

Article

Multivariate Generalized Multiscale Entropy Analysis

Anne Humeau-Heurtier

Univ Angers, LARIS-Laboratoire Angevin de Recherche en Ingénierie des Systèmes, 62 avenue Notre-Dame du Lac, 49000 Angers, France; anne.humeau@univ-angers.fr; Tel.: +33-244-687-587

Academic Editor: Kevin H. Knuth

Received: 22 September 2016; Accepted: 14 November 2016; Published: 17 November 2016

Abstract: Multiscale entropy (MSE) was introduced in the 2000s to quantify systems' complexity. MSE relies on (i) a coarse-graining procedure to derive a set of time series representing the system dynamics on different time scales; (ii) the computation of the sample entropy for each coarse-grained time series. A refined composite MSE (rcMSE)—based on the same steps as MSE—also exists. Compared to MSE, rcMSE increases the accuracy of entropy estimation and reduces the probability of inducing undefined entropy for short time series. The multivariate versions of MSE (MMSE) and rcMSE (MrcMSE) have also been introduced. In the coarse-graining step used in MSE, rcMSE, MMSE, and MrcMSE, the mean value is used to derive representations of the original data at different resolutions. A generalization of MSE was recently published, using the computation of different moments in the coarse-graining procedure. However, so far, this generalization only exists for univariate signals. We therefore herein propose an extension of this generalized MSE to multivariate data. The multivariate generalized algorithms of MMSE and MrcMSE presented herein (MG MSE and MGrcMSE, respectively) are first analyzed through the processing of synthetic signals. We reveal that MGrcMSE shows better performance than MG MSE for short multivariate data. We then study the performance of MGrcMSE on two sets of short multivariate electroencephalograms (EEG) available in the public domain. We report that MGrcMSE may show better performance than MrcMSE in distinguishing different types of multivariate EEG data. MGrcMSE could therefore supplement MMSE or MrcMSE in the processing of multivariate datasets.

Keywords: complexity; nonlinear dynamics; entropy; multivariate embedding; multiscale entropy

1. Introduction

Multiscale entropy (MSE) was proposed in the 2000s to quantify the degree of unpredictability of systems across multiple scales [1,2]. Its computation relies on two steps [1,2]: (i) a coarse-graining procedure to derive a set of time series representing the system dynamics on different time scales. The coarse-graining procedure for scale τ is obtained by *averaging* the samples of the time series inside consecutive but non-overlapping windows of length τ . The length of a coarse-grained time series at a scale factor τ is equal to N/τ , where N is the number of samples in the original signal. (ii) Computation of the sample entropy for each coarse-grained time series. Sample entropy is equal to the negative of the natural logarithm of the conditional probability that sequences close to each other for m consecutive data points will also be close to each other when one more point is added to each sequence [3].

Since its introduction, MSE has been used in a variety of applications [4]. However, to increase the accuracy of entropy estimation, and to reduce the probability of inducing undefined entropy for short time series, a refined composite multiscale entropy (rcMSE) has recently been proposed [5]. rcMSE is based on the same steps as MSE (see below).

Recently, Costa et al. reported a generalization of the MSE algorithm [6]: they proposed MSE_n , where n corresponds to the moment used to coarse-grain a time series. MSE_μ uses the mean value

(first moment) and corresponds to the standard MSE. In the generalized MSE algorithm, the second, third, or higher moment can be used in the non-overlapping segments of length τ [6]. Using MSE_{σ^2} , it has been shown that human heartbeat volatility time series exhibit complex bursting behaviors over a wide range of time scales, and that multiscale complexity of the volatility degrades with aging and pathology [6].

MSE and its generalization are able to process univariate datasets only. For multivariate time series, individual time series are processed separately. This may be satisfactory only if the individual signals are statistically independent or at least uncorrelated. This is often not the case when real world signals from a given system are registered simultaneously. To overcome this shortcoming, the multivariate MSE (MMSE) and multivariate rcMSE (MrcMSE) have been proposed [7–9]. MMSE and MrcMSE are able to operate on any number of data channels, and provide a robust relative complexity measure for multivariate data [7–9]. However, MMSE and MrcMSE only consider coarse-grained time series from the first-order moment of the samples contained inside windows.

In this work, we extend the MMSE and MrcMSE algorithms to a more general case. To this end, we introduce the multivariate generalized MSE (MG MSE) and the multivariate generalized rcMSE (MGrcMSE), and evaluate their performance on both synthetic and real-world multivariate processes.

2. Materials and Methods

2.1. Multiscale Entropy

As mentioned above, the MSE algorithm is composed of two steps [1,2]

1. A coarse-graining procedure. For a monovariate discrete signal of length N $\{x_1, \dots, x_i, \dots, x_N\}$, the coarse-grained time series $\{y^{(\tau)}\}$ is computed as

$$y_j^{(\tau)} = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} x_i, \quad 1 \leq j \leq \lfloor N/\tau \rfloor. \quad (1)$$

For scale one, the coarse-grained time series $\{y^{(1)}\}$ corresponds to the original signal. The length of the coarse-grained time series $\{y^{(\tau)}\}$ is $\lfloor N/\tau \rfloor$. Coarse-graining can be seen as an averaging of the data inside a window of length τ (to reduce the high frequency components), followed by a downsampling of the averaged data by a factor τ [10]. Moreover, it has been reported that Equation (1) is similar to the use of a finite-impulse response (FIR) filter on the original time series x and to the downsampling of the filtered signal with a factor τ [11]. This FIR filter is a low-pass filter.

2. Computation of the sample entropy for each coarse-grained time series. The sample entropy quantifies the regularity of finite length time series [3]. A low value for the sample entropy reflects a high degree of regularity, while a random signal has a relatively higher value of sample entropy. Sample entropy is a conditional probability measure that quantifies the likelihood that a sequence of m consecutive data points—that matches another sequence of the same length (match within a tolerance of r)—will still match the other sequence when their length is increased by one sample (sequences of length $m + 1$); m therefore defines the length of the patterns that are compared to each other [3]. For a time series $\{x_1, \dots, x_i, \dots, x_N\}$, the sample entropy is computed as

$$\text{SampEn}(x, m, r) = -\ln \frac{n^{m+1}}{n^m}, \quad (2)$$

where n^m represents the total number of m -dimensional matched vector pairs, and n^{m+1} represents the total number of $(m + 1)$ -dimensional matched vector pairs. More precisely,

$$n^m = \frac{1}{N - m} \sum_{i=1}^{N-m} B_i^m(r), \tag{3}$$

with

$$B_i^m(r) = \frac{n_i^m(r)}{N - m - 1}, \tag{4}$$

and $n_i^m(r)$ is the number of vectors $x_m(j)$ such that $d[x_m(i), x_m(j)] \leq r$. In the definition of sample entropy, the distance d between two vectors is defined as the maximum absolute difference of their corresponding scalar components [3].

From Equation (2), MSE can be written as

$$MSE(x, \tau, m, r) = - \ln \frac{n_\tau^{m+1}}{n_\tau^m}, \tag{5}$$

where n_τ^m represents the total number of m -dimensional matched vector pairs, and is constructed from the coarse-grained time series at the scale factor τ . In the MSE algorithm, two patterns are considered similar if they are closer than a parameter r . The value of r is usually chosen as a percentage of the standard deviation of the signal under study. In the original algorithm of MSE proposed by Costa et al. (algorithm described above; [1,2]), the value of r is constant for all scale factors.

2.2. Refined Composite Multiscale Entropy

MSE has been used widely. However, because the coarse-graining procedure reduces the length of the time series, MSE may yield an inaccurate estimation of entropy or induce undefined entropy at large scales or for short time series. rcMSE has been proposed to overcome this drawback [5]. Compared to MSE, rcMSE increases the accuracy of entropy estimation and reduces the probability of inducing undefined entropy for short time series. For a monovariate discrete signal of length N $\{x_1, \dots, x_i, \dots, x_N\}$, the rcMSE algorithm is based on the following steps [5]:

1. For a scale factor τ , τ coarse-grained time series are generated. The k -th coarse-grained time series for a scale factor τ is defined as $y_k^{(\tau)} = \{y_{k,j}\}_{j=1}^{N/\tau}$, where [12]

$$y_{k,j}^{(\tau)} = \frac{1}{\tau} \sum_{i=(j-1)\tau+k}^{j\tau+k-1} x_i, \quad 1 \leq j \leq \frac{N}{\tau}, \quad 1 \leq k \leq \tau. \tag{6}$$

2. For each scale factor τ , and for all τ coarse-grained time series, the number of matched vector pairs $n_{k,\tau}^{m+1}$ and $n_{k,\tau}^m$ is computed, where $n_{k,\tau}^m$ represents the total number of m -dimensional matched vector pairs and is computed from the k -th coarse-grained time series at a scale factor τ
3. rcMSE is then defined as [5]

$$rcMSE(x, \tau, m, r) = - \ln \left(\frac{\sum_{k=1}^{\tau} n_{k,\tau}^{m+1}}{\sum_{k=1}^{\tau} n_{k,\tau}^m} \right). \tag{7}$$

2.3. Multivariate (Refined Composite) Multiscale Entropy

MMSE, the extension of MSE to multivariate datasets—denoted as $MGMSE_\mu$ in what follows—relies on the same steps as MSE [7,8]: a coarse-graining procedure and a sample entropy computation for each coarse-grained time series. Each of these two steps is adapted to multivariate data. Thus, for the coarse-graining procedure, temporal scales are defined by *averaging* a p -variate time series $\{x_{l,i}\}_{i=1}^N$ ($l = 1, \dots, p$ is the channel index and N is the number of samples in every channel)

over time segments of increasing length. Thus, for a scale factor τ , a multi-channel coarse-grained time series is computed as (here the channel index l goes from 1 to p)

$$y_{l,j}^{(\tau)} = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} x_{l,i}, \quad 1 \leq j \leq \lfloor N/\tau \rfloor. \quad (8)$$

In the sample entropy computation step, the multivariate sample entropy (MSampEn) is used. The MSampEn algorithm is an extension of the univariate sample entropy [3]. In this case, the embedding dimension for the channel k of the multivariate data is noted as m_k . The detailed MSampEn algorithm can be found in [7,8]. Briefly, for a tolerance level r , MSampEn is calculated as the negative of the natural logarithm of the conditional probability that two composite delay vectors close to each other in a m dimensional space will also be close to each other when the dimensionality is increased by one.

A multivariate version of rcMSE ($M\text{GrcMSE}_\mu$) has also recently been proposed [9]. It is based on the same steps as rcMSE, but each of the steps is adapted to multivariate signals. It has been reported that $M\text{GrcMSE}_\mu$ outperforms $M\text{GMSE}_\mu$ because it leads to lower standard deviation values than $M\text{GMSE}_\mu$ [9].

2.4. Multivariate Generalized (Refined Composite) Multiscale Entropy

In the multivariate extension of the generalized MSE ($M\text{GMSE}$) and generalized rcMSE ($M\text{GrcMSE}$) that we propose, the same steps as the ones of $M\text{MSE}$ (resp. $M\text{rcMSE}$) are found, except in the coarse-graining procedure, where the averaging is replaced by the computation of the second moment ($M\text{GMSE}_{\sigma^2}$ and $M\text{GrcMSE}_{\sigma^2}$) or by the computation of the third moment ($M\text{GMSE}_{\text{skewness}}$ and $M\text{GrcMSE}_{\text{skewness}}$), and so on. $M\text{GMSE}_{\sigma^2}$ and $M\text{GrcMSE}_{\sigma^2}$ quantify the dynamics of the volatility over multiple time scales [6]. Thus, for $M\text{GMSE}_{\sigma^2}$, the coarse-grained multivariate time series at scale factor τ is computed as

$$y_{l,j}^{(\tau)} = \frac{1}{\tau - 1} \sum_{i=(j-1)\tau+1}^{j\tau} (x_{l,i} - \mu_{l,j}^\tau)^2, \quad (9)$$

where $\mu_{l,j}^\tau$ corresponds to the mean of $\{x_{l,i}\}_{i=(j-1)\tau+1}^{j\tau}$. $M\text{GMSE}_{\text{skewness}}$ is determined as in Equation (9) where the computation of the variance is replaced by the computation of the skewness (similar steps are used to compute $M\text{GrcMSE}_{\sigma^2}$ and $M\text{GrcMSE}_{\text{skewness}}$, see Figure 1). From Equation (9), we observe that $M\text{GMSE}_{\sigma^2}$ and $M\text{GrcMSE}_{\sigma^2}$ are 0 for scale factor $\tau = 1$. In what follows, we will therefore concentrate on scale factors larger than 1 for $M\text{GMSE}_{\sigma^2}$ and $M\text{GrcMSE}_{\sigma^2}$. For the same reason, we will concentrate on scale factors larger than 2 for $M\text{GMSE}_{\text{skewness}}$ and $M\text{GrcMSE}_{\text{skewness}}$. Hereafter we will study $M\text{GMSE}_{\sigma^2}$, $M\text{GrcMSE}_{\sigma^2}$, $M\text{GMSE}_{\text{skewness}}$, and $M\text{GrcMSE}_{\text{skewness}}$ on synthetic and real datasets.

2.5. Datasets Acquisition

The new algorithms that we propose were tested on synthetic signals, but also on two publicly available experimental biomedical datasets [13,14]. The first set consisted of five two-channel electroencephalogram (EEG) data (bivariate data) recorded at the left and right frontal cortex of male adult WAG/Rij rats. Each signal lasted 5 s and was digitized at 200 Hz. Each one was filtered between 1 and 100 Hz. More details on the acquisition procedure can be found in [15,16]. For this first set, data A correspond to a normal EEG, and data B, C, D, and E contain spike-wave discharges which are the landmark of epileptic activity. These spike-wave discharges are due to abnormal synchronization in an epileptic brain, even when there are no seizures [16]. The synchronization level differs between the data, as has been previously reported [16]. The length of each data segment was only 5 s (1000 samples), because this length was considered to be the largest one in which the signals containing spikes could be visually judged as stationary [15].

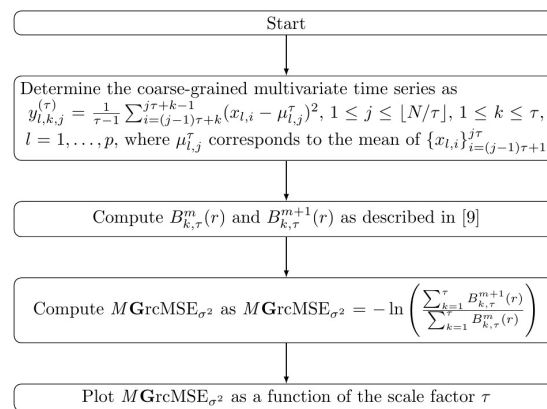


Figure 1. Flow chart describing the computation of multivariate extension of generalized refined composite multiscale entropy ($MGrcMSE_{\sigma^2}$) for a p -variate time series $\{x_{l,i}\}_{i=1}^N$ ($l = 1, \dots, p$ is the channel index, and N is the number of samples in every channel).

The second set of experimental biomedical data consisted of five (data A to E) two-channel EEG data (bivariate data) recorded in humans [17]. The segments processed thereafter were selected and cut out from continuous multichannel EEG recordings after visual inspection for artifacts (e.g., due to muscle activity or eye movements). In addition, the segments had to fulfill a stationarity criterion [17]. Data A and B are segments taken from surface EEG data recorded in healthy volunteers. Volunteers were relaxed in an awake state with eyes open (data A) and eyes closed (data B). Data C, D, and E are EEG from patients who had achieved complete seizure control after resection of one of the hippocampal formations, which was therefore correctly diagnosed to be the epileptogenic zone. Data C were recorded from the hippocampal formation of the opposite hemisphere of the brain, and data D were recorded from within the epileptogenic zone. Data C and D contain only activity measured during seizure-free intervals. Data E contain only seizure activity. The data were selected from all recording sites exhibiting ictal activity. All these data were recorded during 23.6 s with a sampling frequency of 173.61 Hz. They therefore contain 4097 samples. Further details on the acquisition can be found in [17].

3. Experimental Results and Discussion

3.1. Results for Synthetic Signals

We recall that $1/f$ noise time series are more complex than uncorrelated (white) noise time series due to long-range correlations of $1/f$ noise data [1]. Furthermore, a multivariate time series is considered more structurally complex than another if (for most of the scale factors) its multivariate entropy values are higher than those of the other time series. When the multivariate entropy values decrease with the scale factor, the data that are processed contain information only at the smallest scales. They are thus not structurally complex.

Moreover, multivariate data showing a constant entropy measure or a monotonic increase of entropy value with increasing scale factors are data presenting long-range correlations.

We first computed $MGMSE_{\sigma^2}$, $MGrcMSE_{\sigma^2}$, $MGMSE_{\text{skewness}}$, and $MGrcMSE_{\text{skewness}}$ of a trivariate time series, where all the data channels were originally realizations of mutually independent white noise [7]. We then gradually decreased the number of variates representing white noise (from 3 to 0), and simultaneously increased the number of data channels representing independent $1/f$ noise (from 0 to 3), as already proposed in [7–9]. The total number of variates was always three. The embedding dimension m_k was chosen to be equal to 2, and the threshold r was fixed to $0.15 \times$ (standard deviation of the normalized time series) for each data channel [9]. For each kind of

trivariate data, 10 independent realizations were simulated. For each realization, 15,000 samples were generated in each variate. Scales were lower than 40. Therefore, all the coarse-grained time series were longer than 300 samples, as recommended by others [8,18].

$MGMSE_{\mu}$ and $MGrCMSE_{\mu}$ for similar kinds of data have been previously reported [7,9]: the multivariate sample entropy decreases with increasing scale factors, whatever the composition of the multivariate data (see Figures 2a,b). Moreover, for scale factors larger than 2, the higher the number of variates representing $1/f$ noise, the higher the multivariate entropy value. We observe that the results obtained with $MGMSE_{\mu}$ and $MGrCMSE_{\mu}$ are similar. However, for the largest scales, we observe some wandering on $MGMSE_{\mu}$ that is not visible on $MGrCMSE_{\mu}$.

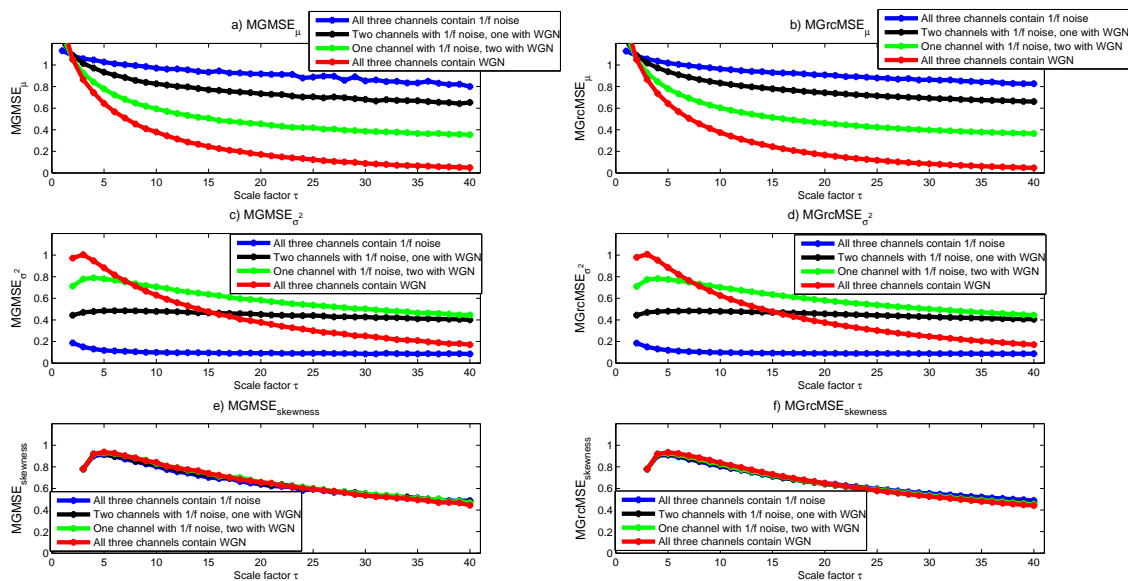


Figure 2. (a) multivariate extension of the generalized MSE ($MGMSE_{\mu}$); (b) $MGrCMSE_{\mu}$; (c) $MGMSE_{\sigma^2}$; (d) $MGrCMSE_{\sigma^2}$; (e) $MGMSE_{skewness}$; and (f) $MGrCMSE_{skewness}$ average values for ten trivariate data containing white Gaussian noise (WGN) and $1/f$ noise, each with 15,000 samples.

The results given by $MGMSE_{\sigma^2}$ and $MGrCMSE_{\sigma^2}$ are shown in Figures 2c,d. We note a slight increase and then a decrease of the multivariate sample entropy with increasing scale factors, except for the trivariate data containing only $1/f$ noise, where a nearly constant multivariate sample entropy is observed for increasing scale factors. For scales lower than 40, the multivariate data with channels containing only white noise show higher multivariate sample entropy than multivariate data with channels containing only $1/f$ noise. Others have reported similar findings for univariate data [19]. Therefore, the scale for which the multivariate data containing only $1/f$ noise shows larger entropy values than those of multivariate data containing only white noise is larger than what was observed with $MGMSE_{\mu}$: for scales lower than 40, trivariate data containing channels with only $1/f$ noise have a lower multivariate sample entropy than trivariate data containing channels with only white noise; for $MGMSE_{\mu}$, for scales larger than 2, trivariate data containing channels with only $1/f$ noise have a larger multivariate sample entropy than trivariate data containing channels with only white noise (see Figures 2a,b). If we focus on the decrease rate observed for $MGMSE_{\sigma^2}$ and $MGrCMSE_{\sigma^2}$, we can say that the results are consistent with what was expected: the more the multivariate data contain $1/f$ noise, the slower the decrease of the multivariate sample entropy values with increasing scale factors ($1/f$ data are theoretically more complex than white noise, because $1/f$ data contain long-range correlations [2]). Moreover, we observe that the results obtained with $MGMSE_{\sigma^2}$ and $MGrCMSE_{\sigma^2}$ are rather similar.

The results obtained by computing $MGMSE_{skewness}$ and $MGrCMSE_{skewness}$ are shown in Figures 2e,f. We observe that—whatever the composition of the multivariate dataset—the multivariate

sample entropy first increases until scale factor $\tau = 5$, and then decreases. For $MGrMSE_{skewness}$, the most rapid decrease is observed for data with channels containing only white noise. The slowest decrease is observed for data with channels containing only $1/f$ noise. For scale factors larger than 20, the highest multivariate sample entropy is observed for data with channels containing only $1/f$ noise, and the lowest for data with channels containing only white noise. The global trend for $MGMSE_{skewness}$ is similar to the one of $MGrMSE_{skewness}$, but the different rates of decrease for each variate are not as clear as for $MGrMSE_{skewness}$.

The same analysis was also performed for synthesized bivariate data (as are the biomedical datasets processed thereafter): originally, the two data channels were realizations of mutually independent white noise [7]. We then gradually decreased the number of variates representing white noise (from 2 to 0), and simultaneously increased the number of data channels representing independent $1/f$ noise (from 0 to 2). The shortest experimental bivariate data processed thereafter (EEG data) contain 1000 samples. Therefore, we computed $MGMSE$ and $MGrMSE$ for two different synthesized bivariate data: one containing 15,000 samples and another containing 1000 samples. Our goal is to analyze if the results obtained for the two lengths (15,000 and 1000 samples) are similar. For this purpose, we considered only scales lower than 10, as data with only 1000 samples do not reasonably allow us to go beyond this limit. Thus, for data with 15,000 samples, the shortest coarse-grained time series has a length of 1500 samples. For data with 1000 samples, the shortest coarse-grained time series has a length of 100 samples. For this last case, the length of the coarse-grained time series will therefore be lower than 300, as recommended by others [8,18]. With a length of 100 samples, the results obtained with $MGrMSE$ should still be acceptable, as shown in [9]. For each length (15,000 samples and 1000 samples), 10 independent realizations were simulated.

For data with 15,000 samples, $MGMSE_{\mu}$ and $MGrMSE_{\mu}$ show consistent results with those observed for trivariate datasets (see Figures 3a,b). However, for data with 1000 samples, we observe that $MGMSE_{\mu}$ leads to undefined entropy values (see Figure 3c). $MGrMSE_{\mu}$ shows similar values to those obtained with 15,000 samples—except for the bivariate data containing only $1/f$ noise, where wandering are observed (see Figure 3d). However, for scales lower than 10, the results obtained with $MGrMSE_{\mu}$ are still acceptable.

For $MGMSE_{\sigma^2}$ and $MGrMSE_{\sigma^2}$ (see Figure 4), the bivariate data containing only $1/f$ time series present almost constant multivariate sample entropy values with increasing scale factors. However, for the bivariate data containing only white noise and for the bivariate data containing both $1/f$ noise and white noise, we observe an increase and then a decrease of the multivariate sample entropy for increasing scale factors. We also observe that $MGMSE_{\sigma^2}$ shows different results for data of 15,000 samples and for data of 1000 samples. However, $MGrMSE_{\sigma^2}$ shows close findings for the two lengths.

For $MGMSE_{skewness}$ and $MGrMSE_{skewness}$, the results show that—whatever the composition of the bivariate dataset—the behavior is the same as scales increase (see Figure 5). We observe an increase of the multivariate sample entropy with increasing scale factors and then a progressive decrease. However, as above, $MGrMSE_{skewness}$ shows more stable results for short time series than $MGMSE_{skewness}$ where wandering are observed.

From all these results obtained with a bivariate dataset, we observe that the findings obtained with 15,000 samples and with 1000 samples show differences when the data are processed with $MGMSE$. When $MGrMSE$ is used, the differences are much less pronounced. For scales lower than 10, $MGrMSE$ computed with 1000 samples leads to consistent results with those obtained with 15,000 samples. This is true for $MGrMSE_{\mu}$, $MGrMSE_{\sigma^2}$, and $MGrMSE_{skewness}$. In what follows, we will therefore concentrate on the refined versions of the multivariate generalized entropy measures: $MGrMSE_{\mu}$, $MGrMSE_{\sigma^2}$, and $MGrMSE_{skewness}$.

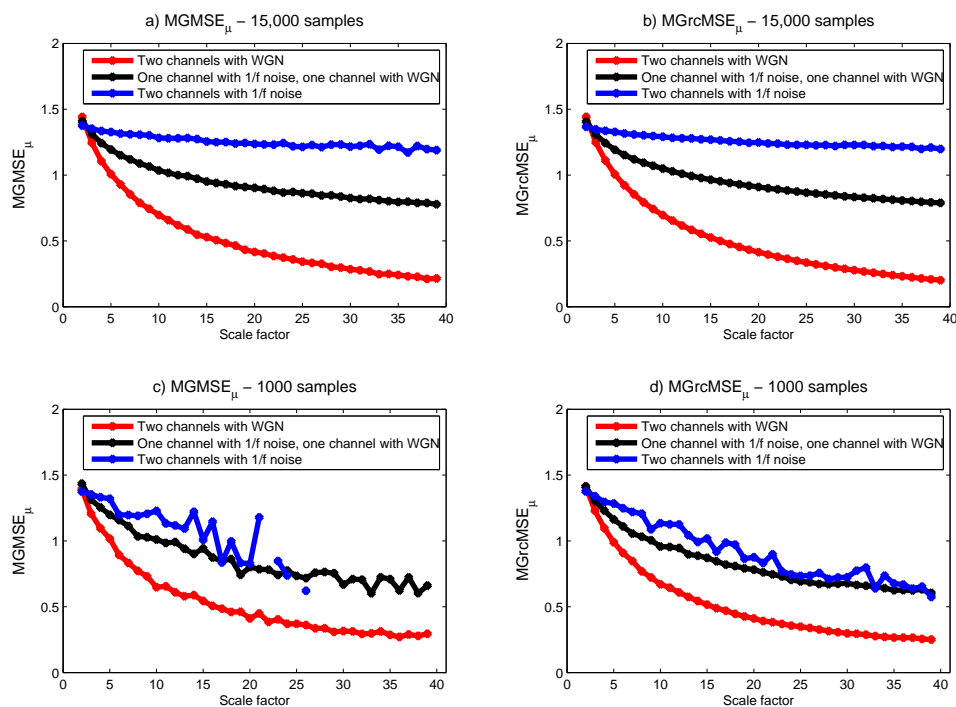


Figure 3. $MGMSE_{\mu}$ (left column) and $MGrcMSE_{\mu}$ (right column) average values for ten bivariate data containing white Gaussian noise (WGN) and $1/f$ noise, each with 15,000 samples (upper panels) and 1000 samples (lower panels).

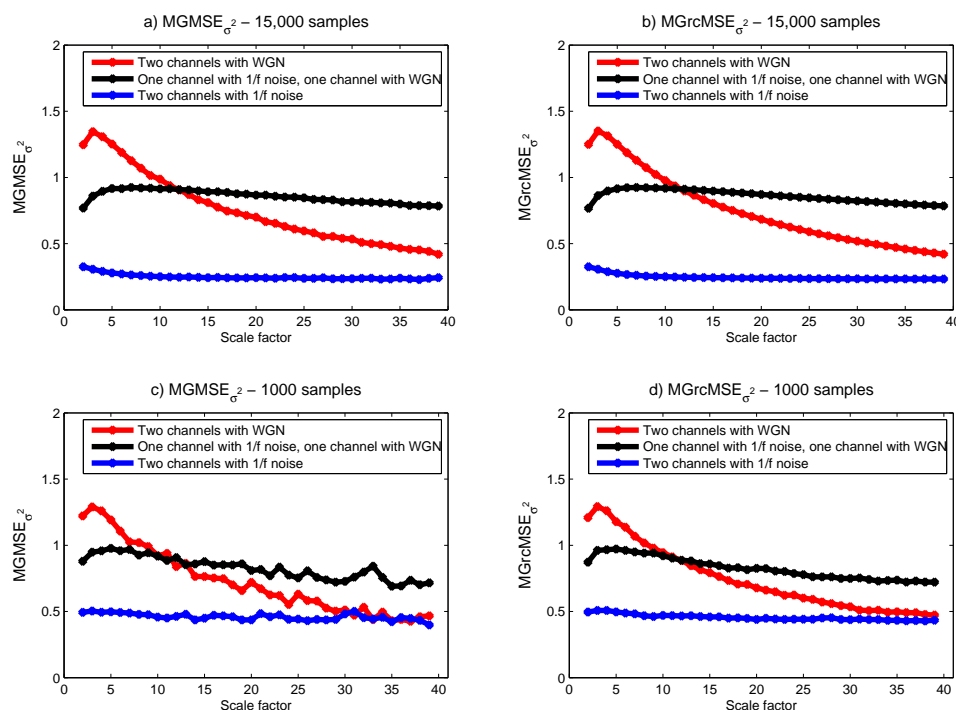


Figure 4. $MGMSE_{\sigma^2}$ (left column) and $MGrcMSE_{\sigma^2}$ (right column) average values for ten bivariate data containing white Gaussian noise (WGN) and $1/f$ noise, each with 15,000 samples (upper panels) and 1000 samples (lower panels).

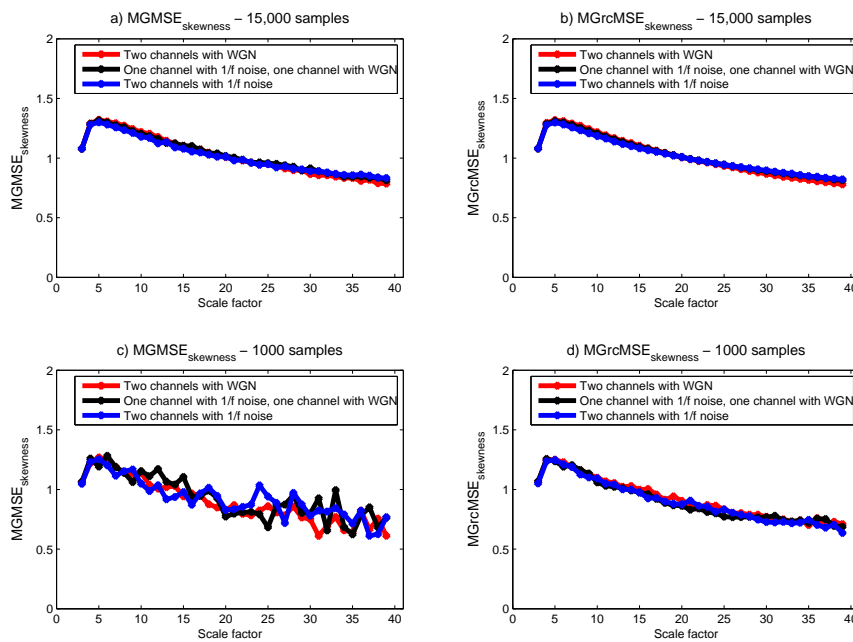


Figure 5. $MGMSE_{skewness}$ (left column) and $MGrCMSE_{skewness}$ (right column) average values for ten bivariate data containing white Gaussian noise (WGN) and $1/f$ noise, each with 15,000 samples (upper panels) and 1000 samples (lower panels).

3.2. Results for Biomedical Datasets

For the first set of EEG data (EEG recorded in rats), the results obtained applying MSE to each channel independently are presented in Figure 6. We note that the sample entropy values for the five pairs of rat EEG signals are close to each other. From this, it seems therefore difficult to distinguish the five cases. The results obtained with $MGrCMSE_{\mu}$, $MGrCMSE_{\sigma^2}$, and $MGrCMSE_{skewness}$ are presented in Figure 7. We observe that $MGrCMSE_{\mu}$, $MGrCMSE_{\sigma^2}$, and $MGrCMSE_{skewness}$ show different patterns for each of the five cases. However, the highest differences between the five cases are given by $MGrCMSE_{\sigma^2}$. In order to more easily quantify these differences, we propose the computation of a complexity index, defined here as the sum of the (multivariate) sample entropy for all scales studied. The computation of a similar index has been performed in other studies (e.g., [20–23]). For the univariate case, the sum was computed for each channel independently, and then the mean of the two sums was considered. The results for each case are presented in Figure 8. We observe that $MGrCMSE_{\sigma^2}$ leads to a better distinction of the different cases than $MGrCMSE_{\mu}$. Data B, C, D, and E all contain spike-wave discharges. Therefore, we could have expected to obtain similar signatures for $MGrCMSE_{\sigma^2}$ and/or for $MGrCMSE_{skewness}$. However, we obtain different results on these datasets. This is consistent with other synchronization measures which were also able to distinguish them [15]. Contents of data B, C, D, and E are therefore different. Our findings suggest that the multivariate multiscale complexity of the volatility may be more interesting than $MGrCMSE_{\mu}$ when studying multivariate datasets.

For the second set of EEG data (EEG recorded in humans), the results obtained applying MSE to each channel independently are presented in Figure 9. We note that the sample entropy of the five cases are close to each other. Therefore, their differentiation is hardly possible. The results obtained with $MGrCMSE_{\mu}$, $MGrCMSE_{\sigma^2}$, and $MGrCMSE_{skewness}$ are presented in Figure 10. We observe that $MGrCMSE_{\mu}$, $MGrCMSE_{\sigma^2}$, and $MGrCMSE_{skewness}$ facilitate the differentiation between the five cases. However, the highest differences between the five cases are given by $MGrCMSE_{\sigma^2}$. As above (and with the same definition as above), we computed the complexity index. The results for each case are presented in Figure 11. As previously, we observe that $MGrCMSE_{\sigma^2}$ leads to a better distinction of the five cases than $MGrCMSE_{\mu}$.

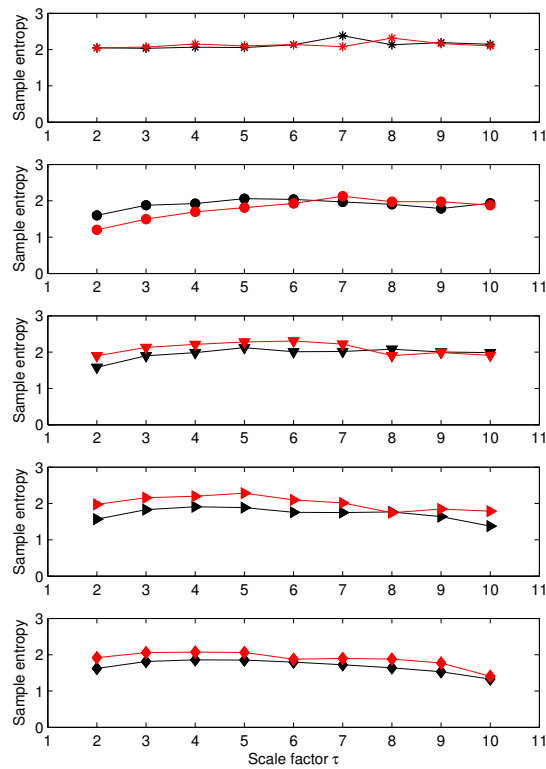


Figure 6. MSE values for five electroencephalogram (EEG) cases (data A to E, from top to bottom). For each subplot, each curve corresponds to a data channel. Data A correspond to a normal EEG, and data B, C, D, and E contain spike-wave discharges [15,16].

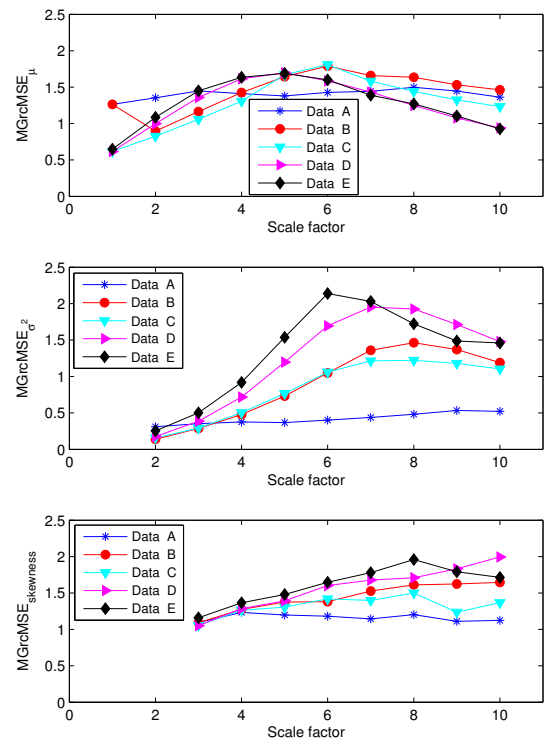


Figure 7. From top to bottom: $MGrcMSE_{\mu}$, $MGrcMSE_{\sigma^2}$, and $MGrcMSE_{skewness}$ for five pairs of rat EEG signals—see Figure 6.

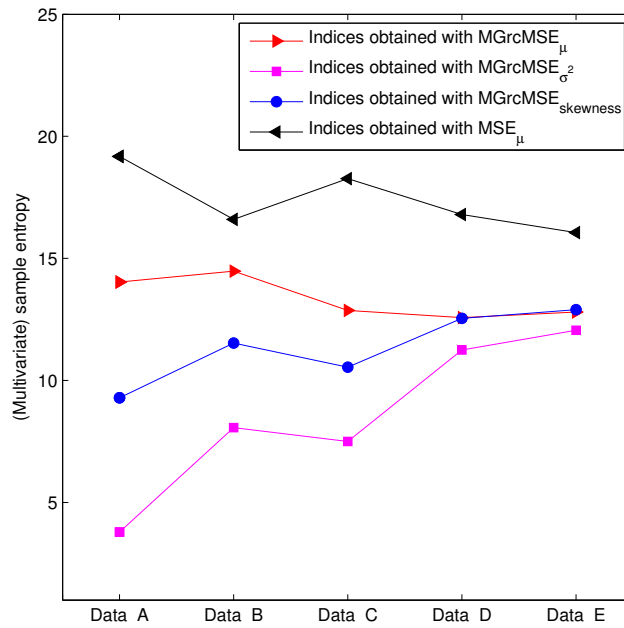


Figure 8. Complexity index for five EEG cases—see Figure 6.

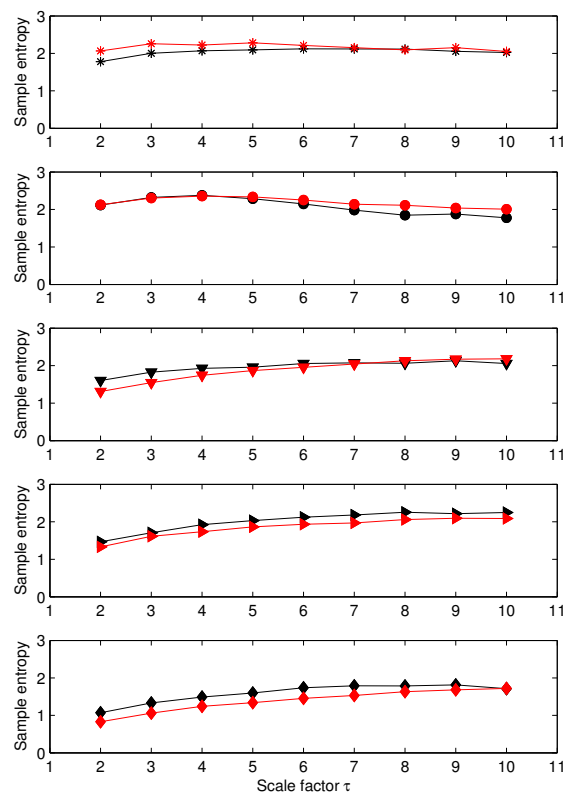


Figure 9. MSE values for five EEG cases (data A to E, from top to bottom). For each subplot, each curve corresponds to a data channel. See text for details on each data set, and [17] for the data acquisition procedure.

As mentioned above, the coarse-graining procedure can be seen as a FIR (low-pass filter) on the original time series and to the downsampling of the filtered signal with a factor τ [11]. This multiscale procedure is reminiscent of multiscale transforms such as wavelet, ridgelet, or curvelet transforms [24–27]. Such transforms can also be used with entropy measures.

Our work is still in progress. The analysis performed on the synthetic data has to be continued, and many more real datasets have to be analyzed in order to obtain a better understanding of the potentialities of MGMSE and MGrcMSE. However, the first findings presented herein are encouraging. Nevertheless, our study presents limitations. A first limitation is the non-existence of an inverse transform for this multiscale transform. Another limitation is the computational time of MGMSE (and even more, the one of MGrcMSE): this computational time prevents real-time analyses. This is particularly annoying, as we need to record long time series if large scales are to be analyzed.

Therefore, for the future, several directions could be studied. To improve the time complexity of MSE, several algorithms have been proposed [28–31]. It could now be interesting to adapt these algorithms to the multivariate case, and to study MGMSE and MGrcMSE based on these faster codes. Furthermore, the original MSE relies on a sample entropy-based approach, but can be used with different types of entropic measures: permutation entropy, cross-approximate entropy, compression entropy, etc. The multivariate generalization proposed herein could be extended to MSE based on other entropic measures. In addition to rcMSE, several other variants to MSE have been proposed: refined MSE [11], composite MSE [12], modified MSE for short-term time series [10], and short time MSE [32], to cite only a few. The multivariate approach of these variants could now be proposed, as well as their generalization to higher moments. For our simulated time series, $MGrcMSE_{skewness}$ leads to close values, whatever the composition of the multivariate dataset. These results have to be studied more thoroughly in the future. However, we observe that different findings are obtained for EEG data. For the univariate case, $MSE_{skewness}$ has already been shown to give interesting results for the biomedical field [33].

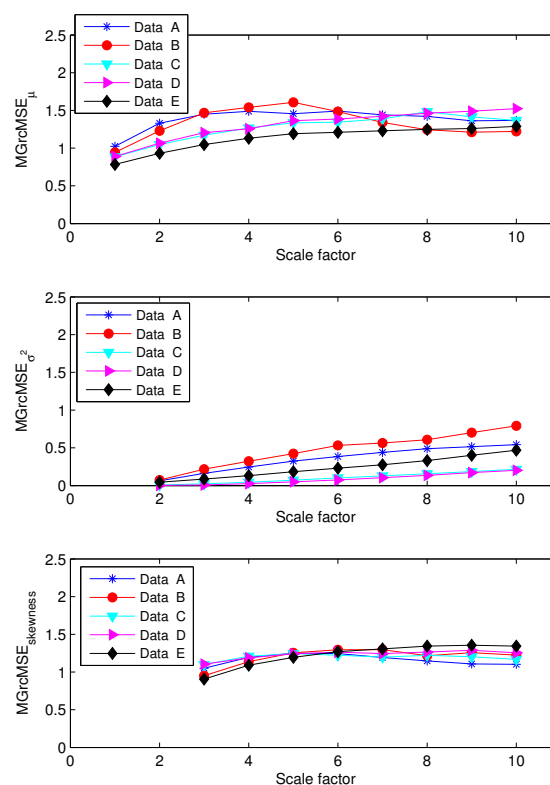


Figure 10. From top to bottom: $MGrcMSE_{\mu}$, $MGrcMSE_{\sigma^2}$, and $MGrcMSE_{skewness}$ for five pairs of human EEG signals—see Figure 9.

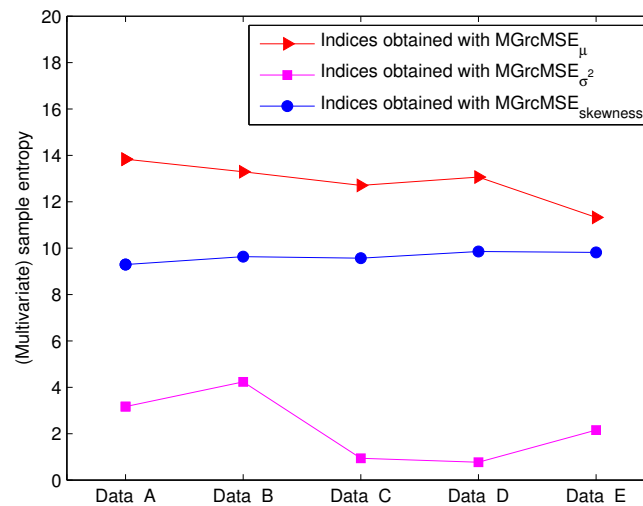


Figure 11. Complexity index for five EEG cases—see Figure 9.

4. Conclusions

MSE is now widely used in a large number of areas. Its refined version (rcMSE) has also been proposed to improve the results when processing short time series. Their adaptation to multivariate data exist (MMSE and MrcMSE). MSE, rcMSE, MMSE, and MrcMSE rely (among others) on a coarse-graining procedure to obtain a set of time series representing the system dynamics on different time scales. In this coarse-graining procedure, the computation of the mean (first moment) of samples inside consecutive windows is performed. A generalization of MSE has recently been proposed. This generalization consists of using other moments in the coarse-graining procedure. Unfortunately, this generalization did not exist for multivariate datasets. We therefore herein proposed an extension of this generalized MSE to multivariate datasets. We first analyzed the behavior of MGMSE and MGrcMSE on synthetic signals, and processed two different publicly available bivariate EEG data sets. Our results show that MGrcMSE may present better performance than MrcMSE in differentiating different types of EEG signals. MGrcMSE could therefore become an interesting signal processing tool for multivariate datasets. Moreover, we have shown that MGMSE and MGrcMSE lead to similar results when long data are analyzed. However, MGrcMSE shows better performance than MGMSE when short data are processed. The time complexity of MGrcMSE being worse than the one of MGMSE, we suggest the use of MGMSE for long data and MGrcMSE for short data.

In our work, the parameter r was kept constant, as done in other studies [6,19]. However, the coarse-graining procedure is similar to smoothing and decimation of the original sequences [34]. Because the same parameter r is used for all scales, in MSE, some authors concluded that the changes in MSE on each scale depend on both the regularity and variation of the coarse-grained time series [34]. More work could therefore be conducted to analyze the results when adapting the parameter r to each scale or for each moment used [33].

Our new method has been applied to EEG data. After a deeper analysis of synthetic time series, multivariate magnetoencephalography data could also be analyzed, but also data recorded simultaneously by different modalities. This could help in the diagnosis of pathologies that affect several organs or systems. For example, the simultaneous acquisition of data from the peripheral and central cardiovascular systems and their analysis with MG(rc)MSE could help in the understanding and diagnosis of some cardiovascular diseases. MG(rc)MSE could also be used for many other kinds of data: financial time series, chemical data, etc.

Acknowledgments: Anne Humeau-Heurtier would like to acknowledge Thomas Kreuz for fruitful discussions.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Costa, M.; Goldberger, A.L.; Peng, C.K. Multiscale entropy analysis of complex physiologic time series. *Phys. Rev. Lett.* **2002**, *89*, 068102.
2. Costa, M.; Goldberger, A.L.; Peng, C.K. Multiscale entropy analysis of biological signals. *Phys. Rev. E* **2005**, *71*, 021906.
3. Richman, J.S.; Moorman, J.R. Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol. Heart Circ. Physiol.* **2000**, *278*, H2039–H2049.
4. Humeau-Heurtier, A. The multiscale entropy algorithm and its variants: A review. *Entropy* **2015**, *17*, 3110–3123.
5. Wu, S.D.; Wu, C.W.; Lin, S.G.; Lee, K.Y.; Peng, C.K. Analysis of complex time series using refined composite multiscale entropy. *Phys. Lett. A* **2014**, *378*, 1369–1374.
6. Costa, M.D.; Goldberger, A.L. Generalized multiscale entropy analysis: Application to quantifying the complex volatility of human heartbeat time series. *Entropy* **2015**, *17*, 1197–1203.
7. Ahmed, M.U.; Mandic, D.P. Multivariate multiscale entropy: A tool for complexity analysis of multichannel data. *Phys. Rev. E* **2011**, *84*, 061918.
8. Ahmed, M.U.; Mandic, D.P. Multivariate multiscale entropy analysis. *IEEE Signal Process. Lett.* **2012**, *19*, 91–94.
9. Humeau-Heurtier, A. Multivariate refined composite multiscale entropy analysis. *Phys. Lett. A* **2016**, *380*, 1426–1431.
10. Wu, S.D.; Wu, C.W.; Lee, K.Y.; Lin, S.G. Modified multiscale entropy for short-term time series analysis. *Physica A* **2013**, *392*, 5865–5873.
11. Valencia, J.F.; Porta, A.; Vallverdú, M.; Claria, F.; Baranowski, R.; Orlowska-Baranowska, E.; Caminal, P. Refined multiscale entropy: Application to 24-h Holter recordings of heart period variability in healthy and aortic stenosis subjects. *IEEE Trans. Biomed. Eng.* **2009**, *56*, 2202–2213.
12. Wu, S.D.; Wu, C.W.; Lin, S.G.; Wang, C.C.; Lee, K.Y. Time series analysis using composite multiscale entropy. *Entropy* **2013**, *15*, 1069–1084.
13. Available online: <http://www2.le.ac.uk/departments/engineering/research/bioengineering/neuroengineering-lab/software> (accessed on 5 June 2016).
14. Available online: http://epileptologie-bonn.de/cms/front_content.php?idcat=193&lang=3 (accessed on 22 October 2016).
15. Quian Quiroga, R.; Kraskov, A.; Kreuz, T.; Grassberger, P. Performance of different synchronization measures in real data: A case study on electroencephalographic signals. *Phys. Rev. E* **2002**, *65*, 041903.
16. Quian Quiroga, R.; Kreuz, T.; Grassberger, P. Event synchronization: A simple and fast method to measure synchronicity and time delay patterns. *Phys. Rev. E* **2002**, *66*, 41904.
17. Andrzejak, R.G.; Lehnertz, K.; Rieke, C.; Mormann, F.; David, P.; Elger, C.E. Indications of nonlinear deterministic and finite dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state. *Phys. Rev. E* **2001**, *64*, 061907.
18. Ahmed, M.U.; Rehman, N.; Looney, D.; Rutkowski, T.M.; Mandic, D.P. Dynamical complexity of human responses: A multivariate data-adaptive framework. *Bull. Pol. Acad. Sci. Tech. Sci.* **2012**, *60*, 433–445.
19. Azami, H.; Fernandez, A.; Escudero, J. Refined Multiscale Fuzzy Entropy Based on Standard Deviation for Biomedical Signal Analysis. 2016, arXiv:1602.02847.
20. Schnettler, W.T.; Goldberger, A.L.; Ralston, S.J.; Costa, M. Complexity analysis of fetal heart rate preceding intrauterine demise. *Eur. J. Obstet. Gynecol. Reprod. Biol.* **2016**, *203*, 286–290.
21. Costa, M.; Ghiran, I.; Peng, C.-K.; Nicholson-Weller, A.; Goldberger, A.L. Complex dynamics of human red blood cell flickering: Alterations with in vivo aging. *Phys. Rev. E Stat. Nonlinear Soft Matter Phys.* **2008**, *78*, 020901.
22. Costa, M.D.; Henriques, T.; Munshi, M.N.; Segal, A.R.; Goldberger, A.L. Dynamical glucometry: Use of multiscale entropy analysis in diabetes. *Chaos* **2014**, *24*, 033139.
23. Costa, M.D.; Schnettler, W.T.; Amorim-Costa, C.; Bernardes, J.; Costa, A.; Goldberger, A.L.; Ayres-de-Campos, D. Complexity-loss in fetal heart rate dynamics during labor as a potential biomarker of acidemia. *Early Hum. Dev.* **2014**, *90*, 67–71.

24. Starck, J.-L.; Murtagh, F.; Fadili, J.M. *Sparse Image and Signal Processing: Wavelets and Related Geometric Multiscale Analysis*, 2nd ed.; Cambridge University Press: Cambridge, UK, 2015; pp. 1–423.
25. Shuman, D.I.; Faraji, M.J.; Vandergheynst, P. A multiscale pyramid transform for graph signals. *IEEE Trans. Signal Process.* **2016**, *64*, 2119–2134.
26. Xiao, M.; He, Z. Remote sensing image fusion based on gaussian mixture model and multiresolution analysis. *Proc. SPIE* **2013**, *8921*, doi:10.1117/12.2031385.
27. Grohs, P. Geometric multiscale decompositions of dynamic low-rank matrices. *Comput. Aided Geom. Des.* **2013**, *30*, 805–826.
28. Sugisaki, K.; Ohmori, H. Online estimation of complexity using variable forgetting factor. In Proceedings of the 2007 SICE Annual Conference, Takamatsu, Japan, 17–20 September 2007; pp. 1–6.
29. Samani, A.; Madeleine, P. Permuted sample entropy. *Commun. Stat. Simul. Comput.* **2010**, *39*, 1506–1516.
30. Pan, Y.H.; Lin, W.Y.; Wang, Y.H.; Lee, K.T. Computing multiscale entropy with orthogonal range search. *J. Mar. Sci. Technol.* **2011**, *19*, 107–113.
31. Jiang, Y.; Mao, D.; Xu, Y. A fast algorithm for computing sample entropy. *Adv. Adapt. Data Anal.* **2011**, *3*, 167–186.
32. Chang, Y.C.; Wu, H.T.; Chen, H.R.; Liu, A.B.; Yeh, J.J.; Lo, M.T.; Tsao, J.H.; Tang, C.J.; Tsai, I.T.; Sun, C.K. Application of a modified entropy computational method in assessing the complexity of pulse wave velocity signals in healthy and diabetic subjects. *Entropy* **2014**, *16*, 4032–4043.
33. Shi, W.; Shang, P.; Ma, Y.; Sun, S.; Yeh, C.H. A comparison study on stages of sleep: Quantifying multiscale complexity using higher moments on coarse-graining. *Commun. Nonlinear Sci. Numer. Simul.* **2017**, *44*, 292–303.
34. Nikulin, V.V.; Brismar, T. Comment on “Multiscale entropy analysis of complex physiologic time series”. *Phys. Rev. Lett.* **2004**, *92*, 089803.



© 2016 by the author; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).