Review

The Multiscale Entropy Algorithm and Its Variants: A Review

Anne Humeau-Heurtier

University of Angers, LARIS—Laboratoire Angevin de Recherche en Ingénierie des Systèmes, Angers 49016, France; E-Mail: anne.humeau@univ-angers.fr; Tel.: +33-244-687-587; Fax: +33-244-687-561

Abstract: Multiscale entropy (MSE) analysis was introduced in the 2002 to evaluate the complexity of a time series by quantifying its entropy over a range of temporal scales. The algorithm has been successfully applied in different research fields. Since its introduction, a number of modifications and refinements have been proposed, some aimed at increasing the accuracy of the entropy estimates, others at exploring alternative coarse-graining procedures. In this review, we first describe the original MSE algorithm. Then, we review algorithms that have been introduced to improve the estimation of MSE. We also report a recent generalization of the method to higher moments.

Keywords: entropy; complexity; nonlinear dynamics; multiscale entropy; refined multiscale entropy; composite multiscale entropy; generalized multiscale entropy; intrinsic mode entropy; hierarchical entropy; fuzzy sample entropy

1. Introduction

Complexity of time series can be studied through several measures, e.g., approximate entropy [1] and sample entropy [2]. However, some of these measures do not take into account the multiple time scales in physical systems. In the 2000s, the multiscale entropy (MSE) approach has been proposed by Costa *et al.* to represent the complexity of a signal [3]. MSE relies on the computation of the sample entropy over a range of scales: in the MSE algorithm, coarse-grained time series—that represent the system dynamics on different scales—are analyzed with the sample entropy algorithm. Since its introduction, MSE has become a prevailing method to quantify the complexity of signals. It has been used successfully in different research fields, including biomedical time series [4–6], electroseismic time series [7], vibration of rotary machine [8], financial time series [9]. However, the sample entropy

algorithm may give an imprecise estimation of entropy or even induce undefined entropy values for short time series (see, e.g., [10]). Moreover, it has been reported that the coarse-graining procedure (that eliminates the fast temporal scales) exploits a filter with a frequency response that cannot prevent aliasing and which is therefore suboptimal, especially in the presence of fast oscillations [11]. Finally, some authors mentioned that the entropy-based complexity is partially linked to the reduction of variance generated by the elimination of the fast temporal scales [11]. This is why, other algorithms—based on MSE—have been developed to try to improve the original MSE algorithm. We herein propose to present a state-of-the art of these algorithms that have been conceived to overcome some limitations of the original MSE.

2. Original Multiscale Entropy Algorithm

The MSE algorithm is composed of two steps [3,4]

(1) 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 i is obtained by averaging the samples of the time series inside consecutive but non overlapping windows of length i. Thus, for a monovariate discrete signal of length N $\{x_1, \ldots, x_i, \ldots, x_N\}$ the coarse-grained time series $\{y^{(\tau)}\}$ is computed as

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

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 N/τ.

(2) computation of the sample entropy for each coarse-grained time series

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 of one sample (sequences of length $m + 1$); m therefore defines the length of the patterns that are compared to each other [2]. In this definition, the distance between two vectors is computed as the maximum absolute difference of their corresponding scalar components [2]. More precisely, sample entropy is determined as

$$SampEn(m, r) = \lim_{N \to \infty} -\ln \frac{A^m(r)}{B^m(r)}, \tag{2}$$

where $A^m(r)$ is the probability that two sequences will match for $m + 1$ points and $B^m(r)$ is the probability that two sequences will match for m points (with a tolerance of r), where self-matches are excluded [2]. Equation (2) is estimated by the statistics [2]

$$SampEn(m, r, N) = -\ln \frac{A^m(r)}{B^m(r)}. \tag{3}$$

By setting $B = \{[(N - m - 1)(N - m)]/2\}B^m(r)$ and $A = \{[(N - m - 1)(N - m)]/2\}A^m(r)$ [2], we have $A/B = A^m(r)/B^m(r)$ and the sample entropy can be expressed as [2]

$$SampEn(m, r, N) = -\ln \frac{A}{B}, \tag{4}$$

where A is the total number of forward matches of length $m + 1$ whereas B is the total number of template matches of length m [2].

Regular and periodic signals have a theoretical sample entropy of 0 whereas uncorrelated random signals have a maximum entropy (the value depends on the signal length) [2]. The sample entropy for coarse-grained white noise time series can analytically be written as [4]

$$- \ln \int_{-\infty}^{+\infty} \frac{1}{2} \sqrt{\frac{\tau}{2\pi}} \Big[\text{erf} \Big(\frac{x + r}{\sqrt{2/\tau}} \Big) - \text{erf} \Big(\frac{x - r}{\sqrt{2/\tau}} \Big) \Big] \exp(-\frac{x^2 \tau}{2}) dx, \qquad (5)$$

where τ and $\text{erf}(\,\cdot\,)$ refer to the scale factor and to the error function, respectively. Moreover, Costa *et al.* have shown that the sample entropy of a coarse-grained $1/f$ noise is equal to 1.8 for all scales, for the following parameters values: $m = 2$, $r = 0.15\times$ the standard deviation, and $N = 30,000$ samples [4]. For scale one, white noise time series present a higher value of entropy compared to the one of $1/f$ time series. Moreover, the value of entropy for the coarse-grained white noise time series monotonically decreases and becomes smaller than the entropy value of coarse-grained $1/f$ time series for scale factor $\tau > 4$ [3]. These results underline that $1/f$ noise contains complex structures across multiple time scales whereas white noise does not.

However, it has been reported that the computation of MSE values as described in the above-mentioned algorithm presents several drawbacks (see, e.g., [11,12]):

- The coarse-graining procedure can be seen as a two-steps procedure [12]: (I) averaging the data inside a window of length τ in order to reduce the high frequency components; (II) downsampling of the averaged data by a factor τ. The coarse-graining procedure therefore reduces the length of the time series: at a scale factor τ, the coarse-grained time series has a length that is equal to the one of the original time series divided by τ. Therefore, the larger the scale factor, the shorter the coarse-grained time series. It has been reported that to obtain a reasonable entropy value, the time series length should be in the range of 10^m to 20^m [2]. For shorter time series, the variance of the entropy estimator grows very fast as the number of data points is reduced. Large variance of estimated entropy leads to a reduction of reliability. Consequently, the statistical reliability of sample entropy for the coarse-grained time series is reduced, as the scale factor τ increases. The sample entropy algorithm therefore leads to an imprecise estimation of entropy—unreliable results with great variance (errors), or even undefined entropy values (when no template vectors are matched to one another)—for short time series or at large time scales. Thus, it has been shown that for synthetic signals for which the theoretical MSE values are known, the estimated MSE values (numerical solutions) may significantly differ from the analytic solutions (see, e.g., [10]). This is particularly annoying for practical applications where it is difficult to obtain long recordings (biomedical field for example).
- It has been reported that Equation (1) is similar to the use of a finite-impulse response (FIR) filter [11]

$$y_j^{(\tau)} = \frac{1}{\tau} \sum_{k=0}^{\tau-1} x(j - k), \qquad 1 \le j \le N, \qquad (6)$$

on the original time series x and to the downsampling of the filtered signal with a factor τ. This FIR filter is a low-pass filter. The authors of [11] underline that the features of the frequency

response of this low-pass filter are poor, among others because it shows side lobes in the stopband. For the latter reason, the low-pass filter cannot prevent aliasing when the downsampling procedure is applied [11]. As a result, the filter does not eliminate fast temporal scale above the filter's cutoff frequency. The downsampling procedure that follows produces aliasing generating spurious oscillations in the frequency between 0 and the filter's cutoff frequency [11]. The authors of [11] therefore conclude that the evaluation of the complexity of the downsampled time series is biased by these artifacts.

- As mentioned previously, 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 proposed by Costa *et al.*, the value of r is constant for all scale factors. However, some authors considered this point as a drawback [11,13]. Indeed, as mentioned above, Equation (1) can be seen as a low-pass filtering followed by a downsampling. As a result, when the scale factor increases, the standard deviation of the resulting filtered time series may become lower and lower. Therefore, the patterns may become closer and closer. If the parameter r is constant while the scale factor τ increases, more and more patterns will be considered indistinguishable. This will lead to a decrease of the entropy when the scale factor τ increases. Valencia *et al.* concluded that MSE measures not only the variations of entropy as a function of the scale factor τ but also the variations in the power of the signal [11,14]. On this question, Costa *et al.* argue that subsequent changes of the variance due to the coarse-graining procedure should be accounted for by the entropy measure [4].

From the three points mentioned above and others (see below), several authors proposed modified MSE algorithms in order to overcome the possible drawbacks of the original MSE. These algorithms are described below.

MSE is a method that can be used with different types of entropic measures. Thus, some works did not use sample entropy but permutation entropy (see, e.g., [8,15]), cross-approximate entropy (see, e.g., [16,17]), compression entropy (see, e.g., [18]), ... These works will not be described herein: we will focus only on MSE algorithms with a sample entropy-based approach. However, some of the algorithms described below could be adapted to use not sample entropy but other entropic measures. Moreover, in what follows, the computation of the sample entropy for a time series x with the parameters m and r will be noted as $S_E(x, m, r)$.

3. Refined Multiscale Entropy

In 2009, Valencia *et al.* proposed the refined MSE (RMSE) [11]. In the latter algorithm, the authors proposed a way to remove the fast temporal scales and used a coarse-graining that prevents the influence of the reduced variance on the complexity evaluation [11]. In order to improve the elimination of the fast temporal scales, the FIR filter is replaced by a low-pass Butterworth filter. The squared magnitude of the filter frequency response is chosen as [11]

$$|H(e^{2\pi jf})|^2 = \frac{1}{1 + (f/f_c)^{2n}},\tag{7}$$

where n is the filter order and f_c is the cutoff frequency [11]. This filter has the advantage of presenting a flat magnitude for its frequency response in the passband. Moreover, it has no side lobes in the stopband and the roll-off is fast. It thus reduces aliasing when the filtered time series is downsampled [11].

Regarding the coarse-graining procedure, Valencia *et al.* proposed to reduce the dependence of the estimated entropy rate on the decrease of variance by updating the value of r as a percentage of the standard deviation of the filtered signal [11].

In order to analyze the performance of RMSE, Valencia *et al.* processed (among others) a fully unpredictable process (Gaussian white noise) and a signal with long-range correlation ($1/f$ noise). They reported that RMSE is flat with scale factor τ for the Gaussian white noise and presents a slow but progressive increase with scale factor τ in the case of $1/f$ noise. By opposition, the original MSE presents a monotonic decrease with scale factor τ both for the Gaussian white noise and the $1/f$ noise [11]. Since aliasing is more significant at short time scales when the fast oscillations are dominant, the largest differences between RMSE and the original MSE algorithms are found in the presence of high frequency oscillations [11]. The RMSE algorithm has also been applied on experimental data (see, e.g., [11,14]).

4. Composite Multiscale Entropy

In 2013, Wu *et al.* proposed the composite multiscale entropy (CMSE) in order to reduce the variance of estimated entropy values at large scales [10]. For this purpose, and for a discrete time series $x = \{x_1, \ldots, x_i, \ldots, x_N\}$, the kth coarse-grained time series for a scale factor τ is defined as [10] $y_k^{(\tau)} = \{y_{k,1}^{(\tau)} \ y_{k,2}^{(\tau)} \ldots y_{k,p}^{(\tau)}\}$ where

$$y_{k,j}^{(\tau)} = \frac{1}{\tau} \sum_{i=(j-1)\tau+k}^{j\tau+k-1} x_i, \qquad 1 \le j \le \frac{N}{\tau}, 1 \le k \le \tau. \tag{8}$$

Then, at a scale factor τ the sample entropy values of all coarse-grained time series are computed and the CMSE is defined as the means of the τ entropy values [10]

$$CMSE(x, \tau, m, r) = \frac{1}{\tau} \sum_{k=1}^{\tau} S_E(y_k^{(\tau)}, m, r), \tag{9}$$

where $S_E(y_k^{(\tau)}, m, r)$ corresponds to the sample entropy for the time series $y_k^{(\tau)}$, which can also be written as [19]

$$CMSE(x, \tau, m, r) = \frac{1}{\tau} \sum_{k=1}^{\tau} \left(-\ln \frac{n_{k,\tau}^{m+1}}{n_{k,\tau}^m} \right), \tag{10}$$

where $n_{k,\tau}^m$ represents the total number of m-dimensional matched vector pairs and is computed from the kth coarse-grained time series at a scale factor τ.

In order to compare the original MSE and CMSE, Wu *et al.* applied the two algorithms on white and $1/f$ noises [10]. Moreover, different lengths for the data have been tested. For the white noise, they reported that the estimated MSE values (numerical solutions) are significantly different from the analytic solution, for short signals. For the $1/f$ noise, they reported differences between the numerical solution and the analytic solution, for all scales [10]. The authors also mentioned that the variance of the entropy

estimator is improved by CMSE. However, the over estimation due to the short length of the time series still exists with CMSE on $1/f$ noises [10]. The authors report that CMSE shows better performances on short time series than the original MSE algorithm (the original MSE and CMSE give entropy values that are almost the same but CMSE can estimate entropy values more accurately than the original MSE: the standard deviation values of sample entropy are smaller with CMSE). They also mention that for white noise and $1/f$ noise, CMSE gives a more reliable estimation of entropy than the original MSE algorithm, as shown by the simulation results. CMSE has also been applied on real data (see, e.g., [9,10]).

5. Refined Composite Multiscale Entropy

In 2014, Wu *et al.* proposed the refined CMSE (RCMSE) to improve CMSE because CMSE does not resolve undefined entropy: CMSE estimates entropy more accurately but increases the probability of inducing undefined entropy [19]. From Equation (10), one can observe that in CMSE the logarithms of $\frac{n_{k,\tau}^{m+1}}{n_{k,\tau}^m}$ for all τ coarse-grained time series are first computed and only then the average of the results is computed to give rise to the entropy value. As a result, CMSE leads to an undefined value when either $n_{k,\tau}^{m+1}$ or $n_{k,\tau}^m$ is zero. Consequently, when short time series are analyzed, CMSE leads to much more undefined values than does the original MSE algorithm. To address this drawback, the RCMSE has been proposed, based on the following three steps [19]

(1) the same coarse-graining procedure as in CMSE is used (see Equation (8))
(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
(3) RCMSE is then defined as [19]

$$RCMSE(x, \tau, m, r) = -\ln\left(\frac{\sum_{k=1}^{\tau} n_{k,\tau}^{m+1}}{\sum_{k=1}^{\tau} n_{k,\tau}^m}\right). \tag{11}$$

From Equation (11), one can note that RCMSE leads to undefined entropy values only when all $n_{k,\tau}^{m+1}$ or $n_{k,\tau}^m$ are zeros. The RCMSE algorithm therefore reduces the probability of having undefined entropy values, compared to CMSE [19].

Wu *et al.* also processed white and $1/f$ noises with the RCMSE algorithm [19]. They concluded that, regarding validity, the RCMSE algorithm is superior to both MSE and CMSE algorithms (the probability of obtaining undefined entropy was zero). Moreover, RCMSE leads to more consistent entropy values than CMSE or MSE (the standard deviations of the entropy values obtained with RCMSE were lower than those obtained with CMSE or MSE). Finally, Wu *et al.* reported that RCMSE is better than the CMSE and MSE algorithms regarding independence of the data length [19]. Nevertheless, the computational cost of RCMSE is worse than the one of MSE (but better than the one of CMSE) [19]. RCMSE has recently been applied (in conjunction with empirical mode decomposition) on real data [20].

6. Modified Multiscale Entropy for Short-Term Time Series

In 2013, Wu *et al.* proposed the modified MSE (MMSE) in order to overcome the imprecise estimation of entropy and undefined entropy values obtained with short-term time series (short-term series generated

by the coarse-graining procedure) [12]. In the MMSE algorithm, the coarse-graining procedure is replaced by a moving-average procedure [12]

$$y_j^{(\tau)} = \frac{1}{\tau} \sum_{i=j}^{j+\tau-1} x_i, \qquad 1 \le j \le N - \tau + 1, \tag{12}$$

where $x = \{x_1, \ldots, x_i, \ldots, x_N\}$ is the discrete time series under study. Moreover, the regularities of the moving-averaged time series at a scale factor τ are quantified by computing the sample entropy value with a time delay τ [12]

$$MMSE(x, \tau, m, r) = S_E(y^{(\tau)}, m, r). \tag{13}$$

Therefore, the number of template vectors used in the MMSE algorithm is larger than the one used in the MSE algorithm. This can avoid obtaining an undefined entropy value [12]. However, the computational cost of MMSE is larger than the one of MSE [12].

Applying the MMSE algorithm on white and $1/f$ noises, Wu *et al.* reported that MMSE can detect more effectively the behaviors of white and $1/f$ noises than MSE and provides a more accurate estimation (lower standard deviation values obtained with MMSE than with MSE) [12]. The authors also reported that MMSE can avoid obtaining undefined entropy and is able to provide a more precise estimation of entropy than MSE when analyzing a short-term time series [12]. MMSE has also been applied on experimental data [12].

7. Short Time Multiscale Entropy

In 2014, Chang *et al.* reported another algorithm, the short time MSE (sMSE), in order to be able to process data recorded during a short time [21]. Thus, for a discrete signal $x = \{x_1, \ldots, x_i, \ldots, x_N\}$, the sMSE algorithm has the following steps [21]

(1) construction of the coarse-grained time series $y^{(p)(\tau)}$ with $0 \le p \le \tau - 1$ as

$$y_j^{(p)(\tau)} = \frac{1}{\tau} \sum_{i=(j-1)\tau+1+p}^{j\tau+p} x_i, \qquad 1 \le j \le (N-p)/\tau. \tag{14}$$

(2) the τ $y^{(p)(\tau)}$ time series are subjected to sample entropy computation and averaged, giving sMSE of scale factor τ

$$sMSE(x, \tau, m, r) = \frac{1}{\tau} \sum_{p=0}^{\tau-1} S_E(y^{(p)(\tau)}, m, r), \tag{15}$$

where $S_E(y^{(p)(\tau)}, m, r)$ corresponds to the sample entropy for the time series $y^{(p)(\tau)}$.

The authors applied the sMSE algorithm on white and $1/f$ noises [21]. They noticed that, for white noise, MSE exhibits fluctuations whereas sMSE shows a relatively steady decrease [21]. Moreover, sMSE has also been applied on experimental data [21,22]. Let us note that sMSE and CMSE present the same steps.

8. Instrinsic Mode Entropy

Amoud *et al.* reported that MSE does not reflect the presence of high frequencies in the signal [23]. Moreover, they also reported that MSE is not adapted to nonstationary/nonlinear signals [23]. In order to overcome these drawbacks, they proposed the intrinsic mode entropy (IMEn) that consists in computing the intrinsic mode functions (IMFs) of the signal under study (the IMFs are extracted with the empirical mode decomposition, EMD) and then by computing the sample entropy of the cumulative sums of each of the IMFs [23]. The authors applied their algorithm on stabilogram signals that are a measure of the displacement of the center of pressure of subjects standing on a force plate. They showed that IMEn is able to discriminate between elderly and control subjects [23].

9. Hierarchical Entropy

As other authors, Jiang *et al.* mentioned that MSE focuses on the lower frequency components of different scales, and that it does not take into account the higher frequencies [24]. Jiang *et al.* therefore proposed the hierarchical entropy (HE) in order to quantify the complexity of a physiologic system [24]. The HE algorithm relies on two main steps: (I) the construction of a hierarchical tree; (II) the computation of the sample entropy of each component (node) of this tree [24]. From the authors, the hierarchical decomposition of a signal can be considered as a generalization of both the multiscale decomposition (present in the original MSE algorithm) and the Haar wavelet decomposition [24]. It allows to obtain components of higher frequency in different scales, in addition to lower frequency components. HE was used to analyze heartbeat data and it has been shown that HE is able to extract significant features which may be of help for distinguishing categories of subjects [24].

10. Adaptive Multiscale Entropy

More recently, Hu *et al.* reported that the coarse-graining procedure used in the MSE algorithm essentially represents a linear smoothing and decimation of the original time series [25]. Therefore, only coarse-scale (corresponding to low-frequency components) are captured. They also reported that, because of the linear operations, the way of extracting the different scales is not convenient for nonlinear/nonstationary signals [25]. In order to overcome these drawbacks, Hu *et al.* proposed the adaptive multiscale entropy (AME). AME aims at estimating the entropy of the data over multiple adaptive scales that are intrinsically determined by multivariate EMD. The AME algorithm relies on two steps [25]: (I) a multivariate EMD to decompose the time series into the aligned IMFs at different scales; (II) a sample entropy computation over the selective scales. In order to select the scales, two algorithms have been proposed: the fine-to-coarse AME and the coarse-to-fine AME [25]. From simulations, Hu *et al.* reported that AME is able to adaptively extract the scales inherent in the nonstationary signal and that AME considers both coarse scales and fine scales in the data [25]. AME was also used for biomedical data [25].

11. Multiscale Fuzzy Sample Entropy

Xie *et al.* proposed a modification in the sample entropy algorithm [26]. In order to measure the match degree of two vectors, they suggested to use the nonlinear Sigmoid function instead of the Heaviside function (the Heaviside function presents a discontinuous and hard boundary, whereas the Sigmoid function is smooth and continuous). Moreover, in the modified sample entropy algorithm, the vector sequences are generalized by removing the local baseline. Xie *et al.* reported that the modified sample entropy leads to better performances in term of relative consistency, freedom of parameter selection, robustness to noise and independence on the data length when characterizing time series with different regularities [26]. This new algorithm has been used in different kinds of studies (see, e.g., [27]). Moreover, Kong and Xie applied this modified sample entropy algorithm to classify ventricular tachycardia and fibrillation and reported good results [28]. In 2010, Xie *et al.* proposed to use the modified sample entropy (using a Gaussian function instead of the nonlinear Sigmoid function) in order to determine whether a time series arises from a stochastic or a deterministic process [29]. They reported that the modified sample entropy is robust for detecting determinism in short and noisy signals [29].

In 2009, Chen *et al.* also proposed the use of the concept of fuzzy sets [30]. They thus developed the FuzzyEn measure, which relies on fuzzy membership functions instead of the Heaviside function. This leads to make a fuzzy boundary [30]. Chen *et al.* reported that FuzzyEn corresponds to a more accurate entropy definition than sample entropy; it owns stronger relative consistency and less dependence on data length [30]. Zhang *et al.* used the same approach in a multiscale framework for fault feature extraction: they used a Gaussian type fuzzy membership function in the similarity definition of vectors [31]. Xiong *et al.* went further in the work of Chen *et al.* [30] and proposed a new fuzzy membership function leading to a new fuzzy sample entropy and its multiscale version [32]. They reported that, when used to extract discriminating features of bearing vibration signals, their multiscale fuzzy sample entropy outperforms other multiscale entropy methods, especially when the signals contain heavy noise [32].

12. Multivariate Multiscale Entropy

In order to be able to process multivariate data, multivariate MSE has been proposed [33,34]. For this purpose, a multivariate sample entropy version has been introduced for considering both within and cross-channel dependencies in multichannel data. Multivariate MSE relies on the same steps as MSE [33,34]: (I) a coarse-graining procedure; (II) a sample entropy computation for each coarse-grained time series. However, due to the multivariate nature of the data processed by multivariate MSE, these two steps are adapted to multivariate signals. Thus, for the coarse-graining procedure, temporal scales are defined by averaging a p-variate time series $\{x_{l,i}\}_{i=1}^{N}$ ($l = 1, \ldots, p$ is the channel index and N is the number of samples in every channel) over non overlapping time segments of increasing length. Thus, for a scale factor τ, a coarse-grained multivariate time series is computed as $y_{l,j}^{(\tau)} = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} x_{l,i}$ where $1 \leq j \leq N/\tau$, and the channel index l goes from 1 to p. For the entropy computation, the multivariate sample entropy is used for each coarse-grained multivariate. The multivariate sample entropy algorithm is an extension of the (univariate) sample entropy [2]. For a tolerance level r, multivariate sample entropy 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. The detailed multivariate sample entropy algorithm can be found in [33,34]. In the multivariate MSE algorithm, multivariate sample entropy is evaluated over multiple time scales.

13. Generalized Multiscale Entropy

In 2015, Costa *et al.* introduced the generalized MSE (MSE$_n$ where n corresponds to the moment used to coarse-grain a time series) [35]. The associated algorithm is the following [35]

(1) the signal is divided into non-overlapping segments of length τ
(2) for the data in each of the segments defined in (1), a moment is estimated in order to derive the coarse-grained time series at scale τ
(3) sample entropy is calculated for each coarse-grained time series

The authors mention that, when the second moment is used, MSE$_{\sigma^2}$ quantifies the dynamics of the volatility (variance) of a signal over multiple time scales [35]. MSE$_{\sigma^2}$ has recently been used to study human heartbeat time series [35].

14. Computational Efficiency

The computational time of the original MSE algorithm prevents online applications or even the processing of long data sets (the time complexity of the algorithm is $O(N^2)$). This is due to the computational time of the sample entropy value. In order to overcome this drawback, Sugisaki *et al.* developed, in 2007, a recursive sample entropy algorithm for the situation when rolling windows overlapped, in order to be able to use the complexity measure (sample entropy algorithm) online [36]. Their algorithm uses a variable forgetting factor mechanism. However, the algorithm proposed by Sugisaki *et al.* has two drawbacks [37]: the computed sample entropy is an approximation, and the computational efficiency decays with a decrease in overlap length (the improvement does not exist when overlap length is zero).

Samani and Madeleine proposed another algorithm to compute sample entropy for online applications [38]. Their algorithm uses permuted limited number of samples to estimate sample entropy.

Later, Pan *et al.* went further [37]. They showed that the probability function in entropy could be transformed into an orthogonal range search problem in the field of computational geometry [37]. The algorithm that they proposed for the computation of MSE is based on a k-d tree algorithm. The authors reported that their algorithm reduces the computational time and mentioned that this new algorithm can be used for online applications [37].

In 2011, Jiang *et al.* also proposed an algorithm to compute sample entropy with less computational time. As for Pan *et al.*, their algorithm relies on a k-d tree data structure to accelerate counting the number of matched pairs of the pattern templates [39]. The authors reported that the time complexity of their algorithm is $O(N^{2-\frac{1}{m+1}})$ and its space complexity is $O(N \log_2 N)$, where N is the time series length and m is the length of the pattern templates.

15. Other Improvements

Other improvements than those presented above have been proposed for the original MSE algorithm. Thus, in 2011, Chou proposed the use of MSE on residuals (at various resolution levels) given by the *à trous* redundant wavelet transform [40]. Based on this, the authors proposed to determine the number of resolution levels in the wavelet transformation, according to the Mann-Kendall rank correlation test [40].

Regarding the selection of the parameter r value, and as mentioned in Section 2, Nikulin and Brismar reported that the coarse-graining procedure is similar to smoothing and decimation of the original sequences [41]. Due to the constant value of r with scales, in the original MSE algorithm, they concluded that the changes in MSE on each scale depend on both the regularity and variation of the coarse-grained time series [41]. To this point, Costa *et al.* argued that the degree of irregularity of a complex signal is a property measured by entropy that cannot be entirely captured by the standard deviation or correlation measures, individually or in combination [42]: they argue that, following the normalization, modifications in the variance generated by the coarse-graining procedure are due to the temporal structure of the original signal. Therefore, it should be accounted for by the entropy measure [4].

Regarding the application of MSE on noisy data: it has been reported that MSE is not reliable when applied on heartbeat signals containing outliers [4,43]. Therefore, in order to quantify the complexity of noisy data (data with outliers), Lo *et al.* proposed to analyze the irregularity of the sign time series of the coarse-grained data and, for this purpose, proposed the multiscale symbolic entropy (MSSE) analysis [44]. MSSE uses the sign time series of each coarse-grained series at a given time scale (1 when the time series increases and 0 otherwise). The sign time series is divided into bit sequences of length m. Based on the same concept as the approximate/sample entropy, these bit sequences are divided into different categories based on their temporal patterns (see details in [44]). Then, based on the distribution of the sequences in different categories, the Shannon entropy (which gives information richness) and the mean rank (which gives the degree of uncertainty of the fluctuations) are computed for the sign time series. Lo *et al.* used the MSSE on human heartbeat recordings. While the outliers due to ectopic beats were not removed, the authors were able to stress a modification in the cardiac control of patients with congestive heart failure, and even more [44]. As a result, they mentioned that MSSE enables the application of the complexity theory in clinical practice [44].

We can also report that Govindan *et al.* proposed a modification in the definition of sample entropy for a better characterization of the complexity: in the original sample entropy algorithm, the time series under study is divided into blocks of size m with a time gap of one (sample unit) between the successive components of the block [45]. Govindan *et al.* incorporated a time delay between the successive components of the block [45]. The delay is larger than one in the presence of dynamical correlations. Govindan *et al.* reported that this modification leads to a better characterization of the complexity than the original algorithm of sample entropy [45].

16. Conclusions

As reported above, several algorithms have been proposed to increase the performances of the original MSE. However, other improvements are still possible, among others to reduce the computational time

of the sample entropy values. Moreover, the generalized MSE offers other promising possibilities for analyzing time series.

Conflicts of Interest

The author declares no conflict of interest.

References

1. Pincus, S.M. Approximate entropy as a measure of system complexity. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 2297–2301.
2. 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.
3. Costa, M; Goldberger, A.L.; Peng, C.K. Multiscale entropy analysis of complex physiologic time series. *Phys. Rev. Lett.* **2002**, *89*, 068102.
4. Costa, M; Goldberger, A.L.; Peng, C.K. Multiscale entropy analysis of biological signals. *Phys. Rev. E* **2005**, *71*, 021906.
5. Humeau, A.; Mahé, G.; Chapeau-Blondeau, F.; Rousseau, D.; Abraham, P. Multiscale analysis of microvascular blood flow: A multiscale entropy study of laser Doppler flowmetry time series. *IEEE Trans. Biomed. Eng.* **2011**, *58*, 2970–2973.
6. Humeau, A.; Mahé, G.; Durand, S.; Abraham, P. Multiscale entropy study of medical laser speckle contrast images. *IEEE Trans. Biomed. Eng.* **2013**, *60*, 872–879.
7. Guzman-Vargas, L.; Ramirez-Rojas, A.; Angulo-Brown, F. Multiscale entropy analysis of electroseismic time series. *Nat. Hazards Earth Syst. Sci.* **2008**, *8*, 855–860.
8. Wu, S.D.; Wu, P.H.; Wu, C.W.; Ding, J.J.; Wang, C.C. Bearing fault diagnosis based on multiscale permutation entropy and support vector machine. *Entropy* **2012**, *14*, 1343–1356.
9. Niu, H.; Wang, J. Quantifying complexity of financial short-term time series by composite multiscale entropy measure. *Commun. Nonlinear Sci. Numer. Simul.* **2015**, *22*, 375–382.
10. 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.
11. Valencia, J.F.; Porta, A.; Vallverdu, 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.; Lee, K.Y.; Lin, S.G. Modified multiscale entropy for short-term time series analysis. *Phys. A* **2013**, *392*, 5865–5873.
13. Nikulin, V.V.; Brismar, T. Comment on "Multiscale entropy analysis of complex physiologic time series". *Phys. Rev. Lett.* **2004**, *92*, 089803.
14. Valencia, J.F.; Vallverdu, M.; Schroeder, R.; Cygankiewicz, I.; Vazquez, R.; Bayes de Luna, A., Porta, A.; Voss, A.; Caminal, P. Heart rate variability characterized by refined multiscale entropy applied to cardiac death in ischemic cardiomyopathy patients. *Comput. Cardiol.* **2010**, *37*, 65–68.
15. Li, Q.; Fu, Z. Permutation entropy and statistical complexity quantifier of nonstationarity effect in the vertical velocity records. *Phys. Rev. E* **2014**, *89*, 012905.

16. Wu, H.T.; Lee, C.Y.; Liu, C.C.; Liu, A.B. Multiscale cross-approximate entropy analysis as a measurement of complexity between ECG R-R interval and PPG pulse amplitude series among the normal and diabetic subjects. *Comput. Math. Methods Med.* **2013**, *2013*, 231762.

17. Wu, H.T.; Liu, C.C.; Lo, M.T.; Hsu, P.C.; Liu, A.B.; Chang, K.Y.; Tang, C.J. Multiscale cross-approximate entropy analysis as a measure of complexity among the aged and diabetic. *Comput. Math. Methods Med.* **2013**, *2013*, 324–325.

18. Humeau-Heurtier, A.; Baumert, M.; Mahé, G.; Abraham, P. Multiscale compression entropy of microvascular blood flow signals: Comparison of results from laser speckle contrast and laser Doppler flowmetry data in healthy subjects. *Entropy* **2014**, *16*, 5777–5795.

19. 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.

20. Wang, J.; Shang, P.; Xia, J.; Shi, W. EMD based refined composite multiscale entropy analysis of complex signals. *Phys. A* **2015**, *421*, 583–593.

21. 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.

22. Humeau-Heurtier, A.; Mahé, G.; Abraham, P. Modified multiscale sample entropy computation of laser speckle contrast images and comparison with the original multiscale entropy algorithm. *J. Biomed. Opt.* **2015**, in press.

23. Amoud, H.; Snoussi, H.; Hewson, D.; Doussot, M.; Duchêne, J. Intrinsic mode entropy for nonlinear discriminant analysis. *IEEE Signal Process. Lett.* **2007**, *14*, 297–300.

24. Jiang, Y.; Peng, C.K.; Xu, Y. Hierarchical entropy analysis for biological signals. *J. Comput. Appl. Math.* **2011**, *236*, 728–742.

25. Hu, M.; Liang, H. Adaptive multiscale entropy analysis of multivariate neural data. *IEEE Trans. Biomed. Eng.* **2012**, *59*, 12–15.

26. Xie, H.B.; He, W.X.; Liu, H. Measuring time series regularity using nonlinear similarity-based sample entropy. *Phys. Lett. A* **2008**, *372*, 7140–7146.

27. Mahajan, R.; Morshed, B.I. Unsupervised eye blink artifact denoising of EEG data with modified multiscale sample entropy, Kurtosis, and wavelet-ICA. *IEEE. J. Biomed. Health Inform.* **2015**, *19*, 158–165.

28. Kong, D.D.; Xie, H.B. Use of modified sample entropy measurement to classify ventricular tachycardia and fibrillation. *Measurement* **2011**, *44*, 653–662.

29. Xie, H.B.; Guo, J.Y.; Zheng, Y.P. Using the modified sample entropy to detect determinism. *Phys. Lett. A* **2010**, *374*, 3926–3931.

30. Chen, W.; Zhuang, J.; Yu, W.; Wang, Z. Measuring complexity using FuzzyEn, ApEn, and SampEn. *Med. Eng. Phys.* **2009**, *31*, 61–68.

31. Zhang, L.; Xiong, G.; Liu, H.; Zou, H.; Guo, W. Applying improved multi-scale entropy and support vector machines for bearing health condition identification. *Proc. Inst. Mech. Eng. C* **2010**, *224*, 1315–1325.

32. Xiong, G.L.; Zhang, L.; Liu, H.S.; Zou, H.J.; Guo, W.Z. A comparative study on ApEn, SampEn and their fuzzy counterparts in a multiscale framework for feature extraction. *J. Zhejiang Univ. Sci. A* **2010**, *11*, 270–279.

33. Ahmed, M.U.; Mandic, D.P. Multivariate multiscale entropy: A tool for complexity analysis of multichannel data. *Phys. Rev. E* **2011**, *84*, 061918.

34. Ahmed, M.U.; Mandic, D.P. Multivariate multiscale entropy analysis. *IEEE Signal Process. Lett.* **2012**, *19*, 91–94.

35. 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.

36. Sugisaki, K.; Ohmori, H. Online estimation of complexity using variable forgetting factor. In Proceedings of the SICE 2007 Annual Conference, Takamatsu, Japan, 17–20 September 2007; pp. 1–6.

37. 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.

38. Samani, A.; Madeleine, P. Permuted sample entropy. *Commun. Stat. Simul. Comput.* **2010**, *39*. 1506–1516.

39. Jiang, Y.; Mao, D.; Xu, Y. A fast algorithm for computing sample entropy. *Adv. Adapt. Data Anal.* **2011**, *3*, 167–186.

40. Chou, C.M. Wavelet-based multi-scale entropy analysis of complex rainfall time series. *Entropy* **2011**, *13*, 241–253.

41. Nikulin, V.V.; Brismar, T. Comment on "Multiscale entropy analysis of complex physiologic time series". *Phys. Rev. Lett.* **2004**, *92*, 089803.

42. Costa, M.; Goldberger, A.L.; Peng, C.K. Comment on "Multiscale entropy analysis of complex physiologic time series". *Phys. Rev. Lett.* **2004**, *92*, 089804.

43. Voss, A.; Schulz, S.; Schroeder, R.; Baumert, M.; Caminal, P. Methods derived from nonlinear dynamics for analysing heart rate variability. *Philos. Trans. A Math. Phys. Eng. Sci.* **2009**, *367*, 277–296.

44. Lo, M.T.; Chang, Y.C.; Lin, C.; Young, H.W.; Lin, Y.H.; Ho, Y.L.; Peng, C.K.; Hu, K. Outlier-resilient complexity analysis of heartbeat dynamics. *Sci Rep.* **2015**, *5*, 8836.

45. Govindan, R.B.; Wilson, J.D.; Eswaran, H.; Lowery, C.L.; Preiβl, H. Revisiting sample entropy analysis. *Phys. A* **2007**, *376*, 158–164.