entropy

MDPI

Article

Permutation Entropy Based on Non-Uniform Embedding

Mei Tao [1], Kristina Poskuviene [2], Nizar Faisal Alkayem [1], Maosen Cao [1], Minvydas Ragulskis [1,2,*]

[1] Department of Engineering Mechanics, Hohai University, Nanjing 210098, China; taomeihhu@163.com (M.T.); nizar.alkayem@yahoo.in (N.F.A.); cmszhy@hhu.edu.cn (M.C.)

[2] Center for Nonlinear Systems, Kaunas University of Technology, Studentu 50-146, LT-51368 Kaunas, Lithuania; kristina.poskuviene@ktu.lt

* Correspondence: minvydas.ragulskis@ktu.lt; Tel.: +370-69822456

Received: 11 July 2018; Accepted: 8 August 2018; Published: 17 August 2018

Abstract: A novel visualization scheme for permutation entropy is presented in this paper. The proposed scheme is based on non-uniform attractor embedding of the investigated time series. A single digital image of permutation entropy is produced by averaging all possible plain projections of the permutation entropy measure in the multi-dimensional delay coordinate space. Computational experiments with artificially-generated and real-world time series are used to demonstrate the advantages of the proposed visualization scheme.

Keywords: permutation entropy; ordinal pattern; attractor embedding; multi-dimensional phase space

1. Introduction

Real-world time series and experimental data are usually contaminated with noise. System states of such time series are usually complex, nonstationary and difficult to identify. Computation of the complexity for a given time series helps to quantify the intricacy of the model that governs the evolution of that series.

Different well-known complexity measures can be used to describe a time series, including Lempel-Ziv complexity [1] and Kolmogorov complexity [2]. Chaotic features of a time series can be quantified by assessing its maximal Lyapunov exponent [3]. The complexity of a time series can be assessed by dimension estimation models (such as the information dimension [4] and the fractal dimension [5]); also by entropy assessment methods (such as Shannon entropy [6], approximate entropy [7], K-S entropy [8,9], sample entropy [10], conditional entropy [11], multiscale entropy [12]).

Arguably, the most powerful tool to measure the complexity of a time series is the permutation entropy (PE). Since its introduction in 2002 by Bandt and Pompe in their foundational paper [13], it has been successfully applied in a wide range of scientific areas and for a vast number of purposes [14]. The main advantages of PE are: (i) it is simple to use; (ii) calculations are fast; and (iii) it is robust against noise. PE is based on the permutation patterns or the order relations among values of a signal [15]. PE compares the order of neighboring relative values, rather than apportioning amplitudes according to different levels [15,16]. PE can be computed using fast PE algorithms [17]. These characteristics make PE an appealing tool used in a large number of real-world signal and image processing applications [18].

PE is based on one-dimensional time series reconstruction into a D-dimensional space with the embedding delay τ. The time series embedding scheme into the higher dimension space was first presented by Packard, Crutchfield, Farmer and Shaw [19,20]. This embedding scheme represents the optimal properties of a dynamical system if the embedding dimension Dand time delay τ (the difference between consecutive observations) are estimated adequately. Theoretical foundations by

Takens [21], and expansions of his ideas by Sauer et al. [22] indicate that the embedding dimension of $D > 2m + 1$ (where m is the fractal dimension of the attractor) almost always ensures the reconstruction of the topology of the original attractor [20]. This surprising result states that time series output is sufficient to obtain complete information about hidden states of the dynamical system [23].

Commonly-used methods for finding the optimal time delay τ are the average mutual information method [24], the correlation sum method [25], the phase space expansion method [26,27] and the geometry-based method [28].

The selection of the optimal embedding dimension is usually based on the examination of some invariant on the reconstructed attractor. Usually, these invariants represent the dynamics of the system and are based on the geometrical properties of that system. An invariant value is computed by increasing the embedding dimension until that invariant value settles down [29]. Typical examples of such methods are the box-counting method [20,30], the correlation dimension method [25,31], the largest Lyapunov exponent method [32], the Kolmogorov–Sinai entropy method [33] and the false nearest neighbors (FNN) method [34,35].

The identification of the optimal embedding dimension D and the optimal time delay τ helps to reconstruct the attractor in the delay-coordinate space. However, it is well known that non-uniform embedding (when time delays are not equal) might lead to a better reconstruction of the attractor if compared to uniform embedding (when all time delays are equal) [36,37]. However, the selection of the optimal vector of time delays for non-uniform embedding is a difficult optimization problem that requires massive computational resources. Several approaches have been proposed to tackle this problem. The identification of a near-optimal vector of time delays employing genetic algorithms was proposed in [38,39]. Good near-optimal non-uniform embedding results were obtained by using the ant colony optimization algorithms reported in [40]. Non-uniform time series embedding with special target functions based on the Fourier spectral analysis were presented in [41,42].

The main objective of this paper is to employ non-uniform time series embedding for the construction of a visualization scheme for PE. The optimal embedding dimension and the set of optimal time lags are used to design a computational algorithm for plotting PE as a single surface. The paper is structured as follows. The normalized PE and the target function used to identify optimal time delays are introduced in Section 2. The proposed visualization scheme for PE is introduced in Section 3. The results of computational experiments with the sine wave, the Rossler time series and a real-world time series are discussed in Section 4. A discussion and concluding remarks are given in the last section.

2. Preliminaries

2.1. Permutation Entropy

For a given time series $\{x_1, x_2, \ldots, x_N\}$, uniform embedding yields a trajectory matrix:

$$Y_t = \{x_t, x_{t+\tau}, x_{t+2\tau}, \ldots, x_{t+(D-1)\tau}\}, \quad t = 1, 2, \ldots, N - (D-1)\tau, \tag{1}$$

where D is the embedding dimension and τ is the time delay. PE quantifies the statistics of ordinal permutations in the rows of the trajectory matrix [29].

For example, the sequence $\{3, 7, 6\}$ has ordinal pattern $\pi = 1;3;2$, since its $x_1 \leq x_3 \leq x_2$. The ordinal pattern of the sequence $\{6, 3, 7\}$ is $\pi = 2;1;3$. As a consequence, there are $D!$ possible pattern orders, which represent all unique orderings (permutations $\pi_i, i = 1, 2, \ldots, D!$). The relative frequency of each distribution with which they occur in the trajectory matrix is defined as follows:

$$P_i = \frac{f(\pi_i)}{N - (D-1)\tau}, \quad i = 1, 2, \ldots, D!, \tag{2}$$

where $f(\pi_i)$ represents the occurrence number of the π_i pattern order. Normalized PE is defined here as:

$$H_D(\tau) = -\frac{1}{\ln(D!)} \sum_{i=1}^{D!} P_i \ln(P_i). \tag{3}$$

The range of PE values defined by Equation (3) is from 0–1. PE depends on two predefined parameters: the embedding dimension D and the time lag τ.

2.2. Non-Uniform Embedding

The trajectory matrix produced by non-uniform embedding reads:

$$Y_t = \{x_t, x_{t+\tau_1}, x_{t+\tau_1+\tau_2}, \ldots, x_{t+\Delta}\}, \quad t = 1, 2, \ldots, n - \Delta, \tag{4}$$

where n is the length of the observable time series; Δ is the width of the observation window; $\Delta = \tau_1 + \ldots + \tau_{D-1}$; $\{\tau_1, \tau_2, \ldots, \tau_{D-1}\}$ is the vector of time delays.

The optimal time delays are computed as the $(D-1)$-dimensional argument of the following maximization problem [43]:

$$\max_{1 \le \tau_1, \ldots, \tau_{D-1} \le L} \left(\frac{1}{(n-\Delta)\sqrt{D}} \sum_{k=1}^{n-\Delta} \sqrt{x_k^2 + x_{k+\tau_1}^2 + \ldots + x_{k+\Delta}^2} \right). \tag{5}$$

where L is the upper range for all time lags.

3. The Proposed Visualization Scheme for PE

Visualization of PE as a function of two time delays (at $D = 3$) is proposed in [44]. Such an approach enables one to plot PE as a surface, and the graphical features of that surface leak the underlying complexity of the analyzed time series. Visualizing higher-dimensional arrays is a considerable challenge, and a scheme for rationalizing the information contained in these arrays is of considerable practical benefit, since in practice, one often wishes to use $D > 3$. The proposed visualization scheme for PE at $D > 3$ reads:

1. Determine the optimal embedding dimension D for a given time series.
2. Set L (the upper range for all time lags). Determine the set of optimal time lags $\tau_1^\star, \tau_2^\star, \ldots, \tau_{D-1}^\star$ according to Equation (5).
3. Repeat in lexicographical order and construct planar images of PE:
   (a) Freeze $D-3$ optimal time lags;
   (b) Vary two time lags from 1–L, and compute PE according to Equations (2) and (3);
   (c) Construct a two-dimensional digital image of PE.
4. Average all $(D-1)(D-2)/2$ planar digital images of PE.

Example 1. *In the case of $D = 5$, the visualization procedure can be illustrated by a schematic diagram in Table 1.*

Table 1. Schematic diagram for PE at $D = 5$.

τ_1	τ_2	τ_3	τ_4	$H_5(\tau)$
$1, \ldots, L$	$1, \ldots, L$	$\tau_3^\star$	$\tau_4^\star$	$H_5\left(\tau_1, \tau_2, \tau_3^\star, \tau_4^\star\right)$
$1, \ldots, L$	$\tau_2^\star$	$1, \ldots, L$	$\tau_4^\star$	$H_5\left(\tau_1, \tau_2^\star, \tau_3, \tau_4^\star\right)$
$1, \ldots, L$	$\tau_2^\star$	$\tau_3^\star$	$1, \ldots, L$	$H_5\left(\tau_1, \tau_2^\star, \tau_3^\star, \tau_4\right)$
$\tau_1^\star$	$1, \ldots, L$	$1, \ldots, L$	$\tau_4^\star$	$H_5\left(\tau_1^\star, \tau_2, \tau_3, \tau_4^\star\right)$
$\tau_1^\star$	$1, \ldots, L$	$\tau_3^\star$	$1, \ldots, L$	$H_5\left(\tau_1^\star, \tau_2, \tau_3^\star, \tau_4\right)$
$\tau_1^\star$	$\tau_2^\star$	$1, \ldots, L$	$1, \ldots, L$	$H_5\left(\tau_1^\star, \tau_2^\star, \tau_3, \tau_4\right)$

Note that only two different time delays (τ_k, τ_l) are not fixed in each plane projection of H_D. In other words, H_D is a two-dimensional image of PE when all time delays τ_s, $s \neq k$, $s \neq l$, $k \neq l$ are fixed to the corresponding components of the vector of optimal time delays. Finally, a two-dimensional averaged digital image of PE is computed as an arithmetic average of all planar PE images. The upper range L should be high enough to ensure that the maximum point of the target function (Equation (5)) does not lie on the boundary of the search space (all optimal time lags should be lower than L).

Note that the computation of all possible digital images of PE is not necessary. It is possible to compute the averaged two-dimensional digital image (a square matrix of $L \times L$ pixels) of PE directly:

$$\bar{H}_D(\zeta = i, \eta = j) = \frac{2}{(D-1)(D-2)} \sum_{k=1}^{D-2} \sum_{l=k+1}^{D-1} H_D(\tau_k = i, \tau_l = j), \tag{6}$$

where ζ and η are coordinates of a pixel in the x- and the y-axis; $i, j = 1, \ldots, L$.

4. Computational Experiments

4.1. The Sine Wave

The sine wave used in this computational experiment is generated by the formula $x_t = \sin(\frac{2\pi t}{40})$, $t = 1, \ldots, 10{,}000$; L is set to 100. The optimal embedding dimension is determined by FNN; $D = 4$. Full sort yields the optimal set of time delays $\tau_1 = \tau_2 = \tau_3 = 5$. Planar projections $H_4(\tau_1, \tau_2, 5)$, $H_4(\tau_1, 5, \tau_3)$, $H_4(5, \tau_2, \tau_3)$ and the averaged $\bar{H}_4(\zeta, \eta)$ are shown in Figure 1.

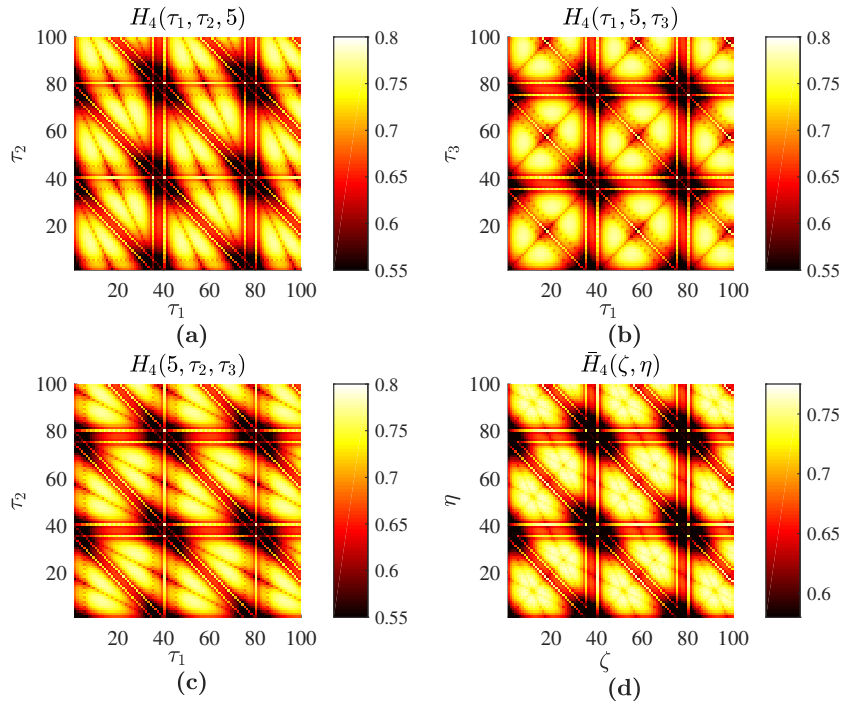

Figure 1. Permutation entropy for the sine wave: $H_4(\tau_1, \tau_2, 5)$ is depicted in (**a**); $H_4(\tau_1, 5, \tau_3)$ in (**b**); $H_4(5, \tau_2, \tau_3)$ in (**c**); and the averaged $\bar{H}_4(\zeta, \eta)$ in (**d**); Numerical values of permutation entropy are indicated in color bars.

It is well known that planar PE projections of the sine wave do exhibit a periodic pattern of periodic and symmetric cells (the size of these cells equals the ratio between the sampling frequency and the period of the sine wave) [44]. The averaged $\bar{H}_4$ based on the optimal embedding into the delay coordinate space also reveals the periodicity of the embedded time series. All three images

$H_4(\tau_1, \tau_2, 5)$, $H_4(\tau_1, 5, \tau_3)$ and $H_4(5, \tau_2, \tau_3)$ are comprised of periodic cells with the same geometric boundaries. Though the structure of the geometric pattern in every cell is similar, the distortion of the image and the angles of the characteristic inclination lines are different (Figure 1a–c). Each cell in the averaged image yields traces of the three different inclination lines (Figure 1d). Thus, in principle, it is possible to read the optimal embedding dimension from the averaged image. The three inclination lines visible in each cell of Figure 1d suggest that the averaged image is comprised of three plain PE images, which in turn suggests that the optimal embedding dimension is four (Figure 2).

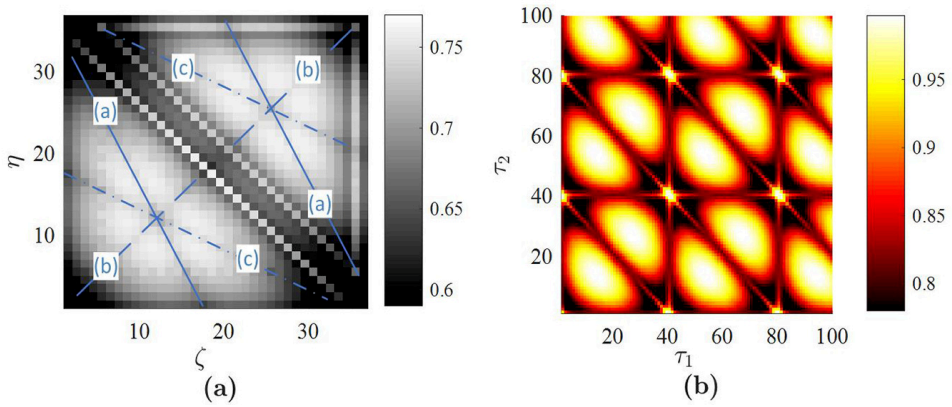

Figure 2. The geometric structure of a single periodic cell produced by the averaged PE reveals the three different inclination lines from three planar projections of the Permutation entropy (PE). (**a**) The PE reconstructed by the technique presented in [44] is depicted in (**b**).

Computational experiments are continued with the sine wave using the technique presented in [44] (Figure 2b). The structure of the reconstructed image reveals the periodicity of the sine wave. However, the complexity of the image in Figure 2b is lower than the complexity of PE images depicted in Figure 1. This can be explained by the fact that the ordinal patterns are divided into 24 bins in Figure 1a–c versus six bins in Figure 2b.

4.2. The Rössler Time Series

The Rössler system is a paradigmatic model of chaotic dynamics [45,46]:

$$\begin{cases} \frac{dx}{dt} = -y - z, \\ \frac{dy}{dz} = x + ay, \\ \frac{dz}{dt} = b + z(x - c), \end{cases} \tag{7}$$

where constants a, b, c are set to $a = 0.1, b = 0.1, c = 14$. The Rössler time series is generated by integrating the three coupled ordinary differential equations and by selecting every tenth value of x in the time domain $1 \leq t \leq 1000$ (the time step is set to 0.01). FNN yields the optimal embedding dimension $D = 6$; L is set to 100. Full sort yields the set of optimal time lags $\{38, 11, 42, 21, 23\}$. All planar projections of PE produced by fixing three of the five delays at the optimal values are illustrated in Figure 3.

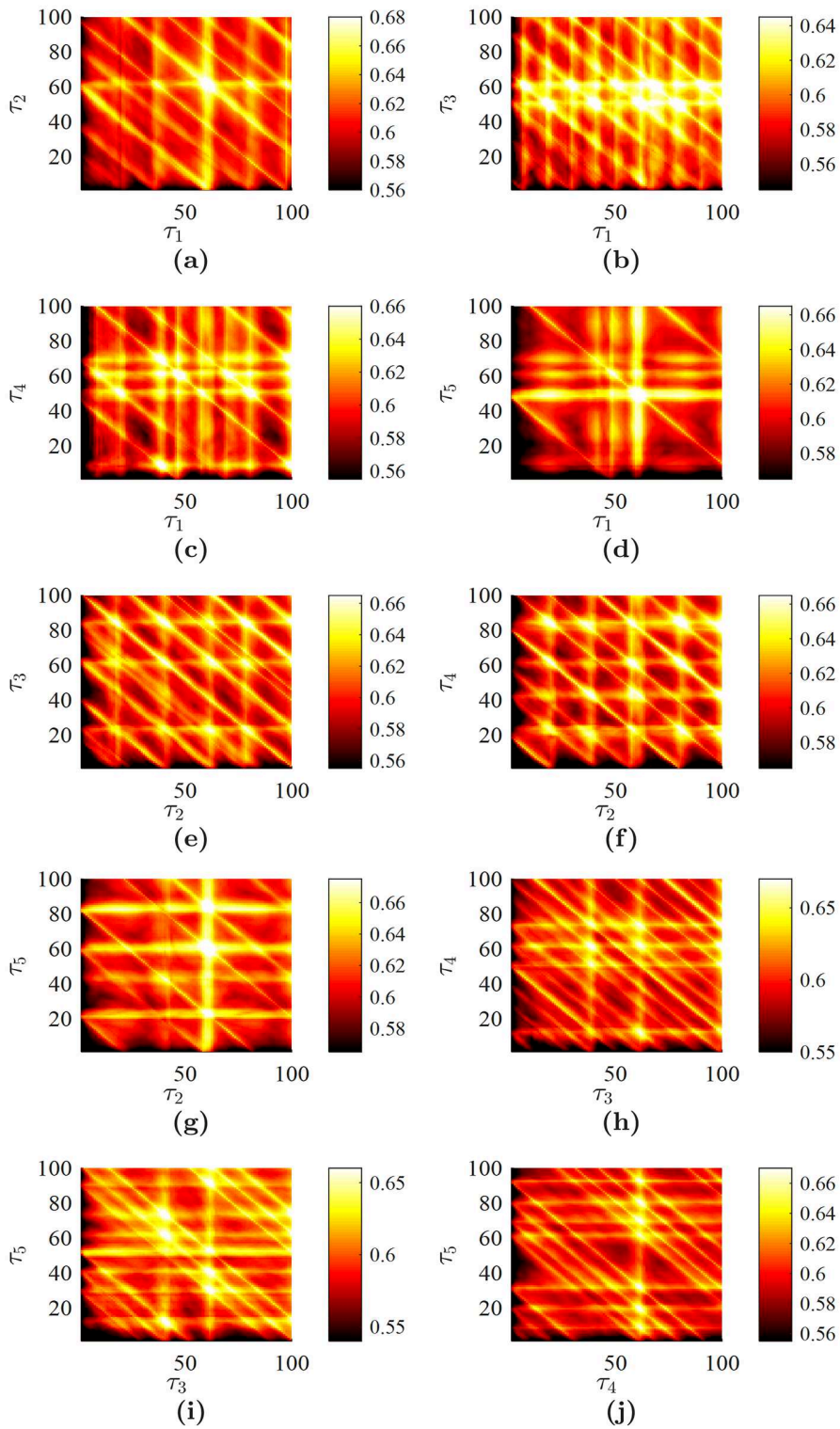

Figure 3. All possible planar projections of PE for the Rössler time series ($D = 6$): (**a**) $H_6(\tau_1, \tau_2, 42, 21, 23)$; (**b**) $H_6(\tau_1, 11, \tau_3, 21, 23)$; (**c**) $H_6(\tau_1, 11, 42, \tau_4, 23)$; (**d**) $H_6(\tau_1, 11, 42, 21, \tau_5)$; (**e**) $H_6(38, \tau_2, \tau_3, 21, 23)$; (**f**) $H_6(38, \tau_2, 42, \tau_4, 23)$; (**g**) $H_6(38, \tau_2, 42, 21, \tau_5)$; (**h**) $H_6(38, 11, \tau_3, \tau_4, 23)$; (**i**) $H_6(38, 11, \tau_3, 21, \tau_5)$; (**j**) $H_6(38, 11, 42, \tau_4, \tau_5)$.

Finally, the averaged image of PE $\bar{H}_6(\zeta, \eta)$ is depicted in Figure 4b. It can be observed that $\bar{H}_6$ integrates specific features represented in planar projections of PE and provides a single image

representing the nonlinear properties of the Rössler time series. However, it is impossible to reconstruct the optimal embedding dimension from Figure 4a. First of all, this is due to the complexity of the time series. Secondly, the number of averaged images is large due to the fact that the optimal embedding dimension is large, as well. The Rössler time series represents a chaotic oscillator. Nevertheless, it is possible to assess an average period of chaotic oscillations (which is around 60 time steps in our computational experiment). The high-PE bands at delays of 60 (Figure 4b) denote this average period of oscillations (compare to Figure 1).

The averaged image $\bar{H}_6(\zeta, \eta)$ can be compared with a similar reconstruction of PE with non-ideal parameters of the embedding. Uniform embedding yields the set of optimal time delays $\{9, 9, 9, 9, 9\}$; the optimal embedding dimension is the same ($D = 6$; Table 2). The averaged $\bar{H}_6(\zeta, \eta)$ produced by uniform embedding is depicted in Figure 4a. It is clear that the uniform embedding produces a much more regular image of averaged PE if compared to the non-uniform embedding. The high-PE bands at delays around 60 are still visible in images produced both by uniform and non-uniform embedding. However, the pattern of PE in Figure 4b is much less regular compared to the one in Figure 4a.

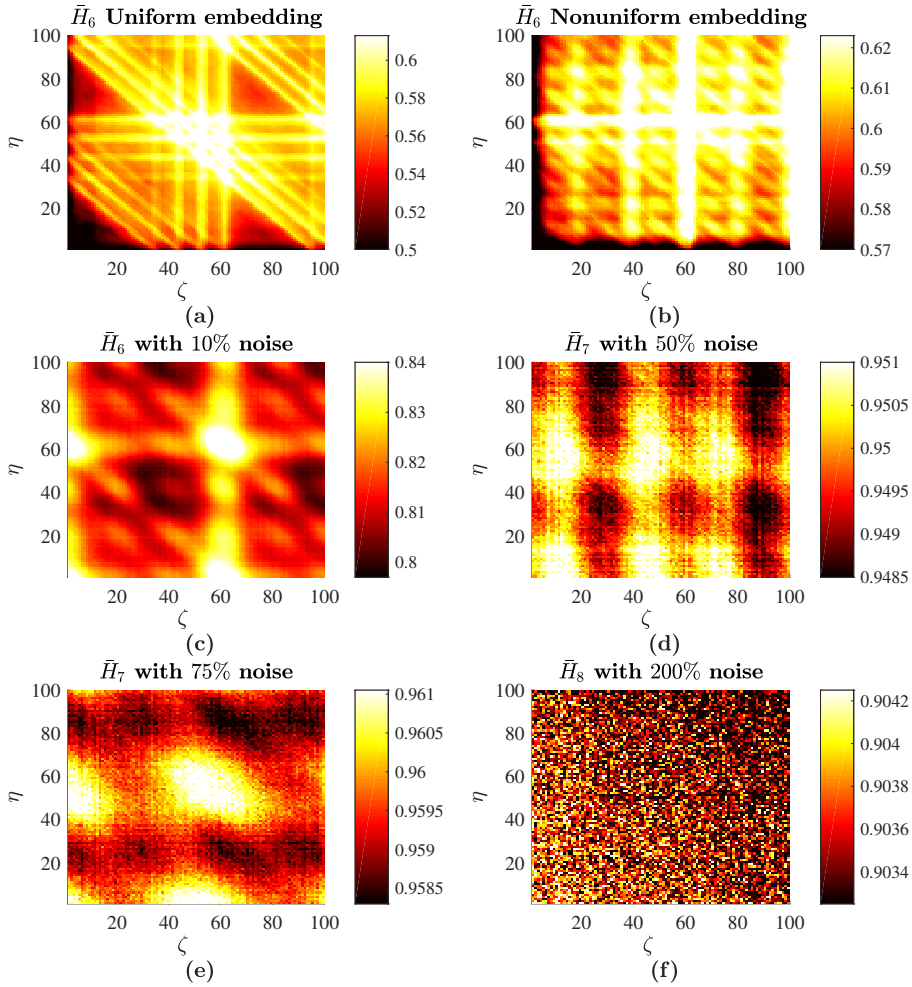

Figure 4. Averaged PE for the Rössler time series: (**a**) uniform embedding with no additive noise; (**b**) non-uniform embedding with no additive noise; (**c**) $\bar{H}_6$ with 10% noise; (**d**) $\bar{H}_7$ with 50% noise; (**e**) $\bar{H}_7$ with 75% noise; (**f**) $\bar{H}_8$ with 200% noise.

Table 2. Optimal embedding dimensions and optimal time lags for the Rössler time series with different noise levels.

Noise Level	Dimension	The Set of Optimal Time Lags	Reference to Figure
0%	6	$\{9, 9, 9, 9, 9\}$	Figure 4a
0%	6	$\{38, 11, 42, 21, 23\}$	Figure 4b
10%	6	$\{37, 45, 17, 38, 6\}$	Figure 4c
50%	7	$\{36, 16, 14, 28, 35, 36\}$	Figure 4d
75%	7	$\{33, 15, 33, 25, 25, 16\}$	Figure 4e
200%	8	$\{49, 41, 15, 36, 33, 16, 40\}$	Figure 4f

All real-world time series are contaminated by noise. Therefore, it is important to investigate the effects induced by noise on the proposed scheme of PE visualization. Computational experiments are continued with the Rössler time series, but different realizations of the Gaussian random noise with zero mean are added to it. The standard deviation of the random noise is set to $\frac{\beta R}{100\%}$, where β is the noise level in percent; R is the range of the stationary Rössler time series (the difference between the maximum and the minimum values of the time series in the observation window). $\bar{H}_6$ for the Rössler time series with different noise levels are depicted in Figure 4 (intermediate images of all possible plain projections of PE are omitted for brevity). The deterministic components of PE are gradually lost when the signal is completely buried in noise. $\bar{H}_6$ with 200% noise (Figure 4f) represents the image of random noise. It is known that PE is robust to noise [47]. However, the amount of noise added to the deterministic time series is so high that the time series becomes a random time series and correlations between data points are lost completely. It is interesting to observe that the average PE is more acutely impacted by noise when one of the coordinates (ζ, η) is low. Note that the addition of noise to the deterministic system induces the change of the optimal embedding dimension and optimal delays in the resulting time series. Therefore, the optimal embedding dimension and the set of optimal time delays are reconstructed for every different noise level (Table 2). Though PE is not a good measure to quantify the structural complexity, it is well suited to assess the randomness of a time series (PE increases with respect to noise level) [48]. This effect is clearly observed in Figure 4 when the range of PE raises from $[0.58; 0.63]$ up to $[0.9937; 0.9940]$.

The presented algorithm for plotting the averaged PE is based not only on the optimal embedding dimension, but also on the set of optimal time delays. Non-uniform embedding results in a different pattern if compared to uniform embedding (Figure 4a,b). However, it is also important to answer a question if the process of optimizing time delays yields any benefit aside from the averaging effect. In other words, it is important to compare PE patterns generated by the optimal non-uniform set of time delays with patterns generated by a set of delays generated by a random number generator.

Computational experiments are continued with the Rössler time series without the additive noise. A standard random number generator is used to produce 10 real numbers uniformly distributed in the interval $[0, 1]$. Every number is multiplied by 50 and rounded, resulting in two sets of random time delays: $\{7, 18, 39, 27, 12\}$ and $\{30, 40, 16, 9, 29\}$. The averaged $\bar{H}_6$ for each set of random time delays are depicted in Figure 5.

It can be seen that particular values of time delays do have a strong impact on the averaged pattern of PE. The averaged PE inherits the properties of individual planar sections of PE. The geometric coordinates of the particular set of time delays in the multi-dimensional coordinate space define specific features of planar projections, which influence the averaged image. The high-PE bands at delays of 60 are still clearly expressed in Figure 5a,b. However, the averaged PE patterns are very much different; and the proposed technique for plotting the averaged PE is particularly based on the set of optimal time delays.

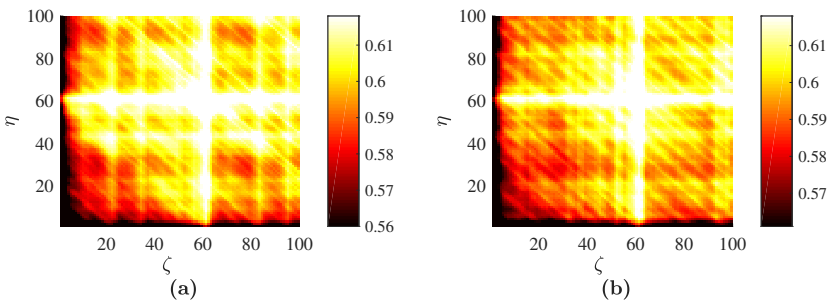

Figure 5. Averaged PE for the Rössler time series with no additive noise: (**a**) the pattern produced by a random set of time delays $\{7, 18, 39, 27, 12\}$; (**b**) by a random set of time delays $\{30, 40, 16, 9, 29\}$.

4.3. Real-World Time Series

Computational experiments are continued with the EEG signal available from the Brain/Neural Computer Interaction (BNCI) Horizon 2020 project [49] (A01 time series from dataset P300 Speller); the graphical representation of the signal is shown in Figure 6.

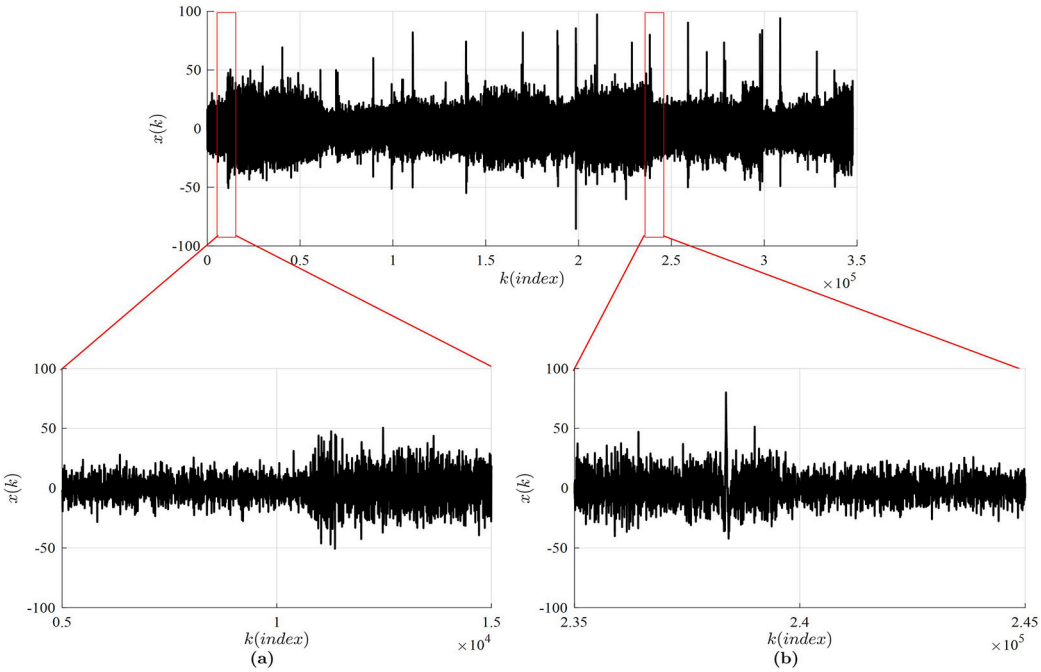

Figure 6. The Electroencephalogram (EEG) signal available from the Brain/Neural Computer Interaction (BNCI) Horizon 2020 project database [49]. Insets (**a**) and (**b**) are used to depict the zoomed parts of the signal.

Two different intervals are selected from the same Electroencephalogram (EEG) signal: $1 \leq k \leq 100,000$ (Interval A) and $200,001 \leq k \leq 300,000$ (Interval B). Every second measurement point is skipped in both intervals in order to make the optimal time delays smaller (the number of remaining points in Intervals A and B is 50,000).

FNN yields the optimal embedding dimension $D = 5$ for Interval A. The set of optimal time lags is $\{\tau_1 = 6, \tau_2 = 9, \tau_3 = 16, \tau_4 = 6\}$; the parameter L is set to 60. The averaged PE for Interval A $\bar{H}_5$ is depicted in Figure 7a.

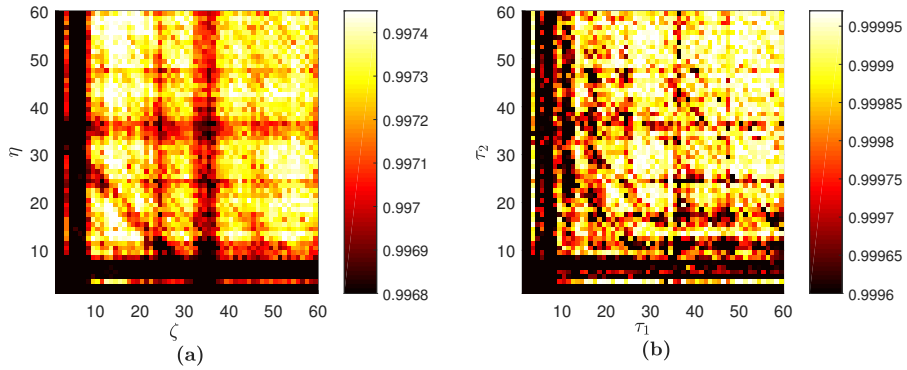

Figure 7. Digital images of PE reconstructed for Interval A. The proposed scheme yields the image in (**a**). The scheme without the assessment of the optimal embedding dimension and the optimal set of time lags results in the image in (**b**).

Computational experiments are continued with Interval A, but now, PE is visualized using the technique proposed in [44]. The embedding dimension D is always three; time delays τ_1 and τ_2 are varied from 1–L (Figure 7b). Some important observations can be done by comparing the two digital images of PE in Figure 7. First of all, the process of averaging makes some features of the PE pattern appear cleaner. However, the averaging results in some information loss; for example, the stripe at $\tau_2 = 5$ in Figure 7b is not visible in Figure 7a. Clearly, some of the structure of the PE pattern in Figure 7a is replicated in Figure 7b, albeit with more noise. Furthermore, the difference in the PE values is a consequence of different D; and the fact that ordinal patterns are being divided into 120 bins in Figure 7a versus six bins in Figure 7b. The straightforward computation of PE (as shown in Figure 7b) yields a high level of noise what results in a high uncertainty of PE reconstruction. The improvement in the signal-to-noise of our PE reconstruction is due to the fact that the proposed algorithm averages over many two-dimensional images.

Analogous computations are performed with Interval B. FNN yields the optimal embedding dimension $D = 5$ again, but the set of optimal time lags is now $\{\tau_1 = 5, \tau_2 = 6, \tau_3 = 14, \tau_4 = 6\}$. The averaged PE for Interval B $\bar{H}_5$ is shown in Figure 8a. The old visualization scheme for Interval B yields the digital image in Figure 8b. This confirms again that the optimal embedding dimension and the optimal set of time delays play a pivotal role in a representative visualization of PE. Comparison of the graphical features in digital images of PE represented in Figures 7a and 8a helps to identify the differences occurring in the evolving time series. Such comparisons are much more difficult with the old visualization scheme (Figures 7b and 8b).

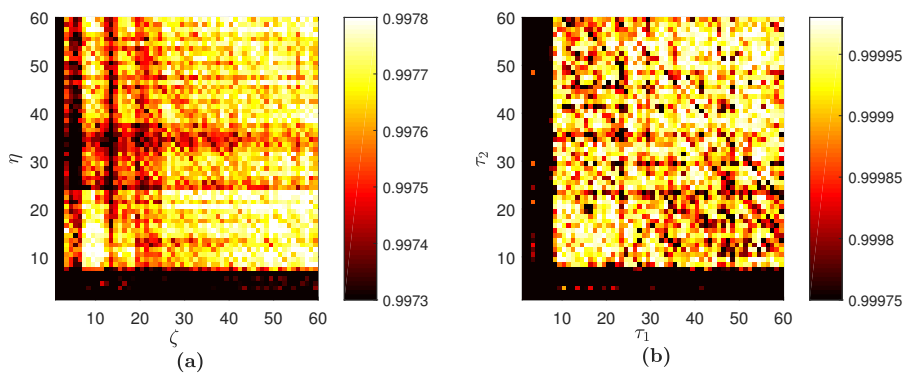

Figure 8. Digital images of PE reconstructed for Interval B. The proposed scheme yields the image in (**a**). The scheme without the assessment of the optimal embedding dimension and the optimal set of time lags results in the image in (**b**).

5. Discussion

A novel visualization scheme for permutation entropy is presented in this paper. This scheme matches permutation entropy with the topological characteristics of the investigated time series: the optimal embedding dimension and the optimal set of time delays. The proposed scheme is based on non-uniform attractor embedding and uses different time delays, but results in a single digital image.

The proposed algorithm is based on the averaging of all possible plain projections of the permutation entropy measure in the multi-dimensional delay coordinate space. Such an approach is a natural extension of the technique used for the quantification of the phase space occupied by the reconstructed attractor [41]. The proposed scheme extends the visualization of permutation entropy from a three-dimensional phase space (with two time delays) to a multi-dimensional phase space.

The proposed scheme is well suited for real-world time series contaminated by the additive noise. Arithmetic averaging of plane projections reduces the optical effects induced by the additive noise and increases the clarity of specific geometric features (which can be used for the interpretation of the investigated time series). It is well known that permutation entropy can be used to identify couplings between time series [50]. The applicability of the proposed visualization scheme for the identification of couplings and synchronization between time series remains a definite objective of future research.

Author Contributions: Conceptualization, M.R.; Methodology, M.R. and K.P.; Software, M.T., K.P. and N.F.A. Validation, M.C. and N.F.A.; Formal analysis, M.T. and K.P.; Investigation, K.P.; Resources, M.T. and M.C.; Data curation, K.P. and N.F.A.; Writing, original draft preparation, M.R.; Writing, review and editing, M.C. and M.R.; Visualization, K.P.; Supervision, M.R.; Project administration, M.T. and M.R.; Funding acquisition, M.C.

Funding: This research was funded by the Jiangsu Provincial Recruitment Program of Foreign Experts (Type B, Grant No. JSB2017007).

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

PE	Permutation Entropy
FNN	False Nearest Neighbors
EEG	Electroencephalogram
BNCI	Brain/Neural Computer Interaction

References

1. Lempel, A.; Ziv, J. On the complexity of finite sequences. *IEEE Trans. Inform. Theory* **1976**, *22*, 75–81. [CrossRef]
2. Kolmogorov, A.N. Three approaches to the definition of the concept "quantity of information". *Probl. Peredachi Inf.* **1965**, *1*, 3–11.
3. Kaspar, F.; Schuster, H.G. Easily calculable measure for the complexity of spatiotemporal patterns. *Phys. Rev. A* **1987**, *36*, 842–848. [CrossRef]
4. Farmer, D.J. Information dimension and the probabilistic structure of chaos. *J. Phys. Sci.* **1982**, *37*, 1304–1325. [CrossRef]
5. Termonia, Y.; Alexandrowicz, Z. Fractal dimension of strange attractors from radius versus size of arbitrary clusters. *Phys. Rev. Lett.* **1983**, *51*, 1265–1268. [CrossRef]
6. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [CrossRef]
7. Pincus, S.M. Approximate entropy as a measure of system complexity. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 2297–2301. [CrossRef] [PubMed]
8. Walters, P. *An Introduction to Ergodic Theory*; Springer Publishing House: New York, NY, USA, 1982; p. 250, ISBN 978-0-387-95152-2.
9. Stolz, I.; Keller, K. A general symbolic approach to Kolmogorov-Sinai entropy. *Entropy* **2017**, *19*, 675. [CrossRef]

10. Richman, J.S.; Moorman, J.R. Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol. Heart Circ. Physiol.* **2000**, *278*, 2039–2049. [CrossRef] [PubMed]
11. Keller, K.; Unakafov, A.M.; Unakafova, V.A. Ordinal patterns. *Entropy* **2014**, *16*, 6212–6239. [CrossRef]
12. Zhou, J.; Xiao, J.; Xiao, H.; Zhang, W.; Zhu, W.; Li, C. Multifault diagnosis for rolling element bearings based on intrinsic mode permutation entropy and ensemble optimal extreme learning machine. *Adv. Mech. Eng.* **2014**, *6*, 803–919. [CrossRef]
13. Bandt, C.; Pompe, B. Permutation entropy: A natural complexity measure for time series. *Phys. Rev. Lett.* **2002**, *88*, 174102. [CrossRef] [PubMed]
14. Zunino, L.; Olivares, F.; Scholkmann, F.; Rosso, O.A. Permutation entropy based time series analysis: Equalities in the input signal can lead to false conclusions. *Phys. Lett. A* **2017**, *381*, 1883–1892. [CrossRef]
15. Zanin, M.; Zunino, L.; Rosso, O.A.; Papo, D. Permutationentropy and its main biomedical and econophysicsapplications: A review. *Entropy* **2012**, *14*, 1553–1577. [CrossRef]
16. Li, Y.; Li, G.; Yang, Y.; Liang, X.; Xu, M. A fault diagnosis scheme for planetary gearboxes using adaptive multi-scale morphology filter and modified hierarchical permutation entropy. *Mech. Syst. Signal Proccess* **2018**, *105*, 319–337. [CrossRef]
17. Unakafova, V.A.; Keller, K. Efficiently measuring complexity on the basis of real-world data. *Entropy* **2013**, *15*, 4392–4415. [CrossRef]
18. Azami, H.; Escudero, J. Amplitude-aware permutation entropy: Illustration in spike detection and signal segmentation. *Comput. Meth. Prog. Biomed.* **2016**, *128*, 40–51. [CrossRef] [PubMed]
19. Packard, N.H.; Crutchfield, J.P.; Farmer, J.D.; Shaw, R.S. Geometry from a time series. *Phys. Rev. Lett.* **1980**, *45*, 712–716. [CrossRef]
20. Theiler, J. Estimating the fractal dimension of chaotic time series. *Lincoln Lab. J.* **1990**, *3*, 63–86. [CrossRef]
21. Takens, F. Detecting strange attractors in turbulence. In *Dynamical Systems and Turbulence, Warwick 1980*; Rand, D.A.,Young, L.S., Eds.; Springer: Berlin/Heidelberg, Germany, 1981; pp. 366–381, doi:10.1007/BFb0091924.
22. Sauer, T.; Yorke, J.A.; Casdagli, M. Embedology. *J. Stat. Phys.* **1991**, *65*, 579–616. [CrossRef]
23. Yap, H.L.; Eftekhari, A.; Wakin, M.B.; Rozell, C.J. A first analysis of the stability of takens' embedding. In Proceedings of the IEEE Global Conference on Signal and Information Processing (GlobalSIP), Anaheim, CA, USA, 3–5 December 2014; pp. 404–408. [CrossRef]
24. Fraser, A.; Swinney, H. Independent coordinates for strange attractors from mutual information. *Phys. Rev. A* **1986**, *33*, 1134–1140. [CrossRef]
25. Grassberger, P.; Procaccia, I. Measuring the strangeness of strange attractors. *Physical D* **1983**, *9*, 189–208. [CrossRef]
26. Buzug, T.; Pfister, G. Comparison of algorithms calculating optimal embedding parameters for delay time coordinates. *Physical D* **1992**, *58*, 127–137. [CrossRef]
27. Buzug, T.; Pfister, G. Optimal delay time and embedding dimension for delay-time coordinates by analysis of the global static and local dynamical behavior of strange attractors. *Phys. Rev. A* **1992**, *45*, 7073–7084. [CrossRef] [PubMed]
28. Casdagli, M.; Eubank, S.; Farmer, J.; Gibson, J. State space reconstruction in the presence of noise. *Physical D* **1991**, *51*, 52–98. [CrossRef]
29. Bradley, E.; Kantz, H. Nonlinear time-series analysis revisited. *Chaos* **2015**, *25*, 097610. [CrossRef] [PubMed]
30. Falconer, K. Box-counting dimension. In *Fractal Geometry: Mathematical Foundation and Applications*; John Wiley and Sons: Chichester, UK, 1990; pp. 27–43, ISBN 10: 0471922870.
31. Grassberger, P.; Procaccia, I. Characterization of strange attractors. *Phys. Rev. Lett.* **1983**, *50*, 346–349. [CrossRef]
32. Wolf, A.; Swift, J.B.; Swinney, H.L.; Vastano, J.A. Determining Lyapunov exponents from a time series. *Physical D* **1985**, *16*, 285–317. [CrossRef]
33. Pesin, Y. Characteristic Lyapunov exponents and smooth ergodic theory. *Russ. Math. Surv.* **1977**, *32*, 55–114. [CrossRef]
34. Kennel, M.; Brown, R.; Abarbanel, H. Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Phys. Rev. A* **1992**, *45*, 3403–3411. [CrossRef] [PubMed]
35. Abarbanel, H.D.I.; Brown, R.; Sidorowich, J.J.; Tsimring, L.S. The analysis of observed chaotic data in physical systems. *Rev. Mod. Phys.* **1993**, *65*, 1331–1392. [CrossRef]

36. Huke, J.P.; Broomhead, D.S. Embedding theorems for non-uniformly sampled dynamical systems. *Nonlinearity* **2007**, *20*, 2205–2244. [CrossRef]
37. Manabe, Y.; Chakraborty, B. A novel approach for estimation of optimal embedding parameters of nonlinear time series by structural learning of neural network. *Neurocomputing* **2007**, *70*, 1360–1371. [CrossRef]
38. Small, M. Optimal time delay embedding for nonlinear time series modeling. *arXiv* **2003**, arXiv:0312011v1. [CrossRef]
39. Vitrano, J.B.; Povinelli, R.J. Selecting dimensions and delay values for a time-delay embedding using a genetic algorithm. In Proceedings of the GECCO'01 3rd Annual Conference on Genetic and Evolutionary Computation, San Francisco, CA, USA, 7–11 July 2001; pp. 1423–1430, ISBN: 1-55860-774-9.
40. Shen, M.; Chen, W.N.; Zhang, J.; Chung, H.S.H.; Kaynak, O. Optimal selection of parameters for nonuniform embedding of chaotic time series using ant colony optimization. *IEEE Trans. Cybern.* **2013**, *43*, 790–802. [CrossRef] [PubMed]
41. Ragulskis, M.; Lukoseviciute, K. Non-uniform attractor embedding for time series forecasting by fuzzy inference systems. *Neurocomputing* **2009**, *72*, 2618–2626. [CrossRef]
42. Lukoseviciute, K.; Ragulskis, M. Evolutionary algorithms for the selection of time lags for time series forecasting by fuzzy inference systems. *Neurocomputing* **2010**, *73*, 2077–2088. [CrossRef]
43. Timofejeva, I.; Poskuviene, K.; Cao, M.S.; Ragulskis, M. Synchronization measure based on a geometric approach to attractor embedding using finite observation windows. *Complexity* **2018**, *2018*, 8259496. [CrossRef]
44. Little, D.J.; Kane, D.M. Permutation entropy with vector embedding delays. *Phy. Rev. E* **2017**, *96*, 062205. [CrossRef] [PubMed]
45. Rössler, O.E. An equation for continuous chaos. *Phys. Lett.* **1976**, *57*, 397–398. [CrossRef]
46. Letellier, C.; Messager, V. Influences on Otto E. Rössler's earliest paper on chaos. *Chaos* **2010**, *20*, 3585–3616. [CrossRef]
47. Amigo, J.M. *Permutation Complexity in Dynamical Systems*; Springer: Berlin, Germany, 2010; p. 280, ISBN 978-3-642-04083-2.
48. Amigo, J.M.; Zambrano, S.; Sanjuan, M.A.F. Combinatorial detection of determinism in noisy time series. *EPL* **2008**, *83*, 60005. [CrossRef]
49. BNCI Horizon 2020 Project Database. Available online: http://bnci-horizon-2020.eu/database/data-sets (accessed on 1 May 2018)
50. Riedl, M.; Müller, A.; Wessel, N. Practical considerations of permutation entropy. *Eur. Phys. J. Spec. Top.* **2013**, *222*, 249–262. [CrossRef]

© 2018 by the authors. 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/).