PHILOSOPHICAL TRANSACTIONS A

rsta.royalsocietypublishing.org

Research

Cite this article: Yeh J-R, Peng C-K, Huang NE. 2016 Scale-dependent intrinsic entropies of complex time series. *Phil. Trans. R. Soc. A* **374**: 20150204.
http://dx.doi.org/10.1098/rsta.2015.0204

Accepted: 4 January 2016

One contribution of 13 to a theme issue 'Adaptive data analysis: theory and applications'.

Subject Areas:
biomedical engineering, analysis, complexity, fractals, biophysics

Keywords:
multi-scale entropy, complexity, empirical mode decomposition, detrending, fractal property

Author for correspondence:
Chung-Kang Peng
e-mail: chungkang.peng@gmail.com

Scale-dependent intrinsic entropies of complex time series

Jia-Rong Yeh[1], Chung-Kang Peng[1,2] and Norden E. Huang[1]

[1]Research Center for Adaptive Data Analysis and Center for Dynamical Biomarkers and Translational Medicine, National Central University, Taoyuan, Taiwan, Republic of China
[2]Division of Interdisciplinary Medicine and Biotechnology, Beth Israel Deaconess Medical Center/Harvard Medical School, Boston, MA, USA

Multi-scale entropy (MSE) was developed as a measure of complexity for complex time series, and it has been applied widely in recent years. The MSE algorithm is based on the assumption that biological systems possess the ability to adapt and function in an ever-changing environment, and these systems need to operate across multiple temporal and spatial scales, such that their complexity is also multi-scale and hierarchical. Here, we present a systematic approach to apply the empirical mode decomposition algorithm, which can detrend time series on various time scales, prior to analysing a signal's complexity by measuring the irregularity of its dynamics on multiple time scales. Simulated time series of fractal Gaussian noise and human heartbeat time series were used to study the performance of this new approach. We show that our method can successfully quantify the fractal properties of the simulated time series and can accurately distinguish modulations in human heartbeat time series in health and disease.

1. Introduction

Multi-scale entropy (MSE) is an innovative assessment method that is derived from statistical physics and complex systems for diagnosing the health condition of biological systems [1–3]. The dynamical complexity, measured by MSE, of a biological system is proposed to be an assessment of a system's capacity to adapt to various unpredictable stresses. Therefore, a biological system with a high degree of complexity is considered

THE ROYAL SOCIETY PUBLISHING

to be a healthy one. Loss of complexity is an important characteristic for biological systems in ageing and poor functional conditions (i.e. diseases) [1–3]. In recent years, MSE has been widely applied in diverse research topics, ranging from electromagnetics [4] and environmental research [5] to statistical mechanics [6]. An important scientific and technical issue when applying MSE to complex time series was noted, but no systematic solution has been developed. The challenge is the following: because entropies calculated by MSE depend on the threshold determined by the standard deviation of time series, trends embedded in time series affect the outcomes of MSE [7,8]. However, a complex time series is often nonlinear and multi-modal and it is difficult to determine the appropriate time scales for detrending a nonlinear time series based on linear assumptions. Two important issues should be taken into consideration: (i) What type of method is appropriate for detrending complex time series? (ii) How many trends embedded in a complex time series should be removed.

In 2007, Wu *et al.* [9] discussed the detrending and variability of nonlinear and multi-modal time series based on the application of an adaptive algorithm of empirical mode decomposition (EMD). They proposed that a 'trend' in the data depends on the time scale of the observation. EMD can decompose a complex time series into a set of intrinsic mode functions (IMFs). Each IMF represents a nonlinear component with oscillations in a narrow range of time scales. A local mean of upper and lower envelopes enclosing the data can be viewed as an intrinsic trend for the shortest time scales. This process can be iterated by treating the trend as data, and to decompose the next level of trend (on a larger time scale). Here, the word 'intrinsic' means that the trend is determined completely by the data themselves. No *a priori* functional form and time scales were used to determine the trend with the EMD-based approach. Thus, we can obtain a set of intrinsic trends from small time scale to large time scale. Effectively, the nth detrended time series is a summation of the first n IMFs. When the trends with different time scales are removed, we can quantify the complexity of the data after excluding the fluctuations beyond a specific time scale of trends. Then, a systematic approach to the complexity of the complex time series can be realized.

The difference in complexity between two successive detrended time series is assumed to represent the complexity contributed by an IMF with a specific time scale. Furthermore, the complexity contributed by an IMF can be simplified so that it is represented by an entropy on a specific time scale because an IMF is a narrow band time series. Then, a nonlinear and multi-modal time series can be expressed as the combination of a set of IMFs and the monotonic final residual. The complexity of the time series can be represented as a set of scale-dependent entropies contributed by IMFs with different corresponding time scales. In this study, the entropy contributed by an IMF is denoted as the *intrinsic entropy* and the time scale of the IMF is denoted as the *intrinsic time scale*.

To examine the characteristics of complexity expressed using scale-dependent intrinsic entropies, this new approach was used to analyse two completely different types of datasets: (i) an artificial time series of simulated fractal Gaussian noise with a known Hurst exponent and (ii) human heartbeat time series of healthy subjects and patients with various diseases.

2. Methodology

(a) Quantify the complexity of a stationary time series by multi-scale entropy

The MSE analysis was developed in recent years [1,2]. Briefly, the complexity of a time series is expressed by the entropy as a function of coarse-graining scales. There are two steps in MSE: (i) coarse-graining the time series using different scales; (ii) quantifying the degree of irregularity for each coarse-grained time series by the SampEn algorithm [10]. The details of SampEn can be found elsewhere [11]. The coarse-graining procedure of a time series is defined as

$$x^{\tau}(l) = \frac{1}{\tau} \sum_{t=(l-1)\tau+1}^{l\tau} x(t), \quad 1 \le l \le \frac{N}{\tau}, \tag{2.1}$$

where $x^{\tau}(l)$ is the coarse-grained time series with coarse-graining scale τ; l is the new index of the coarse-grained time series; $x(t)$ is the time series before coarse-graining; and N is the length of the original time series. The length of each coarse-grained time series shrinks to the length of the original time series divided by the coarse-graining scale τ.

The complexity of a time series $x(t)$ can be expressed as a set of entropies with different coarse-graining scales $\tau = 1, \ldots, L$. The entropy for a coarse-grained time series calculated by SampEn is noted as $SampEn(x^{\tau}(l), m, r, SD)$, where m is the vector length; r is the tolerance coefficient; and SD is the standard deviation. In the calculation of SampEn, we set vector length $m = 2$ and tolerance coefficient $r = 0.2$.

It should be mentioned that $x(t)$ is assumed to be a stationary time series in the calculation of entropy. However, a real-world signal is often non-stationary. An appropriate detrending process, therefore, is considered to be an important step to improve the stationary property of a real-world signal in the complexity analysis by MSE.

(b) Detrending by empirical mode decomposition

EMD can decompose a nonlinear and multi-modal time series into a finite number of IMFs and a monotonic residual [12],

$$X(t) = \sum_{n=1}^{M} C_n(t) + R(t), \tag{2.2}$$

where $X(t)$ is the original time series, $C_n(t)$ is the nth IMF, M is the number of IMFs and $R(t)$ is the final residual.

In 2007, EMD was proposed as an adaptive detrending algorithm for nonlinear and multi-modal time series [13]. It is important to note that the 'trend' of a time series has to be defined with respect to the 'scale' of interest. Thus, the EMD algorithm can be used to derive the scale-dependent trend in the following way: the overall residual term, R, can be considered as the global trend of the entire time series. The second largest scale trend is $R + C_M$. By adding an additional IMF, i.e. C_{M-1}, C_{M-2}, etc., we can obtain trends on shorter and shorter scales. The time scale of each trend can be defined as the averaged period of the corresponding IMF, because the averaged period represents the time scale of fluctuations with the trend removed. In summary, a detrended time series is the reconstructed time series of the first n IMFs,

$$x_n(t) = \sum_{i=1}^{n} C_i(t), \tag{2.3}$$

where $C_i(t)$ is the ith IMF. We denote $x_n(t)$ as the nth detrended time series.

The complexity of a detrended time series can be quantified by the MSE algorithm with a set of entropies for coarse-graining scales 1 to L. Each entropy is calculated by SampEn on the detrended coarse-grained time series.

(c) Define intrinsic entropy as representing the complexity contributed by a specific intrinsic mode function

Equation (2.3) shows that the nth detrended time series is a combination of the first n IMFs. Thus, an IMF can be expressed as the difference between two successive detrended time series, $C_n(t) = x_n(t) - x_{n-1}(t)$. The complexity contributed by $C_n(t)$ can be considered as a set of entropy increments for coarse-graining scales 1 to L when a specific IMF, $C_n(t)$, is added into the detrended time series. In this study, an entropy increment is defined as

$$\Delta E_n^{\tau} = \mathrm{SampEn}(x_n^{\tau}(l), m, r, SD_n) - \mathrm{SampEn}(x_{n-1}^{\tau}(l), m, r, SD_n). \tag{2.4}$$

It should be noted that the standard deviations SD_n used in both entropy calculations for $x_n^{\tau}(l)$ and $x_{n-1}^{\tau}(l)$ are the same. The entropy of $x_n^{\tau}(l)$ is always greater than that of $x_{n-1}^{\tau}(l)$ based on the same standard deviation SD_n. Therefore, ΔE_n^{τ} is always a positive value.

For coarse-graining scales $\tau = 1, \ldots, L$, a row vector of ΔE_n represents a set of entropy increments contributed by IMF n. Because a time series can be decomposed into M IMFs, an entropy increment matrix with size $M \times L$ is used to represent the overall entropy distribution for a time series. Because each IMF has its own intrinsic time scale, the entropy increments contributed by a specific IMF are significant only on a specific coarse-graining scale range correlated to the IMF's intrinsic time scale. The relationship between the intrinsic time scale and the specific coarse-graining scale range will be determined using the real and simulated data in §2d. The values of the elements within the specific scale range are significantly greater than 0. Therefore, to simplify the expression using the entropy increment matrix, we defined a new term of 'intrinsic entropy' as the maximum of a row vector of the entropy increment matrix, which reflects the most significant entropy increment for adding an IMF into a detrended time series. For a set of IMFs decomposed from a time series, a set of intrinsic entropies with different scales represents the complexity for the time series.

In summary, we apply the EMD to decompose a time series into a set of IMFs and the final residual in this new approach. Each IMF contributes an intrinsic entropy on its corresponding time scale. The ensemble of intrinsic entropies reflects a systemic complexity of a time series, and the individual intrinsic entropy represents the entropy contributed by a specific IMF on its corresponding scale.

(d) Define the time scale of an intrinsic mode function

As mentioned above, an intrinsic entropy is the maximum of a row in the entropy increment matrix, which represents the entropy contributed by a specific IMF, with a specific coarse-graining scale. It implies an IMF contributed by its corresponding intrinsic entropy on a specific time scale. It is interesting to determine the specific coarse-graining scale for determining the intrinsic entropy contributed by an IMF. Therefore, we investigate the time scale of an IMF. The time scale of an IMF can be defined as the averaged period of the oscillatory fluctuation according to the definition in a previous study, as expressed in the following equation [14]:

$$\bar{T}_n = \int S_{\ln T,n} d\ln T \left(\int S_{\ln T,n} \frac{d\ln T}{T} \right)^{-1}, \tag{2.5}$$

where N is the sample number of the time series; $S_{\ln T,n}$ is the Fourier spectrum of the nth IMF as a function of $\ln T$; T is the period; and $\bar{T}_n$ is the averaged period of the nth IMF.

In this study, we define the term 'intrinsic time scale' as the averaged period of an IMF to check the correlation between the intrinsic time scale and the specific coarse-graining scale. According to our analysis of the results from simulated data of fractal Gaussian noises and human heartbeat time series in the following sections, a strong correlation between the specific coarse-graining scales and intrinsic time scales can be found. Thus, we can easily determine the coarse-graining scales required to evaluate the intrinsic entropy contributed by a specific IMF.

3. Numerical experiments using simulated fractal Gaussian noise

Fractal Gaussian noise is a time series satisfying the condition of Gaussian distribution and presents a specific fractal property. The Hurst exponent is an index for the fractal property of a time series. For numerical simulation, there are many different methods proposed to generate simulated time series of fractal Gaussian noise, such as a fast fractional Gaussian noise generator [15], Fourier-based algorithms [16,17] and a wavelet-based method [18]. Here, a simple algorithm for generating the simulated fractal Gaussian noise proposed by Wood & Chan [19] was used to generate the fractal Gaussian noises with a Hurst exponent from 0 to 1. Furthermore, fractal Brownian motion can be generated by integration of fractal Gaussian noise, and the Hurst exponent of the fractal Brownian motion is the Hurst exponent of the fractal Gaussian noise plus 1 [20–22]. Therefore, we investigated a total of six different simulated time series of fractal Gaussian noise and fractal Brownian motion to represent a case with anti-correlation ($H = 0.25$), a

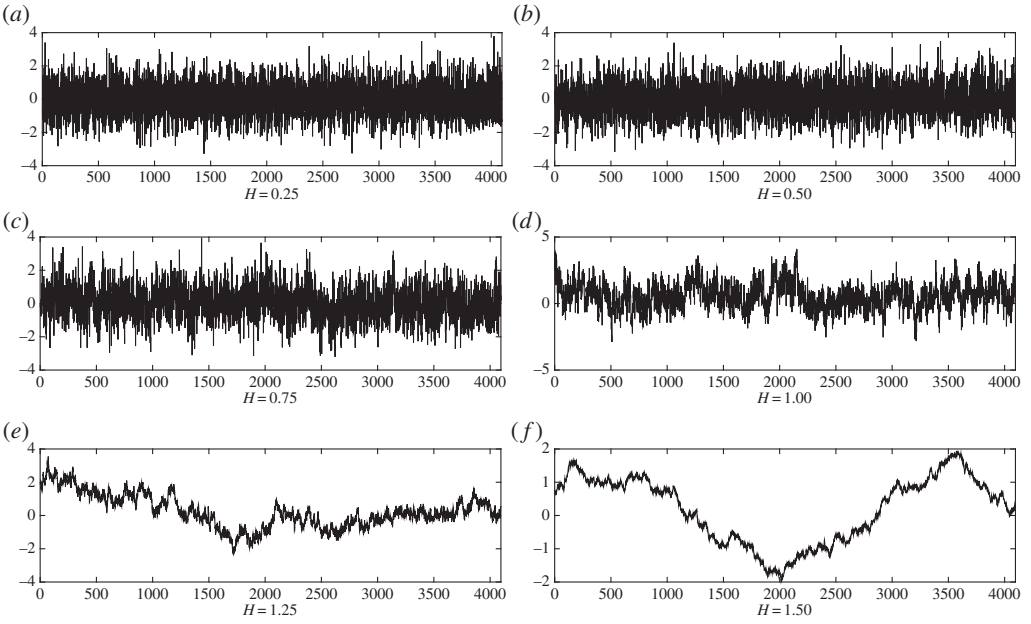

Figure 1. Examples of simulated time series of fractal Gaussian noise with different Hurst exponents ($H = 0.25, 0.50, 0.75, 1.00, 1.25$ and 1.50).

case using uncorrelated noise ($H = 0.5$) and four cases with persistent correlations ($H = 0.75$, 1.0, 1.25 and 1.50), as shown in figure 1.

To examine the complexities of the simulated time series, the first numerical experiment was conducted to estimate the complexities of the simulated time series by MSE. In this numerical simulation, the algorithm of *SampEn* was used to quantify the entropy of a time series. The parameters used in the entropy calculations are the vector length $m = 2$ and tolerance coefficient $r = 0.2$. The length of the simulated time series is 16 384 points. The complexities of the simulated time series with different fractal properties can be expressed by the MSE plots, as shown in figure 2.

In recent studies, a complexity index of a time series was defined as the summation of entropies on the first L coarse-graining scales in MSE. Here, we choose the number of coarse-graining time scales $L = 20$. In other words, we define a complexity index as the summation of entropies of the first 20 coarse-graining scales. The complexity indices of the simulated time series versus their corresponding Hurst exponents are shown in figure 3. In many disciplines of science and engineering, pink noise (also known as $1/f$ noise in the physics literature), with Hurst exponent $H = 1$, is considered to be the most complex in comparison with other fractal noise. This empirical assumption is consistent with our results (shown in figure 3) that pink noise has the highest complexity index.

As mentioned above, an entropy increment matrix with size $M \times L$ represents the entropies for coarse-graining scales $1, \ldots, L$ contributed by M IMFs decomposed from a time series. Figure 4 shows graphical presentations of the entropy increment matrix for the simulated time series with $H = 0.25$, 0.50, 0.75, 1.00, 1.25 and 1.50. The x-axis is the coarse-graining scales and the y-axis is the intrinsic time scale of the IMF, which contributed entropy in the corresponding row of the matrix. Both axes are shown in a logarithmic scale. The simulated time series with Hurst exponents $H = 0.25$ and 0.50 represent anti-correlation and non-correlation in the organization of the signals. In the graphic presentations shown in figure 4, for noises with Hurst exponents $H = 0.25$ and 0.50, only IMFs 1 and 2 contribute significant entropy increments on a low intrinsic time scale. For the simulated noises with persistent correlations, i.e. $H > 0.50$, the IMFs with large

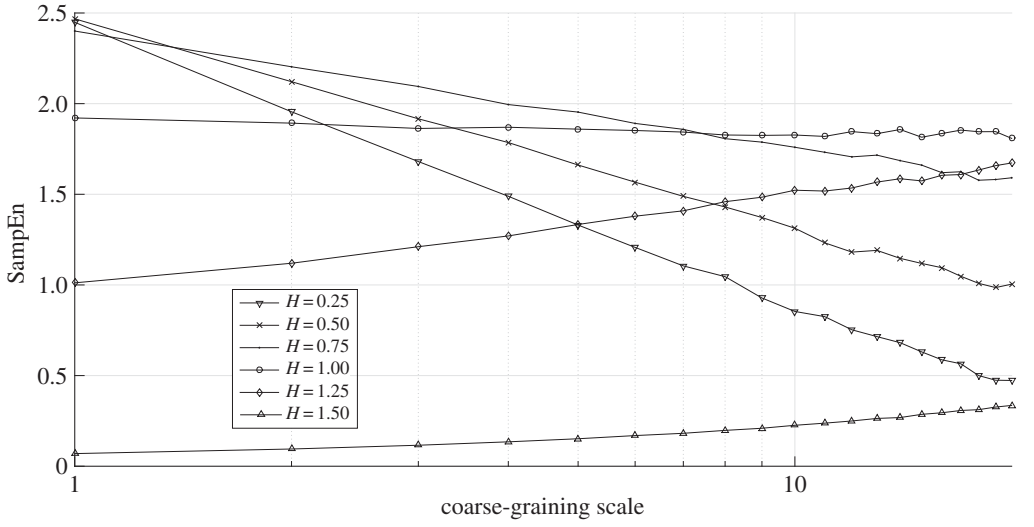

Figure 2. MSE results for the simulated time series of fractal Gaussian noise with Hurst exponents $H = 0.25, 0.50, 0.75, 1.00, 1.25$ and 1.50. Here, the sample entropy was evaluated using the parameters $m = 2$ and $r = 0.20$. The x-axis is shown in logarithmic scale.

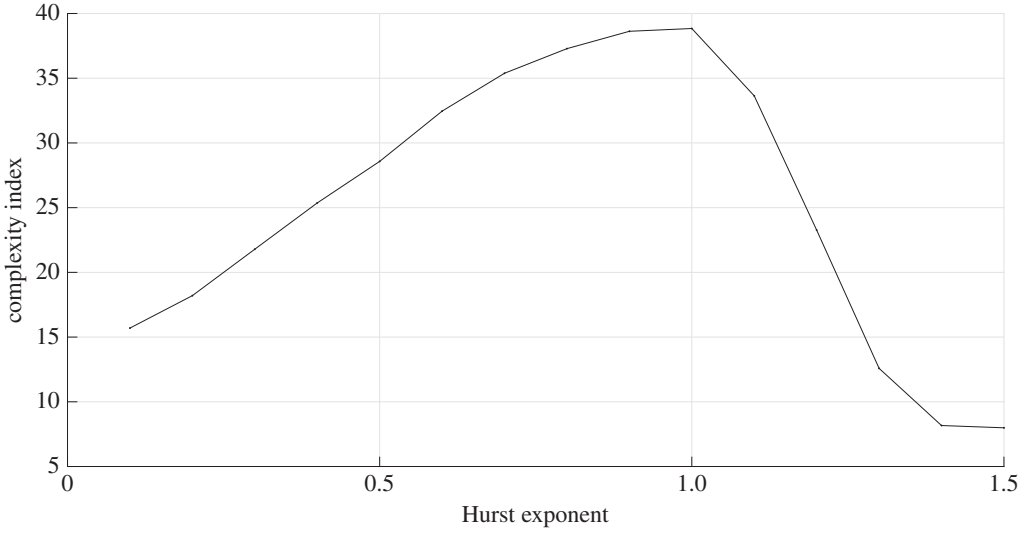

Figure 3. The MSE complexity index, defined as the summation of entropies over the first 20 coarse-graining scales, versus their corresponding Hurst exponents for simulated Gaussian noise.

intrinsic time scales contribute significant entropy increments on large coarse-graining scales. For each row of entropy increment matrix, only the elements around a specific coarse-graining scale have values significantly larger than 0. Moreover, the intrinsic time scale of an IMF has a power-law correlation with its corresponding coarse-graining scales, shown as diagonal distributions in the plots of figure 4 with logarithmic-scale axes. This finding implies that the intrinsic entropy, the maximum of a row of entropy increment matrix, contributed by an IMF can be found on a specific coarse-graining scale and the coarse-graining scale is a function of the intrinsic time scale.

To examine the correlation between intrinsic time scales and the specific coarse-graining scales, we collected samples of all the specific coarse-graining scales of the intrinsic entropies and

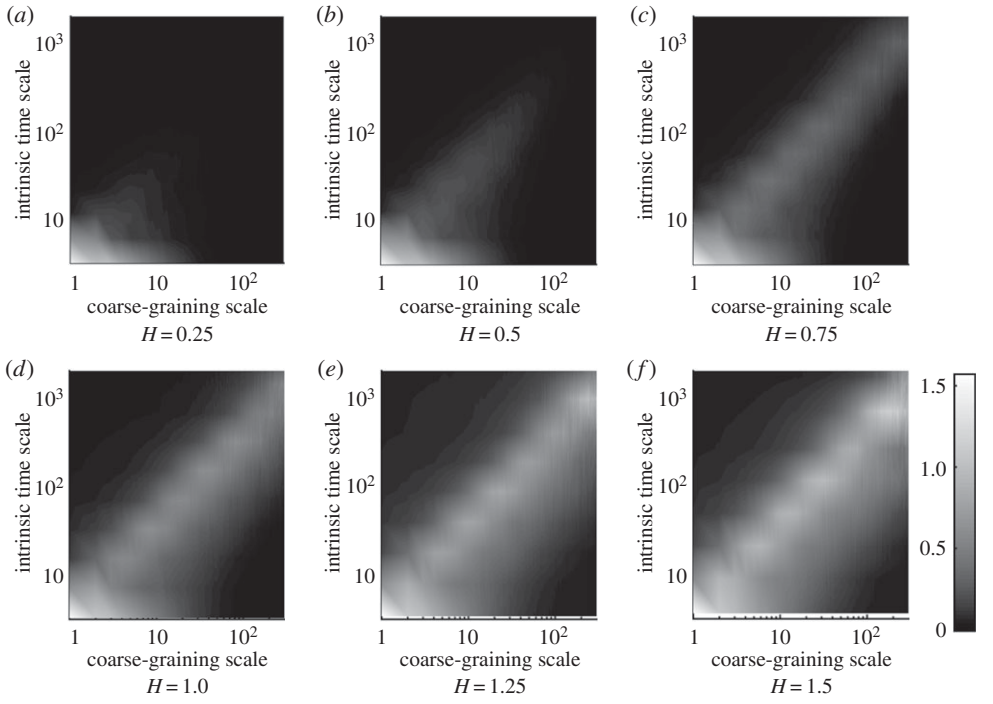

Figure 4. The entropy increment matrices of simulated time series of Gaussian noise with different Hurst exponents.

intrinsic time scales of their corresponding IMFs in the numerical experiment using simulated time series. Because the intrinsic entropies on large intrinsic time scales are small for noises with Hurst exponent $H \leqq 0.50$, the specific time scales cannot be accurately identified, because the values of the intrinsic entropies are not significantly different from the entropies for the other coarse-graining scales. Therefore, only the specific coarse-graining scales for intrinsic entropies greater than 0.05 were checked with the intrinsic time scales of the corresponding IMFs. The correlation between specific coarse-graining scales and intrinsic time scales is shown as figure 5. Then, the correlation can be expressed as

$$S_c = 0.294 \cdot S_i^{1.0}, \tag{3.1}$$

where S_i is the intrinsic time scale and S_c is the specific coarse-graining time scale for an IMF.

According to equation (3.1), only those entropy increments around the coarse-graining scales within the range from 1/4 to 1/3 of the intrinsic time scale should be calculated to derive the intrinsic entropy contributed by an IMF.

In summary, a simulated time series of fractal noise can be decomposed into a set of IMFs. Then, the intrinsic entropy contributed by a specific IMF can be evaluated as the entropy increment for a specific coarse-graining scale, which is located on the scale range at around 0.3 times the intrinsic time scale of the IMF. Moreover, the values of the intrinsic entropies for the IMFs decomposed from the simulated fractal noises depend on the Hurst exponent, as shown in figure 6.

Because a sufficient data length is required for a calculation of entropy, the data length of the coarse-grained time series must be longer than 300 samples. According to equation (3.1), the intrinsic time scale is 3.4 times the specific coarse-graining time scale. The intrinsic time scale represents the averaged period of the IMF and it should be 3.4 times the specific coarse-graining time scale in the calculation of intrinsic entropy. Thus, the sample number of the coarse-grained time series should be 3.4 times the number of periods in an IMF for the calculation of intrinsic

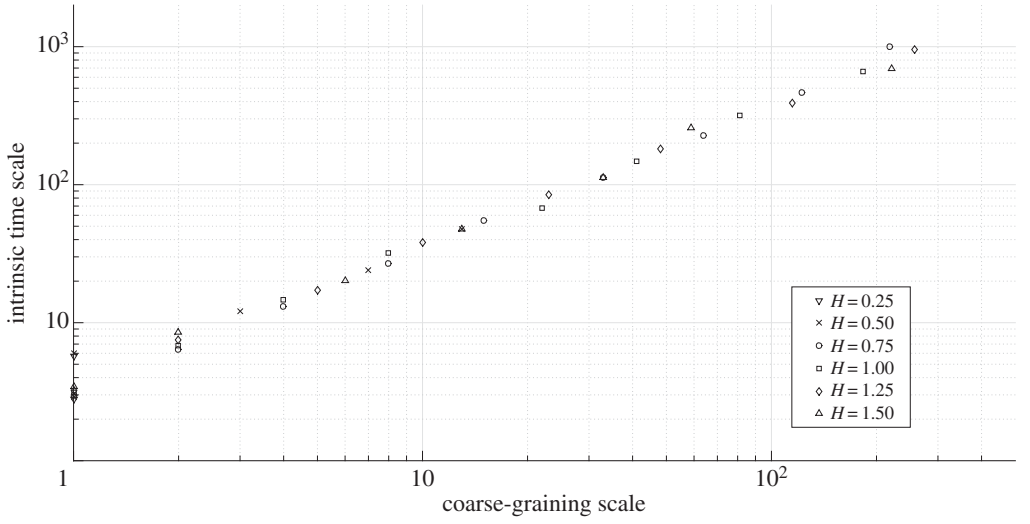

Figure 5. Relationship between the specific coarse-graining scales and the intrinsic time scales for IMFs decomposed from simulated time series of fractal Gaussian noise.

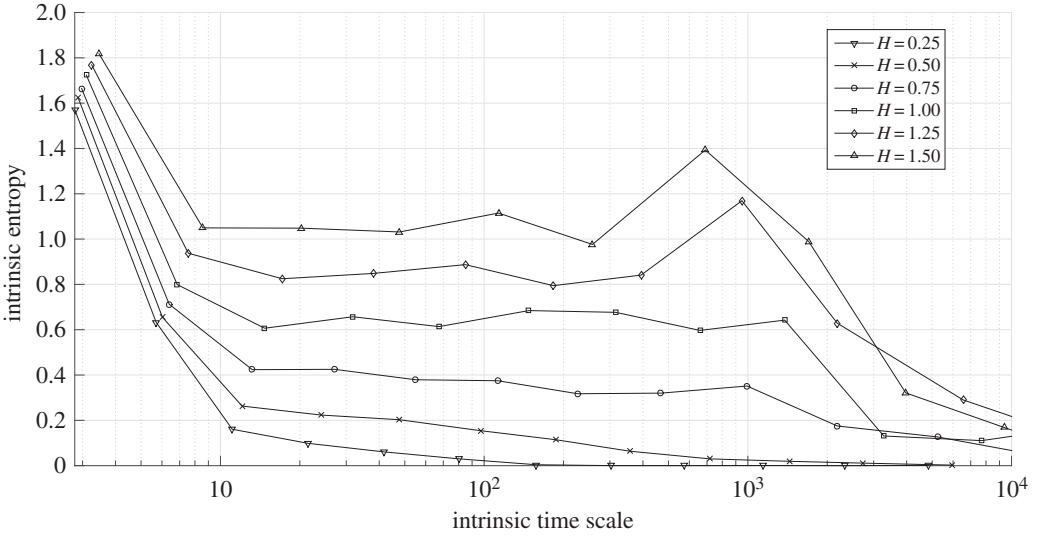

Figure 6. Intrinsic entropies versus their corresponding intrinsic time scales for simulated time series of fractal Gaussian noise with different Hurst exponents.

entropy. An IMF with data length containing 100 periods can be coarse-grained to 340 samples, satisfying the minimum requirement for an entropy calculation. Therefore, only the intrinsic entropy for an IMF with over 100 periods can be accurately quantified in this approach. As shown in figure 6, the expressions using multiple intrinsic entropies were used to display the complexity of fractal noise. Here, the data length of the simulated time series of fractal noise is 32 768. For IMFs with an intrinsic time scale of around 1000 samples, the required coarse-graining time scale is 294 samples and the minimum requirement for data length is 29 400 samples. Therefore, the intrinsic entropies with intrinsic time scales over 1000 are inaccurate, because the data length of the coarse-grained time series is insufficient for the calculation of entropy.

4. Intrinsic entropies for human heartbeat time series

In addition, we also applied this new approach in human heartbeat analysis. Human heartbeat time series for subjects with different physiological and pathological conditions can be downloaded from the website of PhysioNet [23]. A total of 141 heartbeat time series from subjects with different states of normal, ageing and diseases were downloaded from the Automated Teller Machine of PhysioBank. The heartbeat time series of 72 healthy subjects were downloaded from the Normal Sinus Rhythm RR Interval Database and the Massachusetts Institute of Technology/Beth Israel Hospital (now the Beth Israel Deaconess Medical Center, BIDMC) (MIT-BIH) Normal Sinus Rhythm Database; the time series of 44 subjects with congestive heart failure (CHF) were download from the Congestive Heart Failure RR Interval Database and the BIDMC Congestive Heart Failure Database; and 25 recordings from patients with atrial fibrillation (AF) were downloaded from the MIT-BIH Atrial Fibrillation Database. The healthy subjects were divided into two groups by age: 44 subjects aged 60 (66.2 ± 3.7) years or older for the group of healthy elderly and the other 28 subjects aged 36.39 ± 9.4 years for the group of healthy young. A total of 44 patients with CHF were divided into two subgroups of CHF I–II and CHF III–IV according to the criteria of New York Heart Association functional classification. Actually, the heartbeat time series are sequences of heartbeat intervals but not true time series.

Figure 7 shows the entropy increment matrices with a logarithmic coarse-graining scale as the x-axis and a logarithmic intrinsic time scale as the y-axis. Similar to the results of the analysis of the simulated time series of fractal noise, intrinsic entropies contributed by each specific IMF were evaluated with a specific coarse-graining scale. The specific coarse-graining scale depends on the intrinsic time scale of the corresponding IMF. It implies that the specific coarse-graining scale for deriving the intrinsic entropy can be determined by the intrinsic time scale of the corresponding IMF. For both simulated and real-world situations, such as human heartbeat time series, the correlations between the specific coarse-graining scales and intrinsic time scales are consistent. Therefore, the entropy increment matrix for a human heartbeat time series can be simplified to a set of intrinsic entropies with their corresponding intrinsic time scales.

Figure 8 shows the statistical results of intrinsic entropies of the first seven IMFs for human heartbeat time series of subjects with different physiological and pathological states. The heartbeat time series of subjects with AF show significant intrinsic entropies only for the first two IMFs. This implies that a set of intrinsic entropies for the heartbeat time series of a subject with AF is similar to that for noise with a Hurst exponent equal to or less than 0.5. A subject with AF has an anti-correlated or non-correlated heartbeat time series.

Healthy subjects (i.e. healthy young and healthy elderly subjects) behave with high intrinsic entropies on all intrinsic time scales. As a measure of complexity as the complexity index of MSE, the summation of intrinsic entropies was used as a new index of complexity in this new approach. The new complexity index for the healthy young group is significantly higher than that for the other groups with ageing and/or disease states. This result implies that loss of complexity is a significant sign of ageing and disease states, which is consistent with the results obtained by MSE. Furthermore, it is interesting to distinguish the features of intrinsic entropies on different scales for heartbeat time series affected by ageing and disease.

In this study, the healthy young group was chosen as the baseline for comparison with the other groups with ageing and heart diseases. The feature of intrinsic entropies for the heartbeat time series of the healthy young and healthy elderly groups represents a baseline with a healthy condition. Furthermore, the intrinsic time scales of intrinsic entropies with significant differences in comparison with the baseline provide important clues to determine the control mechanisms that are affected by ageing or a specific disease. Table 1 shows the comparison results between heartbeat time series for the healthy young and healthy elderly groups by the approach using intrinsic entropy. According to the results, only the first two intrinsic entropies show significant differences in statistics for these two groups. The p-value of the t-test for the intrinsic entropy of IMF 1 is 0.02 and that for IMF 2 is less than 0.0001, as shown in table 1. Furthermore, the correlation coefficient between age and the intrinsic entropy of IMF 2 is -0.5231. These statistical

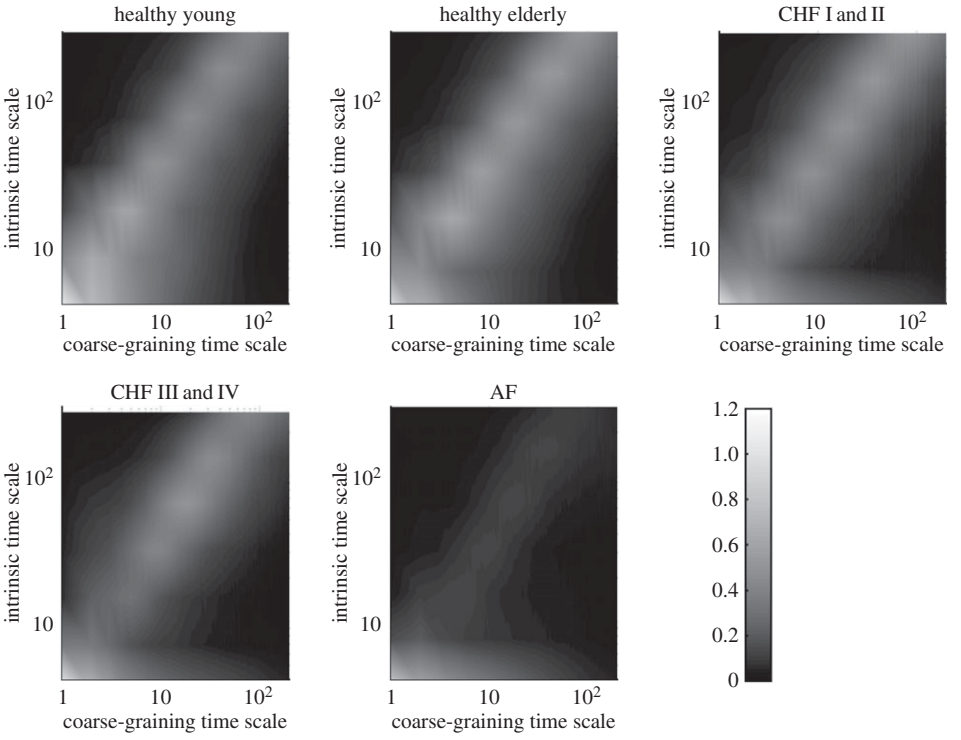

Figure 7. The entropy increment matrices for heartbeat time series from subjects with different physiological and pathological states.

Table 1. The results of comparisons between the intrinsic entropies of the first seven IMFs of the heartbeat time series for two groups of healthy young and healthy elderly. n.s., not significant.

group	IMF 1	IMF 2	IMF 3	IMF 4	IMF 5	IMF 6	IMF 7
young	1.39 ± 0.29	0.75 ± 0.19	0.71 ± 0.13	0.60 ± 0.09	0.55 ± 0.09	0.59 ± 0.12	0.63 ± 0.12
elderly	1.24 ± 0.24	0.56 ± 0.14	0.65 ± 0.17	0.60 ± 0.12	0.59 ± 0.12	0.63 ± 0.13	0.67 ± 0.16
p-value by t-test	0.02	<0.001	n.s.	n.s.	n.s.	n.s.	n.s.

results show a significantly negative correlation between the intrinsic entropies of IMF 2 and age. Ageing causes the loss of intrinsic entropy of IMF 2.

To determine which mechanism was potentially affected by ageing, the control mechanisms with working time scales similar to the intrinsic time scale of IMF 2 is the possible mechanism affected by ageing. The intrinsic time scale of IMF 2 is around eight beats, which is close to 6.4 s per cycle (0.156 Hz) as the high-frequency (HF) band of heart rate variability (HRV). The HF band of HRV reflects the balance of sympathetic activation and vagal modulation. Detrended fluctuation analysis (DFA) $\alpha 1$ (with scale range from 3 to 11 beats) performs as a good assessment of sympathovagal modulation [24]. The correlation between DFA $\alpha 1$ and the intrinsic entropies of IMF 2 for the 72 healthy subjects is shown in figure 9. The Pearson correlation coefficient between these two parameters is 0.7938. This result shows a strong correlation between DFA $\alpha 1$ and the intrinsic entropy of IMF 2. Therefore, the intrinsic entropy of IMF 2 also represents a new measure correlated to sympathovagal modulation. Furthermore, the correlation between DFA $\alpha 1$ and age was checked using the Pearson correlation coefficient and the t-test. The Pearson correlation coefficient between age and DFA $\alpha 1$ is -0.4406 and the p-value of the t-test for DFA $\alpha 1$

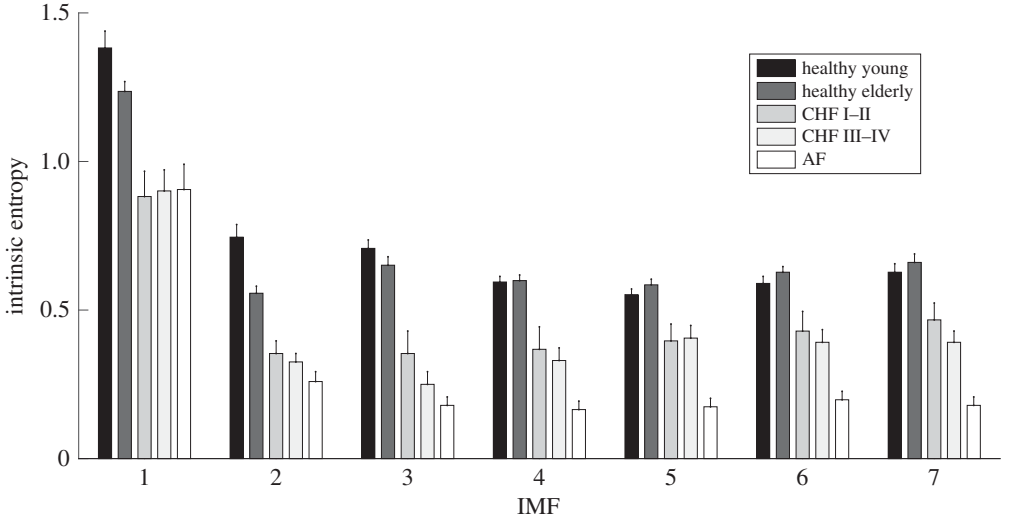

Figure 8. Statistical results of the intrinsic entropies of the first seven IMFs for the human heartbeat time series of five groups with different physiological and pathological states. Data shown are the mean and standard error.

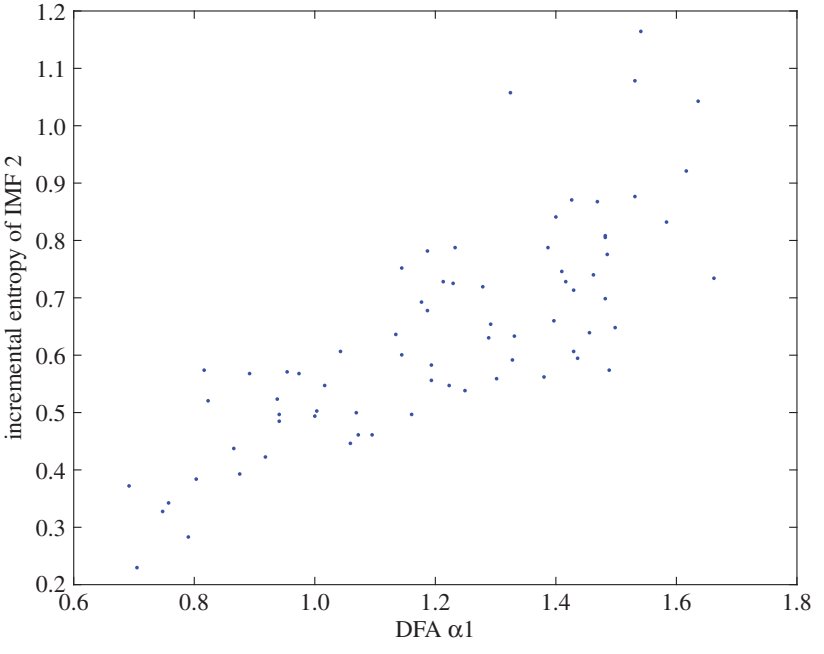

Figure 9. Illustration of the correlation between DFA $\alpha 1$ and the incremental entropy of IMF 2. (Online version in colour.)

for the two groups of healthy young and healthy elderly is 0.0002. Both the correlation coefficient and the *t*-test result show that the correlation between the intrinsic entropy of IMF 2 and age is more significant than that between DFA $\alpha 1$ and age.

In addition, heart diseases cause a decrease in the intrinsic entropies of the heartbeat time series on all intrinsic time scales as shown in figure 8. The loss of complexity can be observed as the decreases in the intrinsic entropies of the first seven IMFs. On the other hand, we were also interested in what underlying mechanism was potentially affected by CHF. Cheyne–Stokes respiration (CSR) is an abnormal pattern of breathing and the pattern repeats with each cycle

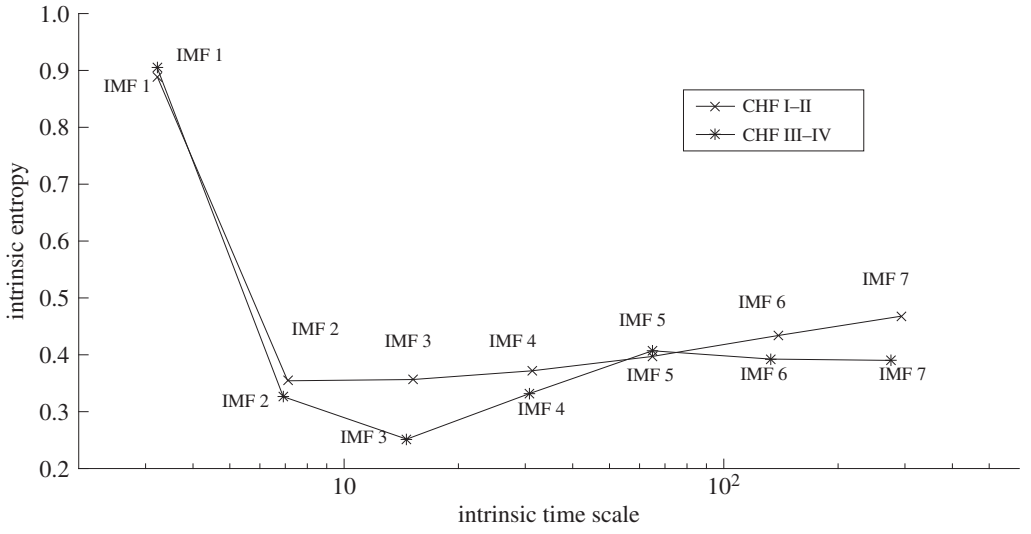

Figure 10. The differences between the intrinsic entropies of the first seven IMFs for the groups with mild CHF (CHF I–II) and severe CHF (CHF III–IV).

Table 2. Means of the intrinsic time scales for the first seven IMFs for the two groups with mild and severe CHF.

group	IMF 1	IMF 2	IMF 3	IMF 4	IMF 5	IMF 6	IMF 7
CHF I–II	3.22	7.10	15.16	31.23	64.72	139.21	293.16
CHF III–IV	3.21	6.91	14.61	30.80	64.84	132.88	275.76

usually taking 30 s to 2 min [25]. CSR is a common symptom of CHF. Therefore, we considered that the intrinsic entropies of IMFs with intrinsic time scales from 36 beats to 180 beats should be potential parameters for distinguishing the two groups of CHF I–II and CHF III–IV, because the severities of CHF are different. Considering the disease state of the subjects with CHF III–IV is more severe than that of the subjects with CHF I–II, the complexities of the subjects with severe CHF should be lower than those of subjects with mild CHF. Moreover, as subjects with severe CHF potentially have more serious CSR symptoms than subjects with mild CHF, the intrinsic entropies should be affected on the corresponding time scales. The corresponding time scales of CSR include the intrinsic time scales of IMFs 4–6, as shown in table 2. Therefore, the intrinsic entropies of IMFs 3–6 should reflect the influences of CSR. Figure 10 shows the intrinsic entropies of IMFs 1–7 for the two groups with mild and severe CHF. Because the intrinsic entropy of IMF 3 reflects the entropy outside the influence of CSR, the intrinsic entropy of IMF 3 for subjects with severe CHF is lower than that for subjects with mild CHF. However, the statistical difference is not significant (p-value of t-test is 0.15). The intrinsic entropies of IMFs 4–6, within the time-scale range affected by CSR, are significantly higher than the intrinsic entropy of IMF 3 for the subjects with severe CHF. The difference between the intrinsic entropies of IMFs 3 and 5 is significant for subjects with mild and severe CHF, as shown in figure 10. Therefore, the ratio of the intrinsic entropy of IMF 5 to the intrinsic entropy of IMF 3 is defined as a new parameter for verifying the difference between heartbeat time series of groups with mild and severe CHF. The mean and standard deviation of this ratio is 2.02 ± 1.24 for subjects with severe CHF and 1.22 ± 0.34 for subjects with mild CHF. The p-value of the t-test is 0.03, which represents a significant difference between the two groups of mild and severe CHF.

In addition, a new complexity index can be defined as the summation of all intrinsic entropies on multiple intrinsic time scales in this scale-dependent approach. As shown in figure 11, the

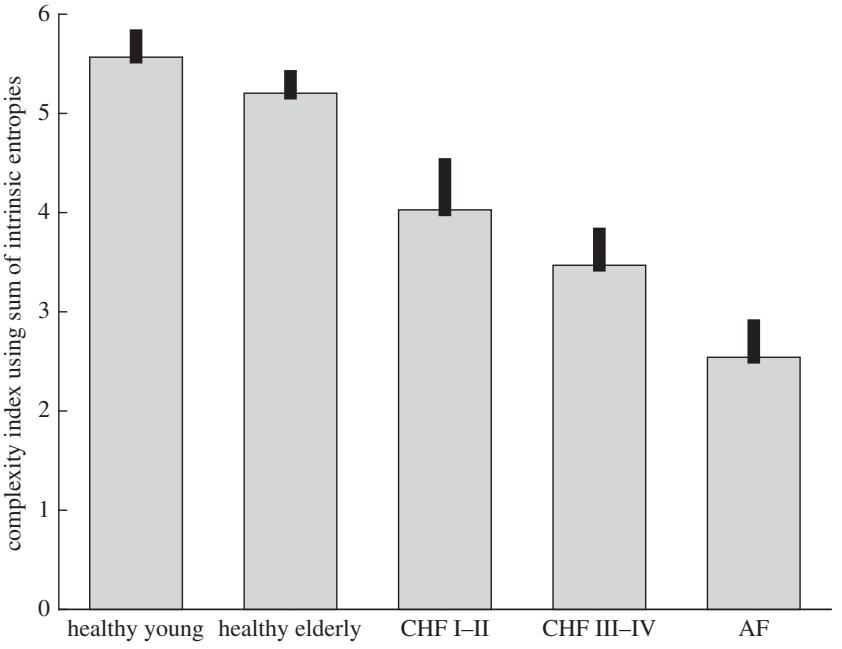

Figure 11. Complexity index using summation of the intrinsic entropies for five groups.

value of the complexity index derived by this scale-dependent approach for the healthy young group is higher than that of the other three groups. Both ageing and heart diseases reduce the complexity of heartbeat time series. These results are similar to the results for the complexity index derived in the MSE analysis.

5. Discussions and conclusion

In this investigation, a new approach by scale-dependent intrinsic entropies was proposed for quantifying the entropies of a time series on different intrinsic time scales. The intrinsic time scales are defined as the averaged periods of IMFs decomposed by EMD. This new approach combined the functions of EMD and MSE. EMD works by decomposing a time series into a set of IMFs with different intrinsic time scales, and a new parameter term as intrinsic entropy was used for quantifying the entropy contributed by a specific IMF on its corresponding intrinsic time scale. The intrinsic entropy represents the relative term between two successive detrended time series but not an additive measure. The value of the intrinsic entropy is a function of the standard deviation of a detrended time series (SD_n). The problem of mode-mixing is an issue arising from using EMD for detrending in this approach. Mode-mixing causes inconsistency of a mode function, which affects the standard deviation of a detrended time series and the outcomes of scale-dependent intrinsic entropy analysis.

In MSE analysis, a complexity index is defined using the summation of entropies on multiple coarse-graining time scales. In the scale-dependent intrinsic entropy analysis, the total of intrinsic entropies on all intrinsic time scales also represents a new complexity index. A complex time series is considered to have high intrinsic entropies on multiple time scales. Comparing the methodologies of MSE and the scale-dependent approach, MSE used the variation (SD) of the last detrended time series (the detrended time series including all IMFs) for the entropy calculations on multiple coarse-graining scales. However, the scale-dependent intrinsic entropies are calculated using the variations of the detrended time series with different intrinsic time scales. According to the results of the MSE analysis for fractal noises, as shown in figure 2, the fractal properties of the simulated noises are represented by the slopes of the MSE plots. The standard

variation of a fractal noise with a high Hurst exponent (Hurst exponent greater than 1) generated by integration of fractal Gaussian noise is large. A large standard deviation reduces the entropies on all coarse-graining scales and results in a low complexity index in MSE analysis. In the scale-dependent approach, the fractal properties are consistent on successive intrinsic time scales, as shown in figure 6, because the variations also change with the intrinsic time scales. Therefore, the results of the complexity analysis by MSE and the scale-dependent approach are incongruous for the fractal noise with a Hurst exponent greater than 1.

For the simulated time series of fractal noise, the fractal property is shown as consistent intrinsic entropies on most of the intrinsic time scales excluding the first IMF, as shown in figure 6. However, the intrinsic entropy of IMF 1 is calculated by MSE directly, and not by the definition mentioned in the section on methodology, because $C_1(t)$ is equal to $x_1(t)$. Theoretically, the intrinsic entropies should be consistent on different scales for a simulated noise with consistent fractal properties on different scales. Therefore, the intrinsic entropy actually reflects the fractal property of the simulated noises.

On the other hand, how many intrinsic entropies can be accurately determined for a time series? For a time series with a limited number of samples, we can obtain a set of IMFs by EMD. As mentioned in the section on methodology, an appropriate coarse-graining scale should be used in the calculation of intrinsic entropy for an IMF and the coarse-graining scale depends on the intrinsic time scale. A long coarse-graining scale reduces the length of the data according to the method of coarse-graining. However, a sufficient length of data is required for an accurate calculation of SampEn. For a coarse-grained time series without sufficient data length, the value of the entropy is not accurate. Therefore, it should be noted that only the intrinsic entropy calculated by a coarse-grained time series with sufficient data length can represent the entropy contributed by an IMF. In figure 6, we can observe the declines in the intrinsic entropies for the last four IMFs because of the insufficient data length of coarse-grained time series.

In the analysis of the results of the human heartbeat time series, the total of the intrinsic entropies represents a new complexity index. According to our results, both ageing and heart disease decreased the complexity of the human heartbeat time series. This is consistent with the results reported in previous studies by MSE. Furthermore, the intrinsic entropy of IMF 2 for the human heartbeat time series is significantly correlated with age and DFA $\alpha 1$. DFA $\alpha 1$ is considered to be a measure of sympathovagal modulation according to previous studies. Therefore, the intrinsic entropy of IMF 2 also performed a new measure similar to DFA $\alpha 1$. In this study, a negative correlation between the intrinsic entropy of IMF 2 and age has been found. This implies that the functional reduction in sympathetic activation and vagal modulation is a sign of ageing.

Finally, the ratio of the intrinsic entropy of IMF 5 to the intrinsic entropy of IMF 3 could be used as a new index for distinguishing the difference between subjects with mild and severe CHF. CSR is thought to act as a control mechanism, which causes fluctuations in heart rate as a result of the abnormal breathing pattern on the corresponding time scales. The sympathovagal effect and CSR represent two different modulations with different intrinsic time scales in the heartbeat time series. In this study, these two modulations were used to describe the potential applications of intrinsic entropy for investigating the entropy on a specific time scale with its corresponding control mechanism.

Data accessibility. In this study, the simulated time series of fractal Gaussian noises were generated using the algorithm proposed by Wood & Chen [19]. The database of the human heartbeat time series was downloaded from the PhysioNet website [23].

Authors' contributions. J.-R.Y. proposed the concept and design, and participated in the data analysis and drafting. N.E.H. participated in the modification of the concept, and interpretation of data. C.-K.P. participated in revising the article critically for important intellectual content. All authors gave final approval of the version to be published.

Competing interests. The authors declare that they have no competing interests.

Funding. We received no funding for this study.

Acknowledgements. The authors are grateful for support from the NSC (Taiwan, ROC), grant no. 99-2627-B-008-003, the joint foundation of the NCU, grant no. CNJRF-99CGH-NCU-A3, VGHUST100- G1-4-3, and NSC

support for the Center for Dynamical Biomarkers and Translational Medicine, National Central University, Taiwan (NSC 99-2911-I-008-100).

References

1. Costa M, Golberger AL, Peng C-K. 2002 Multiscale entropy analysis of complex physiologic time series. *Phys. Rev. Lett.* **89**, 068102. (doi:10.1103/PhysRevLett.89.068102)
2. Costa M, Goldberger AL, Peng C-K. 2005 Multiscale entropy analysis of biological signals. *Phys. Rev. E* **71**, 021906. (doi:10.1103/PhysRevE.71.021906)
3. Peng C-K, Costa M, Goldberger AL. 2009 Adaptive data analysis of complex fluctuations in physiologic time series. *Adv. Adapt. Data Anal.* **1**, 61–70. (doi:10.1142/S1793536909000035)
4. Varotsos PA, Sarlis NV, Skordas ES, Lazaridou MS. 2005 Natural entropy fluctuations discriminate similar-looking electric signals emitted from systems of different dynamics. *Phys. Rev. E* **71**, 011110. (doi:10.1103/PhysRevE.71.011110)
5. Li Z, Zhang YK. 2008 Multi-scale entropy analysis of Mississippi River flow. *Stoch. Environ. Res. Risk Assess.* **22**, 507–512. (doi:10.1007/s00477-007-0161-y)
6. Thuraisingham RA, Gottwald GA. 2006 On multiscale entropy analysis for physiological data. *Physica A, Stat. Mech. Appl.* **366**, 323–332. (doi:10.1016/j.physa.2005.10.008)
7. Yuan H-K, Lin C, Tsai P-H, Chang F-C, Lin K-P, Hu H-H, Su M-C, Lo M-T. 2011 Acute increase of complexity in the neurocardiovascular dynamics following carotid stenting. *Acta Neurol. Scand.* **123**, 187–192. (doi:10.1111/j.1600-0404)
8. Huang NE, Wu Z, Pinzon JE. 2009 Reductions of noise and uncertainty in annual global surface temperature anomaly data. *Adv. Adapt. Data Anal.* **1**, 447–460. (doi:10.1142/S1793536909000151)
9. Wu Z, Huang NE, Long SR, Peng C-K. 2007 On the trend, detrending, and variability of nonlinear and nonstationary time series. *Proc. Natl Acad. Sci. USA* **104**, 14 889–14 894. (doi:10.1073/pnas.0701020104)
10. Richman JS, Moorman JR. 2000 Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol.* **278**, H2039–H2049.
11. Pincus SM. 2001 Assessing serial irregularity and its implications for health. *Ann. N.Y. Acad. Sci.* **954**, 245, and references therein. (doi:10.1111/j.1749-6632.2001.tb02755.x)
12. Huang NE, Shen Z, Long S-R, Wu M-C, Shih H-H, Zheng Q. 1998 The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. Lond. A.* **454**, 903–995. (doi:10.1098/rspa.1998.0193)
13. Yeh J-R, Fan S-Z, Shieh J-S. 2009 Human heart beat analysis using a modified algorithm of detrended fluctuation analysis based on empirical mode decomposition. *Med. Eng Phys.* **31**, 92–100. (doi:10.1016/j.medengphy.2008.04.011)
14. Wu Z, Huang NE. 2004 A study of the characteristics of white noise using the empirical mode decomposition method. *Proc. R. Soc. Lond. A* **460**, 1597–1611. (doi:10.1098/rspa.2003.1221)
15. Mandelbort BB. 1971 A fast fractional gaussian noise generator. *Water Resour. Res.* **7**, 543–553. (doi:10.1029/WR007i003p00543)
16. Paxson V. 1997 Fast, approximate synthesis of fractional Gaussian noise for generating self-similar network traffic. *ACM SIGCOMM, Comput. Commun. Rev.* **27**, 5–18. (doi:10.1145/269790.269792)
17. Huang C, Devetsikiotis M, Lambadaris I, Kaye AR. 1995 Fast simulation for self-similar traffic in ATM networks. *IEEE Int. Conf. Commun. ICC95.* **1**, 439–444. (doi:10.1109/ICC.1995.525208)
18. Jeong HDJ, McNickle D, Pawlikowski K. 1999 Fast self-similar teletraffic generation based on FGN and wavelets. In *Proc. IEEE Int. Conf. on Networks* (*ICON 99*), Brisbane, Australia, 28 September–1 October 1999, pp. 75–82. Piscataway, NJ: IEEE.
19. Wood AT, Chan G. 1994 Simulation of stationary Gaussian processes in $[0, 1]^d$. *J. Comput. Graph. Stat.* **3**, 409–432. (doi:10.1080/10618600.1994.10474655)
20. Li CF. 2003 Rescaled-range and power spectrum analyses on well-logging data. *Geophys. J. Int.* **153**, 201–212. (doi:10.1046/j.1365-246X.2003.01893.x)
21. Shtatland ES. 1991 Fractal stochastic models for acoustic impedance: an explanation of scaling or 1/f geology and stochastic inversion. *Soc. Explor. Geophys.* 1598–1601. (doi:10.1190/1.1888857)

22. Turcotte DL. 1997 *Fractals and chaos in geology and geophysics*. Cambridge, UK: Cambridge University Press.
23. Goldberger AL *et al*. 2000 PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation* **101**, E215–E220. (doi:10.1161/01.CIR.101.23.e215)
24. Tulppo MP, Kiviniemi AM, Hautala AJ, Kallio M, Seppänen T, Mäkikallio TH, Huikuri HV. 2005 Physiological background of the loss of fractal heart rate dynamics. *Circulation* **112**, 314–319. (doi:10.1161/CIRCULATIONAHA.104.523712)
25. Francis DP, Willson K, Davies LC, Coats AJS, Piepoli M. 2000 Quantitative general theory for periodic breathing in chronic heart failure and its clinical implications. *Circulation* **102**, 2214–2221. (doi:10.1161/01.CIR.102.18.2214)