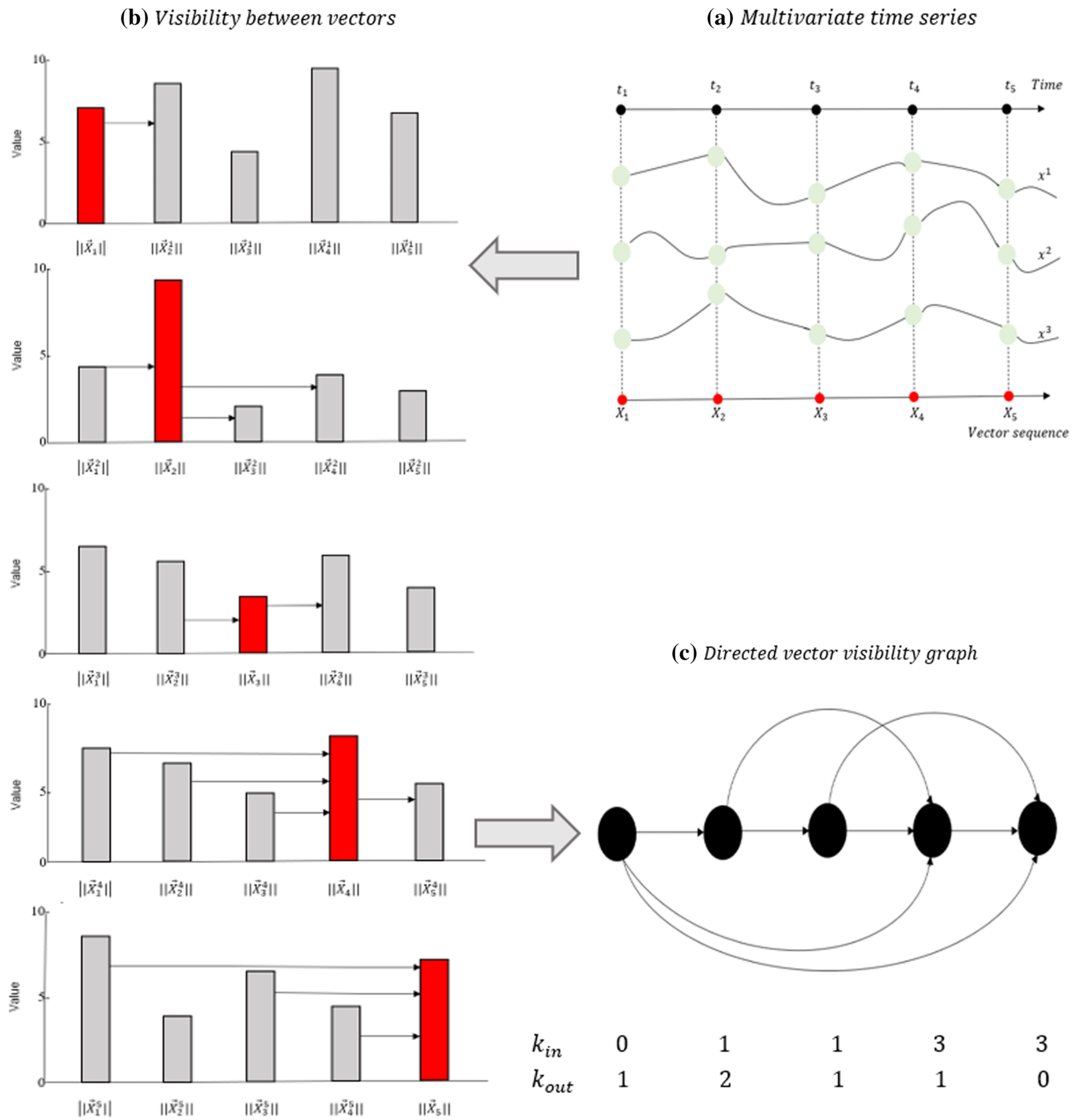


Fig. 1 Procedures from the multivariate time series ($m = 3, N = 5$) to a directed vector visibility graph. **a** Each vector corresponds to a node after mapping the time series into the vector space. **b** The link from a node to others is determined on the basis of the visibility criteria between vectors. **c** The corresponding directed vector visibility graph

specific parameters and considers the global information. From a statistical mechanics point of view, KLD can not only determine the irreversibility of the time series obtained from the non-equilibrium processes, but also can be used to measure its average entropy production [2, 17, 44–46].

Next, we give a simple proof of our proposed algorithm. Here, we only show the *ingoing* and *outgoing* degree distributions obtained from the uncorrelated stochastic series and confirm that they are equal under the condition of infinite size series.

Theorem *Let $X_t = \{x_t^i\}_{i=1,t=-\infty,\ldots,\infty}^{m}$ be a bi-infinite vector sequence of independent and identically distributed random variables obtained from the continuous probability density $f(x^1, x^2, \ldots, x^m)$. Then, the distribution of both the ingoing and the outgoing degrees of its* DVV_g *are.*

$$P_{\text{in}}(k) = P_{\text{out}}(k) = (1/2)^k, k = 1, 2, 3, \ldots \tag{2.5}$$

Proof For the convenience of proof, we choose $m = 3$. Let X_0 be an arbitrary datum. The probability that the vector visibility of X_0 is interrupted by the datum X_1 on its right is independent of $f(x^1, x^2, x^3)$,

$$\Phi_1 = \int\int\int_{-\infty}^{\infty} f\left(x_0^1, x_0^2, x_0^3\right) dx_0$$

$$\int\int \int_{x_0^1x_1^1+x_0^2x_1^2+x_0^3x_1^3 > \sum_{i=1}^{3}\left(x_0^i\right)^2}^{\infty} f\left(x_1^1, x_1^2, x_1^3\right) dx_1$$

$$= \int\int\int_{-\infty}^{\infty} f\left(x_0^1, x_0^2, x_0^3\right)\left[1 - F\left(x_0^1, x_0^2, x_0^3\right)\right] dx_0 = \frac{1}{2},$$

where

$$F\left(x_0^1, x_0^2, x_0^3\right) = \int\int\int_{-\infty}^{\sqrt{\sum_{i=1}^{3}\left(x_0^i\right)^2}} f\left(x^1, x^2, x^3\right) \mathrm{d}x$$

The probability $P(k)$ of the datum X_0 being able to see k data accurately can be established as

$$P(k) = Q(k)\Phi_1 = \frac{1}{2}Q(k), \tag{2.6}$$

where $Q(k)$ is the probability of X_0 at least seeing k data. $Q(k)$ can be recurrently computed by

$$Q(k) = Q(k-1)(1 - \Phi_1) = \frac{1}{2}Q(k-1), \tag{2.7}$$

Because the arbitrary datum X_0 can see at least the first adjacent node to its right, $Q(1) = 1$. The following expression can be got

$$Q(k) = \left(\frac{1}{2}\right)^{k-1}, \tag{2.8}$$

which together with Eq. (2.6) derives the proof. A similar derivation makes available for the *ingoing* case.

It is worth noting that the above result is not affected by the probability density $f(x^1, x^2, \ldots, x^m)$. It holds not only for Gaussian or uniform distribution time series, but also for arbitrary independent and identically distributed random variables with a continuous distribution $f(x^1, x^2, \ldots, x^m)$.

2.3 KLD_τ of multivariate multiscale time series

In order to better understand the intrinsic characteristics of multivariate system, we introduce the multiscale method. Given a m-dimensional time series $X_t =$ with the length of each dimension equaling to N, we define KLD_τ as follows:

Step 1 For the scale factor τ, the original time series is divided into non-overlapping windows of length τ[14]. We gain the coarse-grained m-dimensional time series $Y_k^\tau = y_{i,k,i=1}^{\tau m}$ as

$$y_{j,k}^\tau = \frac{1}{\tau}\sum_{t=(k-1)\tau+1}^{k\tau} x_t^i, 1 \le k \le \frac{N}{\tau} \tag{2.9}$$

Step 2 Get the DVV_g transformed from the coarse-grained m-dimensional time series Y_k^τ.

Step 3 Compute the KLD_τ by the *ingoing* and *outgoing* degree distributions corresponding to the DVV_g.

3 Analyses and results of synthetic data

In this section, we choose several types of systems: uncorrelated stochastic series, correlated stochastic series, dissipative chaotic systems and conservative chaotic systems, to evaluate the degree of irreversibility of their multivariate time series by our new proposed method and Table 1 contains the numerical results.

(a) Uncorrelated stochastic series
(b) Three-dimensional random series

Here, we generate a trivariate time series, where all the data channels are observations of mutually independent time series extracted from the uniform distribution $U[0, 1]$.

Table 1 Values of the *KLD* corresponding to time series of 0.7×10^5 data generated by reversible and irreversible processes

Series description	KLD*[$P_{out}(k)$\\$P_{in}$tk)]*
Reversible stochastic processes	
i/[0_pl] uncorrected	$1.8517 \times 10^{“4}$
A/{0,0,1,2,0.8)	$2.3214 \times 10^{”4}$
Dissipative chaos	
Chen system	0.1463
Duffing system	0.3217
Holmes–Duffing system	0.4983
Lorenz system	0.6162
L *ü*system	0.4217
R*ö*ssler system	1.0711
Conservative chaos	
Sprott-A system	$2.2723 \times 10^{“4}$

(b) Correlated stochastic series

(c) Two-dimensional time series generated by normal distribution

In order to obtain the correlated stochastic series, we consider the bivariate normal distribution $N(0, 0, 1, 2, 0.8)$ as an example of linearly correlated stochastic processes.

(c) Dissipative Chaotic systems

(3) Chen system [47]

$$\begin{cases} x' = a(y - x) \\ y' = (c - a)x - xz + cy \\ z' = xy - bz \end{cases} \tag{3.1}$$

Parameter values: $a = 35, b = 3$ and $c = 28$;
initial conditions: $x_0 = 0, y_0 = 1.001$ and $z_0 = 0$.

(4) Duffing system [48]

$$\begin{cases} x' = y \\ y' = -x - x^3 - ky + f \cos z \\ z' = 1 \end{cases} \tag{3.2}$$

Parameter values: $k = 0.1$ and $f = 80$;
initial conditions: $x_0 = 0, y_0 = 0$ and $z_0 = 1$.

(5) Holmes–Duffing system [48]

$$\begin{cases} x' = y \\ y' = x - x^3 - ky + f \cos z \\ z' = 1 \end{cases} \tag{3.3}$$

Parameter values: $k = 0.1$ and $f = 80$;
initial conditions: $x_0 = 0, y_0 = 0$ and $z_0 = 1$.

(6) Lorenz system [49]

$$\begin{cases} x' = s(y - x) \\ y' = rx - y - xz \\ z' = xy - bz \end{cases} \tag{3.4}$$

Parameter values: $s = 10, r = 28$ and $b = 8/3$;
initial conditions: $x_0 = 10, y_0 = 1$ and $z_0 = 0$.

(7) L *ü* system [50]

$$\begin{cases} x' = a(y - x) \\ y' = -xz + cy \\ z' = xy - bz \end{cases} \tag{3.5}$$

Parameter values: $a = 36, b = 3$ and $c = 20$;
initial conditions: $x_0 = 0, y_0 = 1.001$ and $z_0 = 0$.

(8) R *ö* ssler system [51]

$$\begin{cases} x' = -y - z \\ y' = x + ay \\ z' = b + z(x - c) \end{cases} \tag{3.6}$$

Parameter values: $a = 0.2, b = 0.4$ and $c = 5.7$;
initial conditions: $x_0 = 1, y_0 = 0$ and $z_0 = 0$

(d) Conservative Chaotic systems

(9) Sprott-A system [52]

$$\begin{cases} x' = y \\ y' = -x + yz \\ z' = 1 - y^2 \end{cases} \tag{3.7}$$

initial conditions: $x_0 = 0.1, y_0 = 0.1$ and $z_0 = 0.1$.

Figure 2 shows the *ingoing* and *outgoing* degree distributions of the DVV_g corresponding to the multivariate time series with the size $N = 0.7 \times 10^5$. We can discover that expect the time series from the Lorenz system, their *ingoing* and *outgoing* degree distributions are almost indistinguishable. And their specific numerical values of *KLD* are given in Table 1

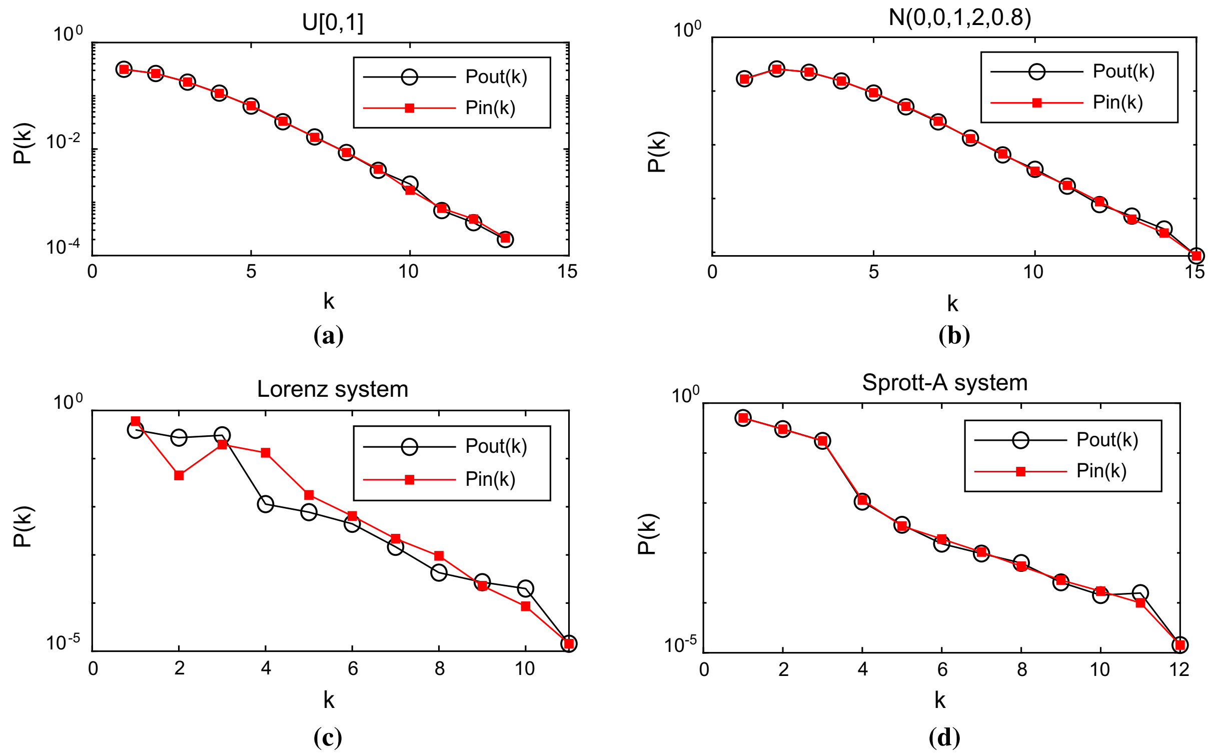


Fig. 2 The *ingoing* and *outgoing* degree distributions of the DVV*g* corresponding to the multivariate time series of 0.7×10^5 data points from: **a** the uniform distribution $U[0,1]$; **b** the two-dimensional normal distribution $N(0,0,1,2,0.8)$; **c** the Lorenz system; **d** the Sprott-A system .

and each of them is very close to 0. It indicates that these time series, which are generated from uncorrelated stochastic series, correlated stochastic series and conservative chaotic systems, respectively, are all reversible time series. And those time series generated from dissipative chaotic systems are irreversible.

The change of *KLD* corresponding to different multivariate time series with gradually increasing series size N is plotted in Fig. 3. With the increase in series size, the *KLD* of reversible time series tends to zero, while the *KLD* of irreversible time series converges to a positive value. Therefore, we can say that the deviation between *KLD* of reversible time series and zero is caused by the finite size effect.

In fact, the research on the influence of series length to *KLD* is similar to the multiscale analysis of time series, but the structure of the coarse-grained time series may be different from that of the original time series. Therefore, the KLD_τ of coarse-grained sequence obtained from the original sequence on a large scale is different from that of a short sequence segment of the real original sequence. Figure 4 exhibits the KLD_τ of nine simulation series with the length of $N = 10^4$ on scale $\tau \in [1, 20]$. For uncorrelated stochastic series, correlated stochastic series and conservative chaotic systems, their KLD_τ increases slightly with the increase in scale τ, respectively, but each of them is still very close to 0 on any scale. For dissipative chaotic systems, their KLD_τ, respectively, shows the downward trend with the increase in scale τ and tends to 0. However, the change trend of KLD_τ with the increase in scale τ is different from that obtained from simply shortening the length of time series. Therefore, we can consider that the coarse-grained process changes the internal structure of the system, which may have an influence on the degree of its internal irreversibility. Nevertheless, for the reversible time series, even if the coarse-grained process changes the internal structure of the system, the coarse-grained sequence is still in a random state, so it is still reversible.

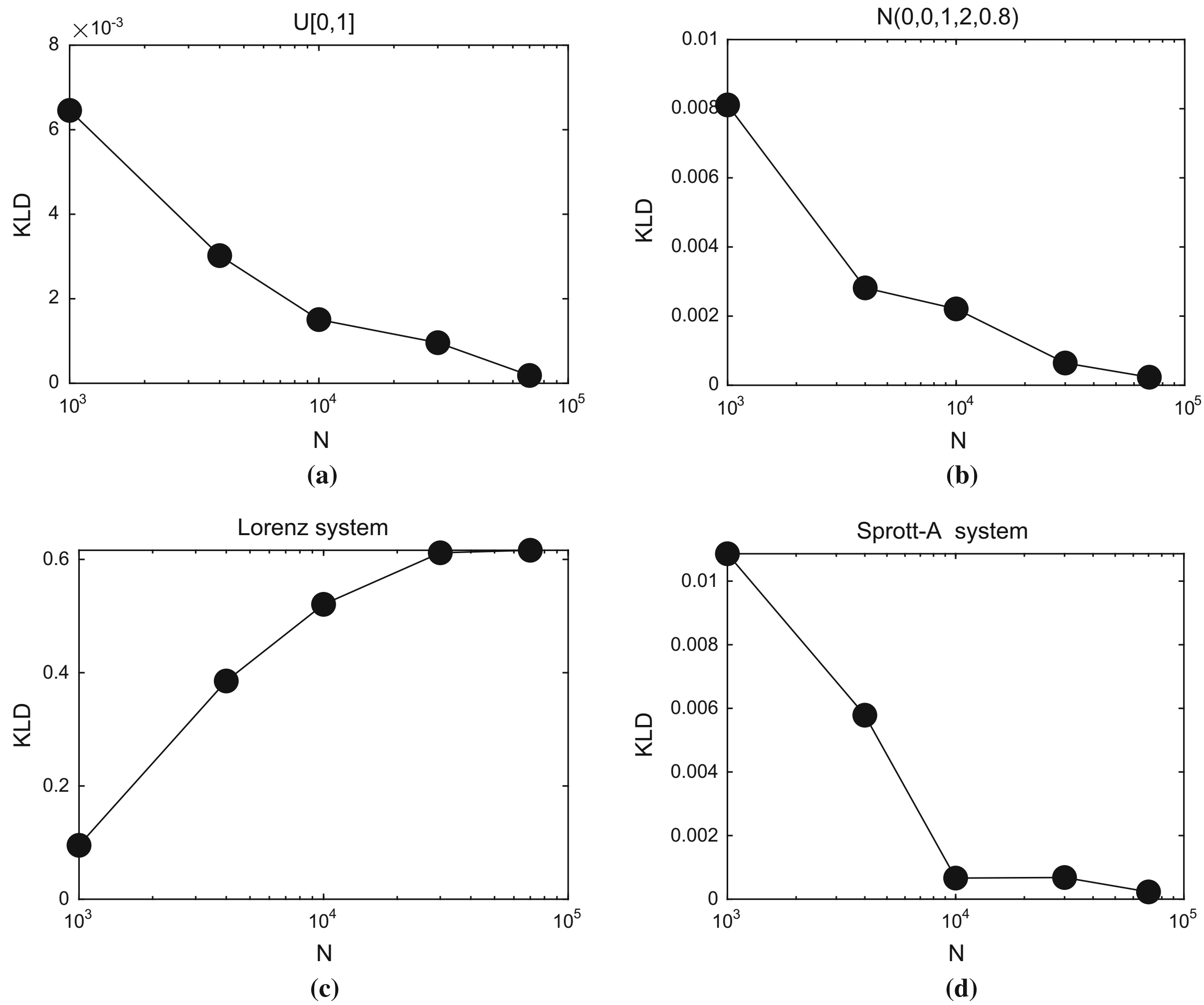


Fig. 3 Semi-log plot of KLD of the graph corresponding to the multivariate time series as a function of the series size N (points are the average of several realizations). **a** the uniform distribution$U[0,1]$; **b** the two-dimensional normal distribution $N(0,0,1,2,0.8)$;**c** the Lorenz system; **d** the Sprott-A system .

4 Analyses and results of financial time series

Stock markets reflect the development of the national economy and also display the economic development and social stability of countries. The commonly used parameters to reflect the fluctuation of stock market are stock trading price and stock trading volume. Here, we use the proposed new irreversible method to analyze the financial time series. The daily closing price and volume of twenty-one stock indices from 2005 to 2019 are gathered from the Web site https://finance.yahoo.com/, and these stock indices are divided into three different regions: Americas, Europe and Asia & Pacific. The exact information of 21 stock indices is shown in Table 2.

Because of the non-stationarity of financial time series, we remove the unwanted data firstly and then use logarithmic price difference as proxies for volatility of stock indices, which is given by

$$x_n = \log(S_n) - \log(S_{n-1}) \tag{4.1}$$

where S_n is the closing price of n th trading day.

For volume, we firstly standardize the data to maintain data consistency. Next, we also select the logarithmic volume difference to handle the standardized data for the sake of weakening the diversity between them.

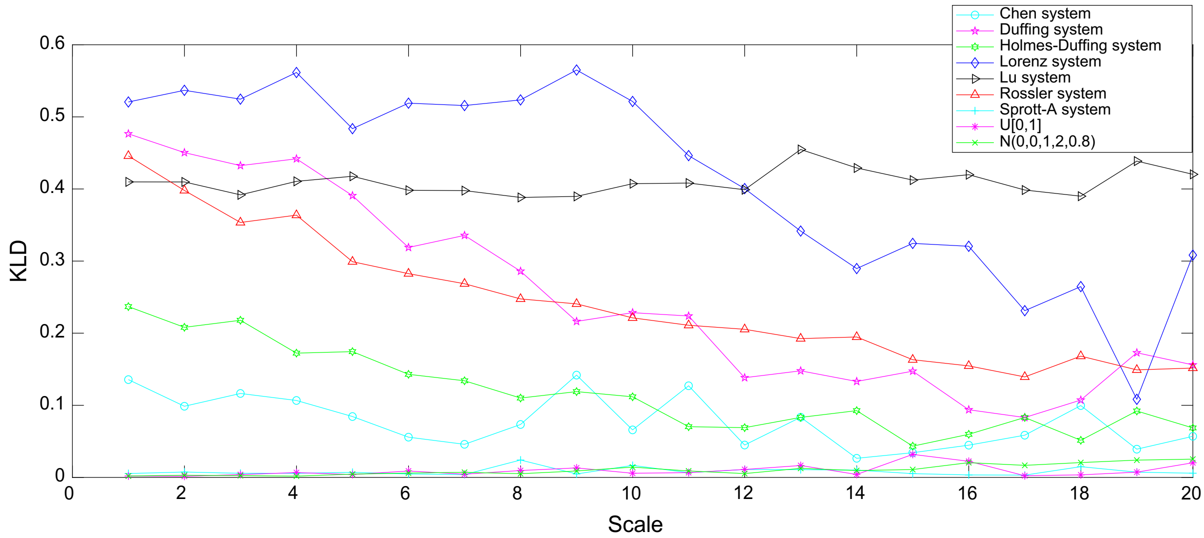


Fig. 4 Irreversibility measures KLD_τ of nine simulated series as a function of τ with $\tau \in [1, 20]$.

Table 2 World stock indices

Region	Country	Symbol	Company	Number	Score rank
Americas	1.Brazil	BVSP	Bovespa Index	3704	11
	2. Canada	GSPTSE	S&P Composite Index	3765	7
	3. Mexico	MXX	IPC Mexico Index	3759	8
	4.USA	DJ1	Dow 30 Index	3775	13
	5.USA	GSPC	S&P 500 Index	3775	20
	6.USA	IXIC	Nasdaq Index	3775	4
	7.USA	RUT	Russell 2000 Index	3775	12
Europe	8.Austria	ATX	ATX Index	3709	2
	9. Spain	IBEX	IBEX 35 Index	3828	17
	10. France	Kill	CAC 40 Index	3833	I
	11.Germany	GDAX1	DAX Index	3803	21
	12. England	LSI:	LSE Index	3783	15
	13.Switzerland	SSMI	SMI Index	3784	5
Asia & Pacific	14.Australia	AORD	ASX All Ordinaries Index	3788	18
	15.India	BSESN	BSE Sensex Index	3676	10
	16.Indonesia	JKSE	Jakarta Composite Index	3650	9
	17.Korea	K.S11	KOSPI Composite Index	3693	19
	18.Japan	N225	Nikkei 225 Index	3671	6
	19.China Mainland	SSE	Shanghai Composite Index	3644	16
	20.China Hong Kong	HS1	Hang Seng Index	3690	14
	21.China Taiwan	TSM	TSMCL Index	3775	3

So as to precisely measure the time irreversibility, we provide Score [s], which is the average of the annual irreversibility value to assess the time irreversibility of a given stock index s[53].

$$\mathrm{Score}[s] = \frac{1}{15} \sum_{\mathrm{year}=2005}^{2019} \mathrm{KLD}_{\mathrm{year}}^{s} \tag{4.2}$$

Furthermore, we calculate several other statistics of the irreversibility. The standard deviation *sd* is established as

$$\mathrm{sd} = \left(\frac{1}{14} \sum_{\mathrm{year}=2005}^{2019} \left(\mathrm{KLD}_{\mathrm{year}} - \overline{\mathrm{KLD}}\right)^2 \right)^{\frac{1}{2}} \tag{4.3}$$

where

$$\overline{\mathrm{KLD}} = \frac{1}{15} \sum_{\mathrm{year}=2005}^{2019} \mathrm{KLD}_{\mathrm{year}} \tag{4.4}$$

And the coefficient of variation C_v is given by

$$C_v = \frac{\mathrm{sd}}{\overline{\mathrm{KLD}}} \tag{4.5}$$

the third central moment v_3 is defined as

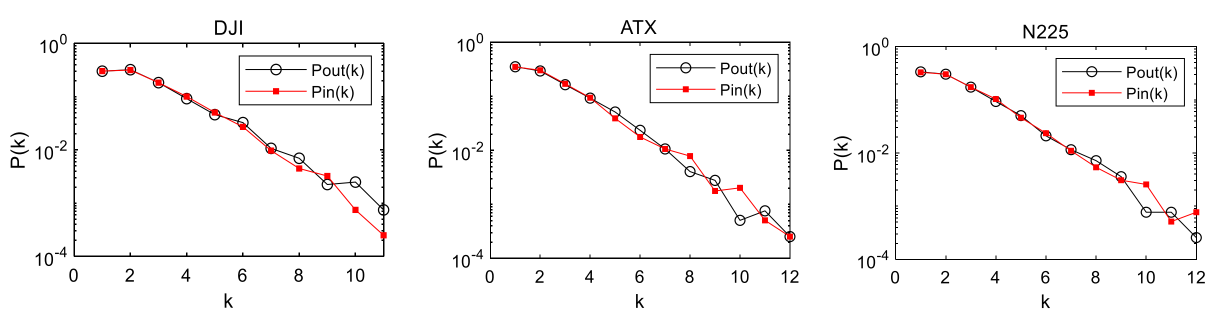
$$v_3 = \frac{1}{15} \sum_{\mathrm{year}=2005}^{2019} \left(\mathrm{KLD}_{\mathrm{year}} - \overline{\mathrm{KLD}}\right)^3 \tag{4.6}$$

the skewness β_s is expressed as

$$\beta_s = \frac{v_3}{\mathrm{sd}^3} \tag{4.7}$$

In the synthetic data analysis, we find that *KLD* can distinguish the irreversible multivariate time series from the reversible ones. However, it can be realized that in the process of financial time series analysis, it cannot well reflect the irreversibility of these multivariate time series as shown in Fig. 5. Here, we only give three different stock markets, respectively, from Americas, Europe and Asia & Pacific. The difference between the distribution of their *ingoing* and *outgoing* degree is not very obvious. Therefore, we consider using KLD_τ to analyze the irreversibility of stock indices on different scales. Figure 6 illustrates the volatility of KLD_τ of each stock index on scale $\tau \in [1, 10]$. As the scale increases, the coarse-grained time series becomes shorter, and the KLD_τ, respectively, shows the increasing trend. However, from the results of synthetic data, we know that the KLD_τ of irreversible time series decreases with the increase in scale and tends to 0, while the KLD_τ of reversible time series is very close to 0 on any scale. By comparing the results of KLD_τ of the simulated series given in Fig. 4, we can obtain that the value of KLD_τ of each stock index series is relatively small, but it is not strictly close to 0, which is different from that of reversible time series. Therefore, we believe that financial time series are multiscale irreversibility.

As we can see from Fig. 6, the stock indices from Americas fluctuate sharply at $\tau = 6\&8$, while the stock indices from Europe fluctuate greatly at $\tau = 7$. However, the stock indices from Asia & Pacific are scattered and it does not fluctuate particularly violently on any scale. These facts are a little different from the situation in Fig. 7, which exhibits the volatility of KLD_τ averaged for all stock indices from each region on scale $\tau \in [1, 10]$. The curve of world reaches its maximum at $\tau = 7$ and has a significant drop at $\tau = 8$. Except for the curve from



(a) **(b)** **(c)**

Fig. 5 The *ingoing* and *outgoing* degree distributions of the DVV_g corresponding to the multivariate time series of data points from each region: Americas, Europe and Asia & Pacific

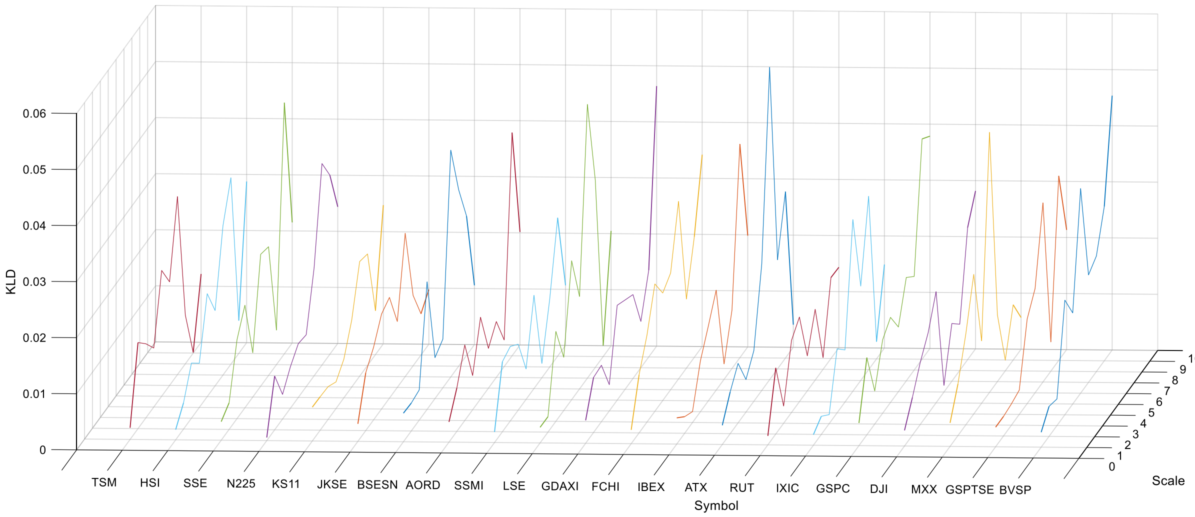


Fig. 6 Irreversibility measures KLD_τ of 21 stock indices as a function of τ with $\tau \in [1, 10]$.

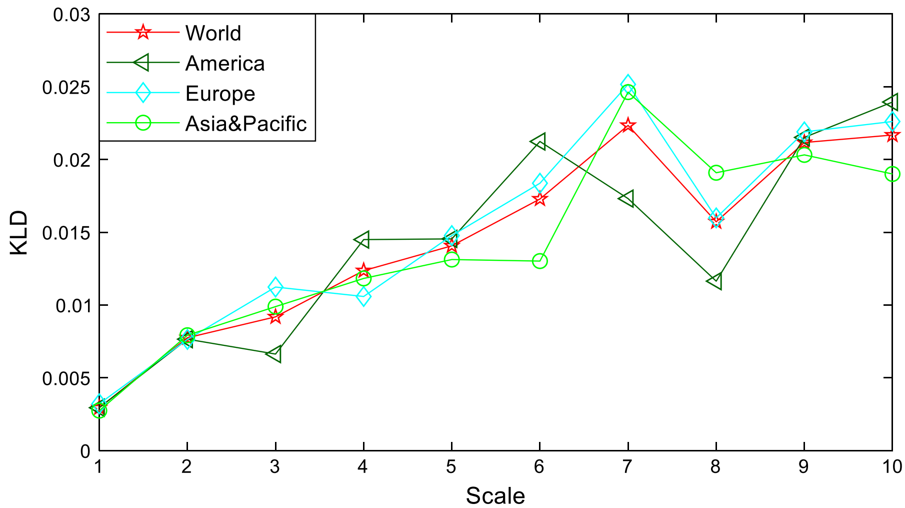


Fig. 7 Irreversibility measures KLD_τ C_{MMJS} C_{MMJS} averaged over all the stock indices from each region: Americas, Europe and Asia & Pacific

Americas, the curves from Europe and Asia & Pacific reach the maximum at $\tau = 7$ and decrease at $\tau = 8$. Curve of Americas shows a downward trend at $\tau = 7\&8$. Because the statistical characteristics of time series irreversibility can reflect the dynamic characteristics of the system, these facts make us consider periodicity. When we handle the time series with scale factor $\tau = 7$, we regard the information of 7 days as a whole, and the information of the next trading day may be identical with that of the first day. Similarly, the scale factor $\tau = 8$ can also be the periodic signal. Owing to the effect of uncertain factors such as the level of development and system mechanism, the periodicity of different stock markets tends to be distinguishing.

These points obtained from the score and its average annual volatility of the irreversibility corresponding to each stock index are shown in Fig. 8. And the standard deviation of the price log-returns of one year is regarded as the proxy for volatility of stock indices. As is known to all, the degree of stability of financial stock indices is generally reflected by its annual volatility. If the two statistics are related, the points in the scatter plot will tend to a smooth curve rather than being as scattered as the plot showing here. Therefore, we can come to a conclusion from Fig. 8

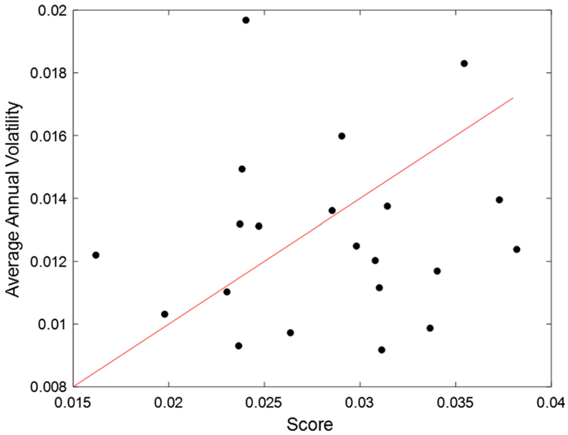


Fig. 8 The plot of points generated by the score and its average annual volatility of the irreversibility of each stock index

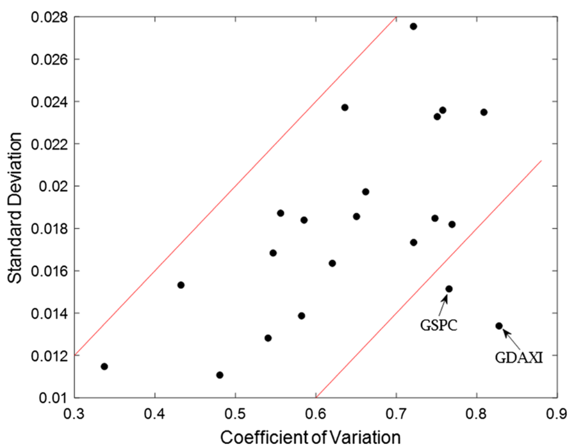


Fig. 9 The plot of points generated by the coefficient of variation C_v and its standard deviation sd of the irreversibility of each stock index

that there is no correlation between the irreversibility and volatility, which implies that we take the time irreversibility as a new statistical measure to research the properties of financial time series is reasonable. In order to further explore the irreversibility of financial time series, some other statistics, such as the standard deviation sd, the coefficient of variation C_v, the third central moment v_3 and the skewness β_s are computed.

Figure 9 presents the relation between the coefficient of variation C_v and its standard deviation sd of the irreversibility of each stock index. The coefficient of variation is a statistic to measure the variation degree of variables, which is not only impacted by the

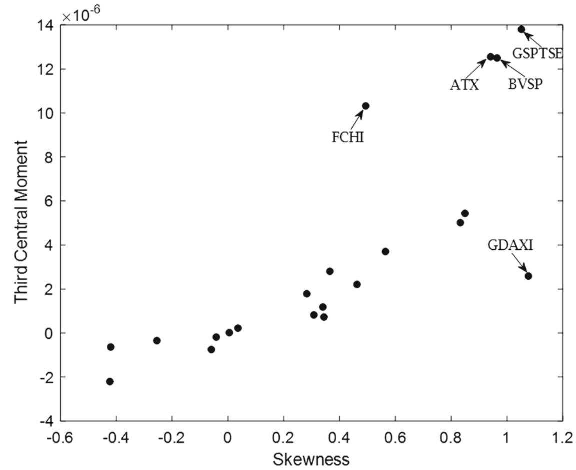


Fig. 10 The plot of points generated by the skewness β_s. and its third central moment v_3. of the irreversibility of each stock index

level of dispersion of variables, but also under the influence of the average of variables. As Fig. 9 shows, the standard deviation of irreversibility approximatively tends to be proportional to the coefficient of variation. It indicates that the average of variables, which is displayed by the average of the annual irreversibility, may have a negligible effect on its coefficient of variation. Consequently, we can replace the standard deviation with the coefficient of variation to explore the potential qualities of financial time series.

Figure 10 manifests the points plotted by the skewness β_s and its third central moment v_3 of the irreversibility of 21 stock indices, respectively. Skewness is the statistic, which reflects the asymmetry of the probability distribution of a random variable with respect to its mean value. By viewing Fig. 10, it can be realized that the skewness values of most points are between -0.6 and 1. Besides, the third central moment corresponding to them is less than others and all close to 0. It suggests that the skewness can express the irreversibility alone. Ranking each stock index according to its score, we can discover that the distributions of the points corresponding to the top two stock indices and the bottom one are all outstanding.

To better analyze the financial time series, we investigate the irreversibility of all stock indices in different years and plot the points generated by the skewness β_s and its third central moment v_3 of the irreversibility of each year in Fig. 11. Due to the essential difference between financial crisis period and

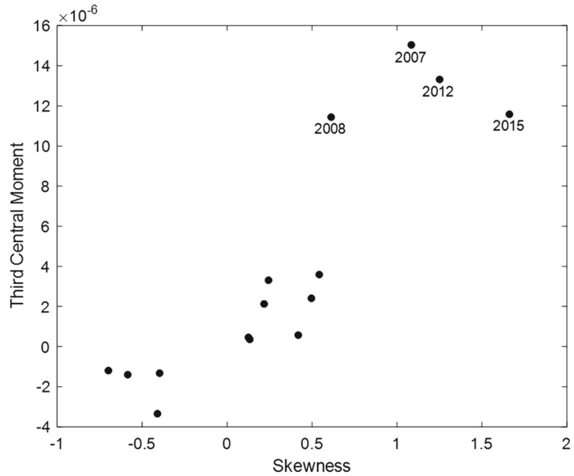


Fig. 11 The plot of points generated by the skewness β_s and its third central moment v_3 of the irreversibility of different years

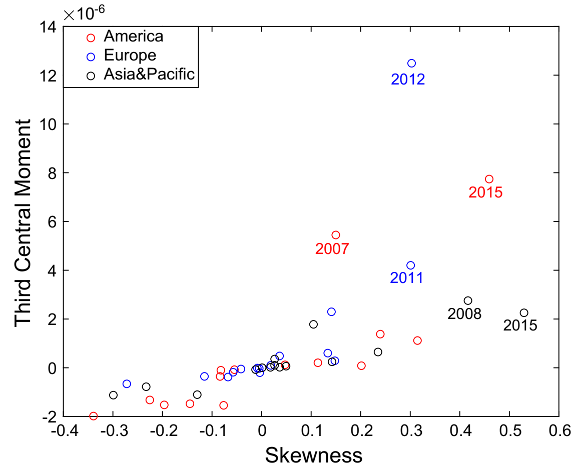


Fig. 12 The plots of points generated by the skewness β_s and its third central moment v_3 of the irreversibility of different years from each region: Americas, Europe and Asia & Pacific

economic stable period, we try to make use of the new proposed method to distinguish them. As we all know, one of the most serious financial crises in the history was triggered by the American subprime mortgage crisis in 2007 and the financial crisis broke out in 2008. The global economy began to recover gradually in 2011. As shown in Fig. 11, it is obvious that the two points corresponding to 2007 and 2008 are deviated from most points, but the points of other two years are also distinguished from most of them. In order to make a more detailed analysis, the location of points given by the skewness β_s and its third central moment v_3 of each year from three different regions is displayed in Fig. 12.

From Fig. 12, we realize that 2015 is a special year for Americas and Asia & Pacific, while for Europe, 2012 presents different characteristics from other years. As a matter of fact, the stock indices from Americas, such as *BVSP*,*GSPTSE* and *RUT*, did show greater volatility in 2015 than in other years, while *DJI* and *GSPC* from Americas did not show significant fluctuations in the whole year, but both had pretty large fluctuations in several months. And the stock indices from Asia & Pacific, such as *N*225, *BSESN*,*SSE* and *HSI*, all fluctuated sharply from 2014 to 2015. For the whole stock indices from Europe, they were basically in a sustained upward phase in 2012.

5 Conclusions

In this paper, we introduce a new method to measure the irreversibility of multivariate time series. By mapping multivariate time series into the directed vector visibility graph (DVV_g), we measure the level of irreversibility of the time series by using the value of Kullback–Leibler divergence (KLD), which is calculated by the *ingoing* and *outgoing* degree distributions corresponding to the DVV_g.

In order to validate the effectiveness of this method, we select uncorrelated stochastic series, correlated stochastic series, dissipative chaotic systems and conservative chaotic systems, and evaluate the degree of irreversibility of their multivariate time series. The results show that when the time series is reversible, their *ingoing* and *outgoing* degree distributions are almost indistinguishable and their specific numerical values of KLD are all very close to 0. The deviation between the KLD of reversible time series and zero is caused by the finite size effect, while the irreversible time series converges to an asymptotical nonzero positive value with the increase in series size.

Later, to further investigate the internal structure of the system, we utilize the multiscale method to analyze the irreversibility of time series. For the reversible time series, their KLD_τ are close to 0 at all scales, while the KLD_τ of irreversible time series decreases and tends to 0 with the increase in scale.

Therefore, we consider that the coarse-grained process changes the internal structure of the system, which may have an influence on the degree of its internal irreversibility. However, for the reversible time series, even if the coarse-grained process changes the internal structure of the system, the coarse-grained time series is still in a random state and its irreversibility is low.

Finally, this method is applied to stock markets. We choose 21 stock indices and, respectively, divide them into three groups on account of different regions. Through experimental analysis, we believe that the financial time series are multiscale irreversibility. To better evaluate the irreversibility of financial time series, we introduce some other statistics and recognize that the new method can be used to identify the special financial periods.

On the whole, the new irreversible method is an effective way to measure the irreversibility of multivariate time series.

Acknowledgements The financial supports from the funds of the National Natural Science Foundation of China (61771035) and the Fundamental Research Funds for the Central Universities (2018JBZ104) are gratefully acknowledged.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Weiss, G.: Time-reversibility of linear stochastic processes. J. Appl. Probab. **12**(4), 831–836 (1975)
2. Kawai, R., Parrondo, J.M.R., Van den Broeck, C.: Dissipation: the phase-space perspective. Phys. Rev. Lett. **98**(8), 080602 (2007)
3. Parrondo, J.M.R., Van den Broeck, C., Kawai, R.: Entropy production and the arrow of time. New J. Phys. **11**(7), 073008 (2009)
4. Lawrance, A.J.: Directionality and reversibility in time series. Int. Stat. Rev. **59**(1), 67–79 (1991)
5. Diks, C., Houwelingen, J.C., Takens, F., DeGoede, J.: Reversibility as a criterion for discriminating time series. Phys. Lett. A **201**(23), 221–228 (1995)
6. Theiler, J., Prichard, D.: Constrained-realization Monte-Carlo method for hypothesis testing. Physica D **94**(4), 221–235 (1996)
7. Rothman, P.: The comparative power of the TR test against simple threshold models. J. Appl. Economet. **7**(7), 187–195 (1992)
8. Ramsey, J.B., Rothman, P.: Time irreversibility and business cycle asymmetry. J. Money Cred. Bank. **28**(1), 1–21 (1996)
9. Hinich, M.J., Rothman, P.: Frequency-domain test of time reversibility. Macroeconom. Dynam. **2**(1), 72–88 (2011)
10. Chen, Y.T., Chou, R.Y., Kuan, C.M.: Testing time reversibility without moment restrictions. J. Econometr. **95**(1), 199–218 (2000)
11. Cheng, Q.S.: On time-reversibility of linear processes. Biometrika **86**(2), 483–486 (1999)
12. Costa, M.D., Goldberger, A.L., Peng, C.K.: Broken asymmetry of the human heartbeat loss of time irreversibility in aging and disease. Phys. Rev. Lett. **95**(19), 198102 (2005)
13. Costa, M.D., Peng, C.K., Goldberger, A.L.: Multiscale analysis of heart rate dynamics: entropy and time irreversibility measures. Cardiovasc. Eng. **8**(2), 88–93 (2008)
14. Costa, M.D., Goldberger, A.L., Peng, C.K.: Multiscale entropy analysis of complex physiologic time series. Phys. Rev. Lett. **89**(6), 068102 (2002)
15. Costa, M.D., Goldberger, A.L., Peng, C.K.: Multiscale entropy analysis of biological signals. Phys. Rev. E **71**(2), 021906 (2005)
16. Brennan, M., Palaniswami, M., Kamen, P.: Do existing measures of Poincare plot geometry reflect non-linear features of heart rate variability. IEEE Trans. Biomed. Eng. **48**(11), 1342–1347 (2001)
17. Guzik, P., Piskorski, J., Krauze, T., Wykretowicz, A., Wysocki, H.: Heart rate asymmetry by Poincare plots of RR intervals. Biomed. Tech. **51**(4), 272–275 (2006)
18. Porta, A., Guzzetti, S., Montano, N., Gnecchi-Ruscone, T., Furlan, R., Malliani, A.: Time reversibility in short-term heart period variability. Comput. Cardiol. **33**, 77–80 (2006)
19. Porta, A., Casali, K.R., Casali, A.G., Gnecchi-Ruscone, T., Tobaldini, E., Montano, N., Lange, S., Geue, D., Cysarz, D., Van Leeuwen, P.: Temporal asymmetries of short-term heart period variability are linked to autonomic regulation. AJP Regulat. Integr. Comparat. Physiol. **295**(2), 550–557 (2008)
20. Porta, A., Daddio, G., Bassani, T., Maestri, R., Pinna, G.D.: Assessment of cardiovascular regulation through irreversibility analysis of heart period variability: a 24 hours Holter study in healthy and chronic heart failure populations. Philos. Trans. R. Soc. A **367**(1892), 1359–1375 (2009)
21. De La Cruz Torres, B., Naranjo Orellana, J.: Multiscale time irreversibility of heartbeat at rest and during aerobic exercise. Cardiovasc. Eng. **10**(1), 1–4 (2010)
22. Hou, F.Z., Zhuang, J.J., Bian, C.H., Tong, T.J., Chen, Y., Yin, J., Qiu, X.J., Ning, X.B.: Analysis of heartbeat asymmetry based on multi-scale time irreversibility test. Phys. A **389**(4), 754–760 (2010)
23. Roldán, E., Parrondo, J.M.: Estimating dissipation from single stationary trajectories. Phys. Rev. Lett. **105**(15), 150607 (2010)
24. Roldan, E., Parrondo, J.M.R.: Entropy production and Kullback-Leibler divergence between stationary trajectories of discrete systems. Phys. Rev. E **85**(3), 031129 (2012)
25. Cover, T.M., Thomas, J.A.: Elements of Information Theory. Wiley (2006)
26. Heyden, M.J., Diks, C.G.C., Hoekstra, B.P.T., DeGoede, J.: Testing the order of discrete markov chains using surrogate data. Physica D **117**(1), 299–313 (1998)

27. Daw, C.S., Finney, C.E.A., Kennel, M.B.: Symbolic approach for measuring temporal "Irreversibility." Phys. Rev. E **62**(2), 1912–1921 (2000)
28. Lehrman, M., Rechester, A.B., White, R.B.: Symbolic analysis of chaotic signals and turbulent fluctuations. Phys. Rev. Lett. **78**(1), 54–57 (1997)
29. Finney, C.E.A., Green, J.B., Daw, C.S.: Symbolic time-series analysis of engine combustion measurements. Soc. Automot. Eng. Trans. **107**(3), 888–897 (1998)
30. Lacasa, L., Nunez, A., Roldan, E., Parrondo, J.M.R., Luque, B.: Time series irreversibility: a visibility graph approach. Eur. Phys. J. B **85**(6), 217–227 (2012)
31. Yang, Y., Wang, J., Yang, H., Mang, J.: Visibility graph approach to exchange rate series. Phys. A **388**(20), 4431–4437 (2009)
32. Long, Y.: Visibility graph network analysis of gold price time series. Phys. A **392**(16), 3374–3384 (2013)
33. Liu, C., Zhou, W.X., Yuan, W.K.: Statistical properties of visibility graph of energy dissipation rates in three-dimensional fully developed turbulence. Phys. A **389**(13), 2675–2681 (2010)
34. Murugesan, M., Sujith, R.I.: Combustion noise is scale-free: transition from scale-free to order at the onset of thermoacoustic instability. J. Fluid Mech. **772**, 225–245 (2015)
35. Ahmadlou, M., Adeli, H., Adeli, A.: New diagnostic EEG markers of the Alzheimer's disease using visibility graph. J. Neural Transm. **117**(9), 1099–1109 (2010)
36. Ren, W., Jin, N.: Vector visibility graph from multivariate time series: a new method for characterizing nonlinear dynamic behavior in two-phase flow. Nonlinear Dyn. **97**(4), 2547–2556 (2019)
37. Lacasa, L., Luque, B., Ballesteros, F., Luque, J., Nuno, J.C.: From time series to complex networks: the visibility graph. Proc. Natl. Acad. Sci. U.S.A. **105**(13), 4972–4975 (2008)
38. Luque, B., Lacasa, L., Luque, J., Ballesteros, F.: Horizontal visibility graphs: exact results for random time series. Phys. Rev. E **80**(4), 046103 (2009)
39. Yang, A.C., Hseu, S., Yien, H., Goldberger, A.L., Peng, C.K.: Linguistic analysis of the human heartbeat using frequency and rank order statistics. Phys. Rev. Lett. **90**(10), 108103 (2003)
40. Kennel, M.B.: Testing time symmetry in time series using data compression dictionaries. Phys. Rev. E **69**(5), 056208 (2004)
41. Cammarota, C., Rogora, E.: Time reversal, symbolic series and irreversibility of human heartbeat. Chaos, Solitons Fractals **32**(5), 1649–1654 (2007)
42. Andrieux, D., Gaspard, P., Ciliberto, S., Garnier, N., Joubaud, S., Petrosyan, A.: Entropy production and time asymmetry in nonequilibrium fluctuations. Phys. Rev. Lett. **98**(15), 150601 (2007)
43. Wang, Q., Kulkarni, S.R., Verdu, S.: Divergence estimation of continuous distributions based on data-dependent partitions. IEEE Trans. Inf. Theory **51**(9), 3064–3074 (2005)
44. Gaspard, P.: Time-reversed dynamical entropy and irreversibility in Markovian random processes. J. Stat. Phys. **117**(3–4), 599–615 (2004)
45. Gaspard, P.: Hamiltonian dynamics, nanosystems, and nonequilibrium statistical mechanics. Phys. A **369**(1), 201–246 (2006)
46. Porporato, A., Rigby, J.R., Daly, E.: Irreversibility and fluctuation theorem in stationary time series. Phys. Rev. Lett. **98**(9), 094101 (2007)
47. Chen, G., Ueta, T.: Yet another chaotic attractor. Int J Bifurcat Chaos **9**(7), 1465–1466 (1999)
48. Duffing, G.: Erzwungene schwingungen bei veränderlicher eigenfrequenz und ihre technische Bedeutung. Vieweg Braunschweig (1918)
49. Lorenz, E.N.: Deterministic nonperiodic flow. J. Atmos. Sci. **20**(2), 130–141 (1963)
50. Lu, J., Chen, G.: A new chaotic attractor coined. Int. J. Bifurcat. Chaos **12**(3), 659–661 (2002)
51. Rössler, O.E.: An equation for continuous chaos. Phys. Lett. A **57**(5), 397–398 (1976)
52. Sprott, J.C.: Some simple chaotic flows. Phys. Rev. E **50**(2), 647–650 (1994)
53. Wang, Y., Shang, P.: A new measurement of financial time irreversibility based on information measures method. Phys. A **503**, 221–230 (2018)