

Article

Surrogate Data Preserving All the Properties of Ordinal Patterns up to a Certain Length

Yoshito Hirata ^{1,2,*} , Masanori Shiro ³ and José M. Amigó ⁴ 

¹ Mathematics and Informatics Center, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

² Faculty of Engineering, Information and Systems, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8573, Japan

³ Human Informatics Research Institute, National Institute of Advanced Industrial Science and Technology, Ibaraki 305-8568, Japan

⁴ Centro de Investigación Operativa, Universidad Miguel Hernández, Avda. de la Universidad s/n, 03202 Elche, Spain

* Correspondence: hirata@cs.tsukuba.ac.jp

Received: 15 June 2019; Accepted: 19 July 2019; Published: 22 July 2019



Abstract: We propose a method for generating surrogate data that preserves all the properties of ordinal patterns up to a certain length, such as the numbers of allowed/forbidden ordinal patterns and transition likelihoods from ordinal patterns into others. The null hypothesis is that the details of the underlying dynamics do not matter beyond the refinements of ordinal patterns finer than a predefined length. The proposed surrogate data help construct a test of determinism that is free from the common linearity assumption for a null-hypothesis.

Keywords: time series analysis; determinism; stochasticity; permutations; hypothesis testing

1. Introduction

Judging whether the underlying dynamics are deterministic or stochastic based on a given time series is an old problem and the first step for modelling such a time series. The current standard approach uses iterative amplitude adjusted Fourier transform (IAAFT) surrogates [1] with some statistics characterizing determinism such as prediction errors [2] and Wayland statistic [3]—but by following this approach, we cannot distinguish nonlinear stochasticity from linear stochasticity or nonlinear determinism.

Recently, we have proposed an alternative approach where we prepare two independent tests for linearity-nonlinearity as well as determinism-stochasticity [4]. For the test of linearity-nonlinearity, we use truncated Fourier transform surrogates (TFTS) [5], an extension of IAAFT surrogates with the mean of $s(t)^2s(t+1)^2$ over a time series $s(t)$ as a test statistic which is not directly related to the determinism that may exist. For the test of determinism-stochasticity, we use the properties of permutations [6–8], which are inequality relations among consecutive measurements: If the underlying dynamics is deterministic and verifies some assumptions (see Section 3 for details), then the number of appearing permutations increases exponentially when the length of permutations is prolonged. But, currently this approach has a problem—we need a long time series of length 1,000,000 to classify stationary time series appropriately [4].

Thus, we propose another approach for testing determinism-stochasticity for the underlying dynamics using the permutation properties of a time series. In this paper, we generate surrogate data which preserve the series of permutations for a given time series almost perfectly and thus the stochastic properties for the underlying dynamics fully up to a certain pre-defined length of the

permutations. We call the surrogate data we propose here as entropy preserving surrogates (EPS). Thus, based on the proposed method, we will be able to identify the determinism for the underlying dynamics based on its time series more firmly than by using the existing methods in the literature, helping researchers to make their mathematical model with more confidence.

For the test of linearity–nonlinearity, we continue using TFTS with the mean of $s(t)^2s(t+1)^2$ as the statistic. If $(s(t), s(t+1))$ follows the multivariate Gaussian distribution, a higher-order moment such as $s(t)^2s(t+1)^2$ can be characterized with the means and variances [9] and becomes pivotal [10] and constant if the underlying dynamics are kept. Thus, any variation can be attributed to a deviation from the linear Gaussianity. Hence, the mean of $s(t)^2s(t+1)^2$ can be used as a test statistic for nonlinearity. We demonstrate the proposed set of methods using time series of length 1000.

2. Our Mathematical Settings

Before starting the main parts of this manuscript, we define our mathematical settings more rigorously.

Our interest is on a dynamical system $f : X \times P \rightarrow X$ on a manifold X driven by a parameter space P , which may change along the time. Thus, typically, we have

$$x(t+1) = f(x(t), p(t)) \quad (1)$$

for $x(t) \in X$ and $p(t) \in P$, starting from the initial conditions $x(0) \in X$ and $p(0) \in P$. We cannot directly observe $x(t)$. Instead, we have an observation function $g : X \rightarrow \mathbf{R}$ such that $s(t) = g(x(t))$. When g is given by a skew product of the state X and its disturbance Q , then we can model observational noise as well.

Then, our question is whether $p(t)$ is constant throughout the time or $p(t)$ changes along the time. If $p(t)$ is constant throughout the time, then we call the underlying dynamics deterministic. If $p(t)$ changes along the time in a deterministic way, then we also call the underlying dynamics deterministic. If $p(t)$ changes along the time randomly, then we call the underlying dynamics stochastic.

3. Background

There have been a number of researches in the existing literature discussing how to characterize determinism and/or stochasticity: The best known approaches could be the ones using the parallelness of neighboring orbits [3,11] and the optimal neighborhood size for local linear predictions [12]. Recently, the most popular one could be that by Amigó et al. [8,13,14], which uses the fact that there exist forbidden ordinal patterns. To explain the approach of Amigó et al. (2008) [8] in more detail, we first define ordinal patterns or permutations [6].

Suppose that a time series is given by $s(t) \in \mathbf{R}$. We focus on inequality relations among consecutive measurements $s(t), s(t+1), \dots, s(t+L-1)$ over time period between t and $t+L-1$. Namely, if we order these measurements in the ascending order, we could have $s(t+i_1) \leq s(t+i_2) \leq \dots \leq s(t+i_L)$, where $i_j \in \{0, 1, 2, \dots, L-1\}$ and are unique. For convenience, we define $s(t+i) \leq s(t+j)$ if $s(t+i) = s(t+j)$ and $i < j$. Then, the corresponding permutation is $\pi(\{s\}, t, L) = (i_1, i_2, \dots, i_L)$.

The number of appearing permutations increases exponentially when the length of permutations L is prolonged if the underlying dynamics is one-dimensional, deterministic and piecewise monotone [13] or, in any dimension, if the underlying dynamics is deterministic and expansive [7].

Thus, Amigó et al. (2008) [8] uses the existence of forbidden permutations as the signature for a deterministic system. The contraposition of this theorem was previously used for identifying if a given time series is generated from a nonlinear and stochastic system [4].

Moreover, the entropies obtained using the permutation statistics can be used for estimating the metric and topological entropies [6,7,15].

Therefore, permutations are good tools for characterizing time series generated from the underlying dynamics.

4. Methods

Here we propose to generate surrogate data that preserves all the statistical properties of permutations up to a certain predefined length L . Here we call such surrogate data as entropy preserving surrogates (EPS). Our method is quite simple and follows a general principle proposed by Schreiber (1998) [16]: we randomly exchange the temporal order of time series through the method of simulated annealing [17] so that we preserve a series of permutations for a given time series as well as a series of permutations for a moving average of the given time series over length L subsampled by an interval L (see Figure 1). In this way, we generate 39 surrogate data for obtaining the significance level of $2 / (39 + 1) = 5\%$ level for each time series. Since a series of permutations is preserved in the entropy preserving surrogates, all the transitions from every permutation to another are preserved. Therefore, our null hypothesis in the entropy preserving surrogates is that the underlying dynamics has significant historical dependence only up to the length L and the dependence over L does not matter for the underlying dynamics. As a by-product, permutation entropies calculated up to length L are preserved. Please find the detail on how to generate EPS in the Appendix A.

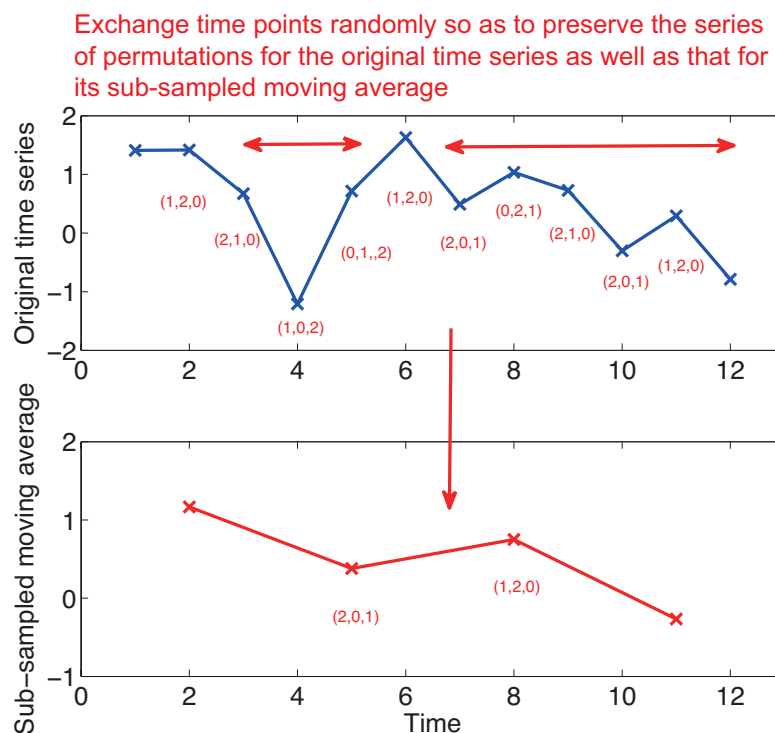


Figure 1. Schematic figure showing how we generate an entropy preserving surrogate.

To compare an original time series with its surrogate data for telling whether the original time series is statistically different from its surrogate data or not, we estimate the maximal Lyapunov exponent in the following way: First, we fit the parameters $a(t)$ and $b(t)$ for the following local linear model for each time t using 20 neighboring points in infinite-dimensional delay coordinates [18]:

$$s(t + 1) - s(\tau + 1) \approx a(t) + b(t)(s(t) - s(\tau)), \tag{2}$$

where $s(\tau)$ is one of 20 spatial neighbors for $s(t)$. Then, we evaluate the following quantity as a test statistic for the second half of each dataset:

$$E_t[\log |b(t)|]. \quad (3)$$

This statistic can be regarded as a proxy for the maximal Lyapunov exponent. We decided to use 20 neighbors for the above estimation because if the number of neighbors is less than 20, then the estimation would heavily depend on the closer neighbors, while the estimation would not be able to characterize local states well if the number of neighbors is greater than 20.

5. Results

5.1. Toy Examples

First, we show some numerical experiments for datasets generated from toy models which are free from observational noise. We set $L = 30$ throughout the paper because we would like to investigate the deterministic structure which finely persists over the pseudo-periodicity evaluated by the pseudo-periodic surrogates [19]. To obtain pseudo-periodic surrogates, we used the three-dimensional delay coordinates with delay 8.

Our first toy model is the first-order autoregressive linear (AR(1)) model [20]. The model we used is as follows:

$$x(t+1) = 0.8x(t) + \eta(t), \quad (4)$$

where $\eta(t)$ follows the Gaussian distribution of mean 0 and standard deviation 1.

The second toy model is the GARCH model [21]. The model equations are

$$y(t) = 0.409933 + 0.095y(t-1) + \epsilon(t), \quad (5)$$

$$h(t) = 14.4038 + 0.095\epsilon(t-1)^2 + 0.895h(t-1), \quad (6)$$

where $\epsilon(t)$ follows the Gaussian distribution of mean 0 and standard deviation $\sqrt{h(t)}$. We observe $y(t)$ to generate a time series.

The third toy model is the model for noise-induced order [22]. We use the following equations:

$$x(t+1) = f(x(t)) + b + 10^{-2.5}u(t), \quad (7)$$

$$f(x) = \begin{cases} \left(-(0.125 - x)^{1/3} + 0.50607357 \right) \exp(-x), & \text{if } x < 0.125, \\ \left((x - 0.125)^{1/3} + 0.50607357 \right) \exp(-x), & \text{if } 0.125 \leq x < 0.3, \\ 0.121205602 \left((10x \exp(-10x/3))^{19} \right), & \text{otherwise,} \end{cases} \quad (8)$$

where $u(t)$ follows the uniform distribution between -1 and 1 .

The fourth toy model is the logistic map [23]. We use the following equation:

$$x(t+1) = 3.8x(t)(1 - x(t)). \quad (9)$$

We also use time-continuous models for testing the proposed method. Our fifth toy model is the Lorenz model [24]. We use the following equations:

$$\dot{x} = -10(x - y), \quad (10)$$

$$\dot{y} = -xz + 28x - y, \quad (11)$$

$$\dot{z} = xy - \frac{8}{3}z. \quad (12)$$

We sampled x every 0.1 unit time.

The sixth model is the Rössler model [25]. Here we use the following equations:

$$\dot{x} = -(y + z), \quad (13)$$

$$\dot{y} = x + 0.36y, \quad (14)$$

$$\dot{z} = 0.4 + z(x - 4.5). \quad (15)$$

We sampled x every 1 unit time.

For each model, we generated 20 time series of length 1000 to examine the robustness for the proposed test. In this paper, we also used pseudo-periodic surrogates [19] with correlation dimensions [26] as test statistics. For pseudo-periodic surrogates, the null hypothesis is that the underlying dynamics has determinism beyond pseudo-periodicity. Such surrogate data can be generated by connecting segments of time series by choosing a neighboring point at each step with a Gaussian uncertainty. If we generate surrogate data in this way, a rough periodicity related to the underlying dynamics is preserved, while fine structure is destroyed. Thus, we can judge if there is determinism beyond this rough periodicity. For the cases without observation noise, we also use the proxy for the maximal Lyapunov exponent as a test statistic for pseudo-periodic surrogates for comparison.

When we use TFTS, we apply the end-to-end matching [27] using the first and last 20 points to suppress the artificial high-frequency components which might be generated during applying the Fourier transforms. When we generate the proposed entropy preserving surrogates, we use the same segments of time series, which could be the reason why we can find slight differences between the values for the proxies of the maximal Lyapunov exponent between Figures 6 and 7 as we will discuss later.

In all the model analyses down below, we used whole the datasets for each time series, meaning that we did not divide each time series into halves or so.

Examples for entropy preserving surrogates are shown in Figures 2 and 3. Especially, such a time series shown for entropy preserving surrogates looks similar to the original time series (Figure 2). When we look at their return plots, we can see that an entropy preserving surrogate (Figure 3B) seems to be perturbed from the original time series (Figure 3A).

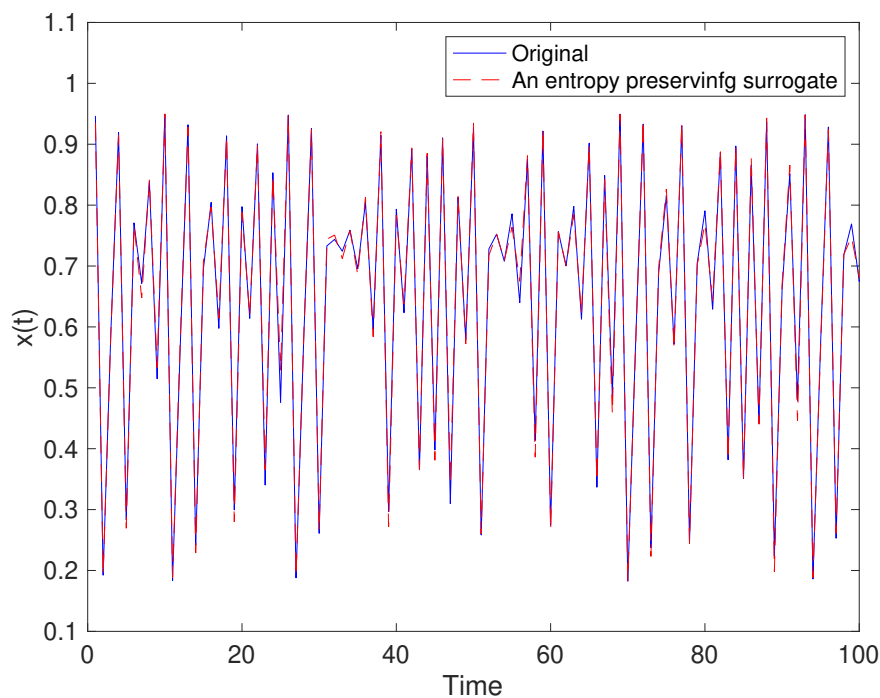


Figure 2. Example of an entropy preserving surrogate for the logistic map.

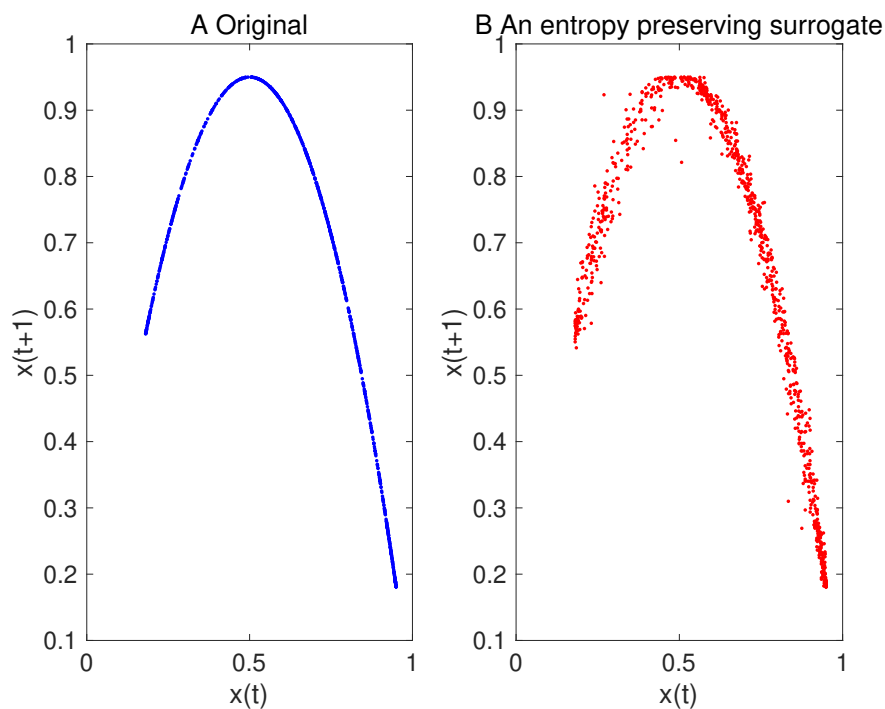


Figure 3. Return plot for the original time series of the logistic map (A) and that for one of its entropy preserving surrogates (B).

The results of the surrogate tests are summarized in Figures 4–7 as well as Table 1. For most cases, the tested time series were classified into the correct classes for the corresponding toy models. To evaluate the numbers of rejections appropriately, consider the binomial distribution with $N = 20$ trials and $p = 0.05$. Then, the cumulative sum of probabilities from 0 becomes more than 95% if the number of positives is 4 or greater. For example, 3 rejections for the test of linearity for the AR(1) model are not statistically significant. For the same reason, 3 rejections for the proposed test of determinism beyond 30 steps for the model of noise-induced order are not statistically significant.

The results presented in Figure 4 and Table 1 show that the nonlinearity test examined here is robust.

Table 1 and comparison of Figures 5B,C and 6B,C mean that the results of pseudo-periodic surrogates heavily depend on test statistics we use. These results may be due to the fact that pseudo-periodic surrogates are typical realizations and need a pivotal statistic for the test [10].

Figure 7C,D mean that a time series with the same value for the permutation entropies up to length 30 is likely to have a positive Lyapunov exponent even if the underlying dynamics is stochastic. This usage could lead to another implication obtained from the entropy preserving surrogates. The results shown in Table 1 mean that the proposed method has some skill for detecting the determinism for the underlying dynamics.

The results presented in Table 1 also show that the proposed method works well even for flows such as the Lorenz model and Rössler models as far as sampling intervals are chosen appropriately.

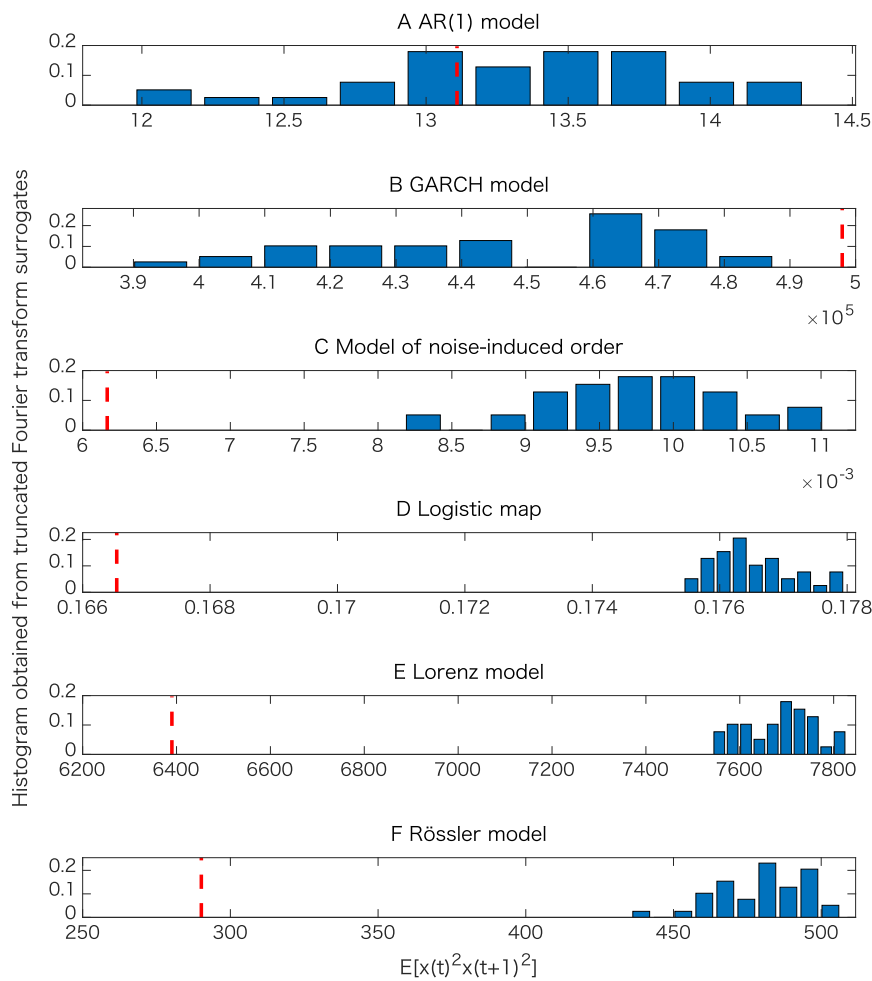


Figure 4. Examples of tests for nonlinearity for various models when the datasets are free from observational noise. Here an important point is whether or not the value obtained from each original time series shown in the red vertical dashed line is within the interval specified with the minimum and the maximum for the test statistic $E[x(t)^2x(t + 1)^2]$ of the 39 truncated Fourier transform surrogates (TFTS) surrogates, which can be interpreted from each histogram. Therefore, it does not matter much whether the test statistic obtained from the original data is smaller or greater than those obtained from TFTS surrogates. (A) result for the AR(1) model; (B) result for the GARCH model; (C) result for the model of noise-induced order; (D) result for the logistic map; (E) result for the Lorenz model; (F) result for the Rössler model.

Table 1. Results of noise free data summarized as classifications. We counted the number of rejections for each test for each model. The italic numbers correspond to the significant numbers of rejections based on the calculations using the binomial distributions.

Property \ Model	AR(1)	GARCH	Noise-Induced Order	Logistic	Lorenz	Rössler
Nonlinearity with $E[x(t)^2x(t + 1)^2]$	3	20	20	20	20	20
Determinism beyond pseudo-periodicity with correlation dimensions	0	20	0	0	0	20
Determinism beyond psuedo-periodicity with maximal Lyapunov exponent	1	1	7	17	2	0
Determinism beyond 30 steps with maximal Lyapunov exponent	2	1	3	20	8	6
Total	20	20	20	20	20	20

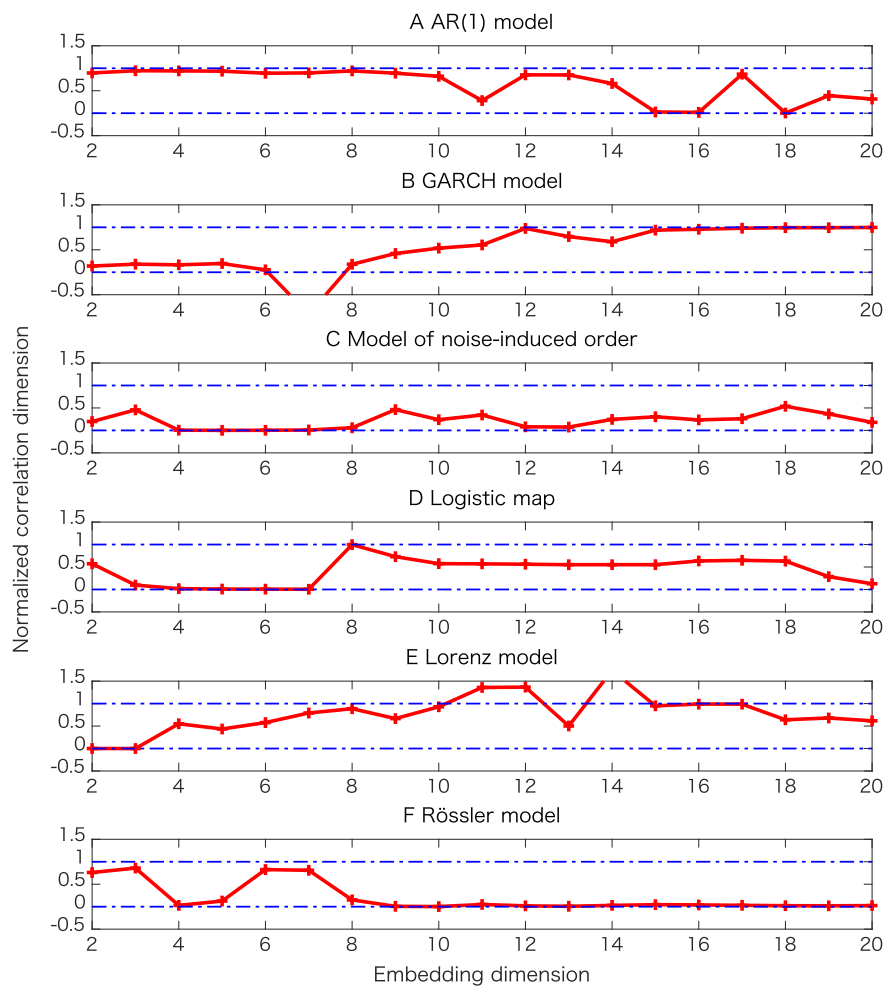


Figure 5. Examples of tests of determinism beyond pseudo-periodicity using pseudo-periodic surrogates for various models when the datasets are free from observational noise. Here we use the correlation dimensions as test statistics. In this surrogate data, rough periodic behavior is preserved, while fine structure related to the possible underlying determinism in question is destroyed. Correlation dimensions are normalized so that the minimum and the maximum values for the correlation dimensions of the 39 pseudo-periodic surrogates for each dimension become 0 and 1, respectively. **(A)** result for the AR(1) model; **(B)** result for the GARCH model; **(C)** result for the model of noise-induced order; **(D)** result for the logistic map; **(E)** result for the Lorenz model; **(F)** result for the Rössler model.

We also tested the cases where for each case, we added Gaussian observational noise of mean 0 and standard deviation which is 5% of the standard deviation of the original time series (Figures 8–10, and Table 2). But, still the proposed method seems to work properly. Determinism beyond pseudo-periodicity was detected for the GARCH model and the model of noise-induced order, while the determinism was weak in the sense that the dependence did not persist beyond 30 steps statistically significantly. On the other hand, the logistic map tended to exhibit determinism beyond 30 steps (Table 2). Overall, Table 2 shows the robustness for the proposed method against observational noise.

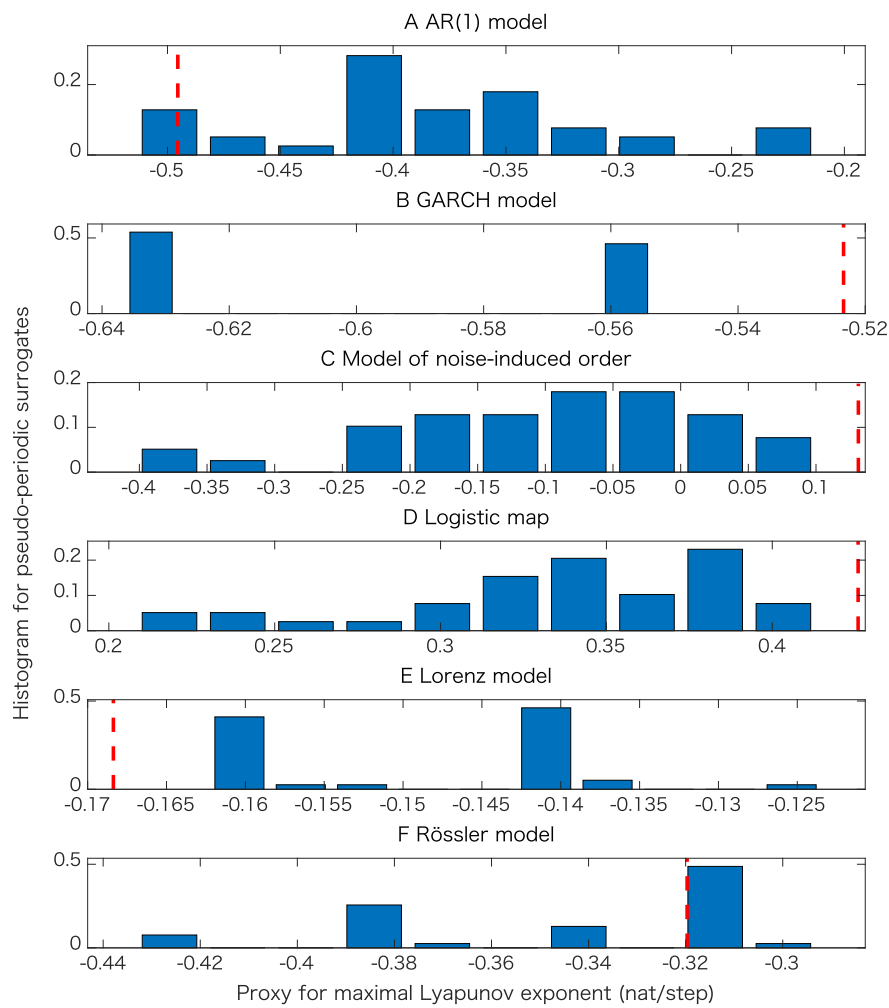


Figure 6. Examples of tests of determinism beyond pseudo-periodicity using pseudo-periodic surrogates when we use the proxy for the maximal Lyapunov exponent as a test statistic. In each panel, the red dashed line corresponds to the value obtained from the original time series and the histogram, obtained from the pseudo-periodic surrogates. (A) result for the AR(1) model; (B) result for the GARCH model; (C) result for the model of noise-induced order; (D) result for the logistic map; (E) result for the Lorenz model; (F) result for the Rössler model.

Table 2. Results of 5% observational noise data summarized as classifications. See the caption of Table 1 to interpret the results.

Property\Model	AR(1)	GARCH	Noise-Induced Order	Logistic	Lorenz	Rössler
Nonlinearity with $E[x(t)^2x(t+1)^2]$	1	19	20	20	20	20
Determinism beyond pseudo-periodicity with correlation dimensions	0	20	20	0	20	11
Determinism beyond 30 steps with maximal Lyapunov exponent	3	0	2	16	6	7
Total	20	20	20	20	20	20

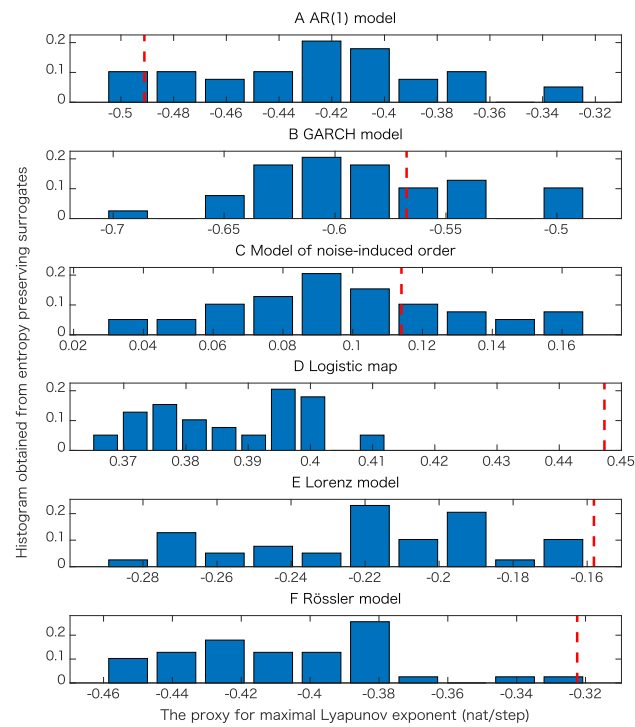


Figure 7. Examples of tests of determinism beyond 30 steps using the proposed entropy preserving surrogates for various models when the datasets are free from observational noise. In each panel, the red dashed line corresponds to the value of test statistic obtained from the original data. (A) result for the AR(1) model; (B) result for the GARCH model; (C) result for the model of noise-induced order; (D) result for the logistic map; (E) result for the Lorenz model; (F) result for the Rössler model.

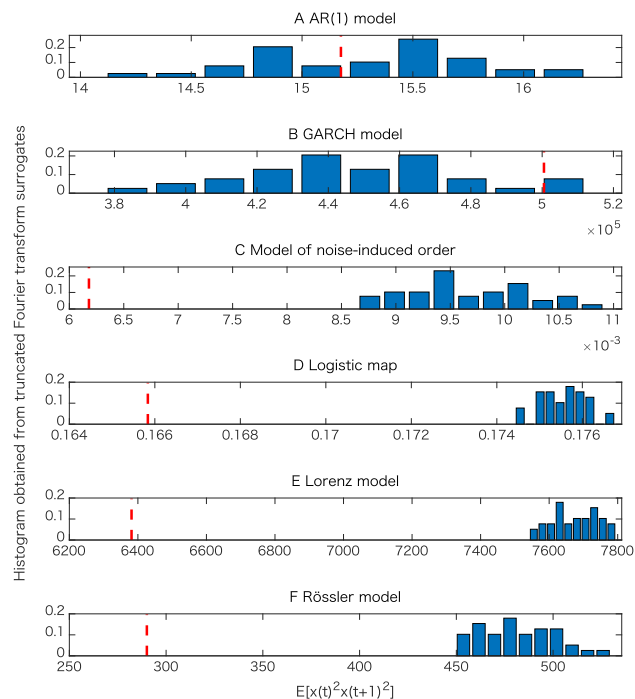


Figure 8. Examples of tests of nonlinearity for various models when 5% observational noise is added. See the caption of Figure 5 to interpret the results. (A) result for the AR(1) model; (B) result for the GARCH model; (C) result for the model of noise-induced order; (D) result for the logistic map; (E) result for the Lorenz model; (F) result for the Rössler model.

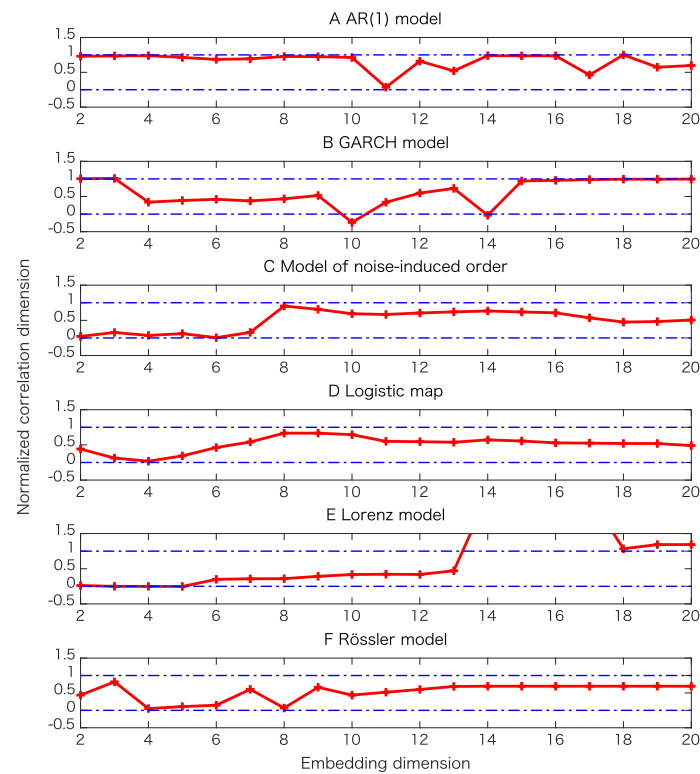


Figure 9. Examples of tests of determinism beyond pseudo-periodicity using pseudo-periodic surrogates for various models when 5% observational noise is added. (A) result for the AR(1) model; (B) result for the GARCH model; (C) result for the model of noise-induced order; (D) result for the logistic map; (E) result for the Lorenz model; (F) result for the Rössler model.

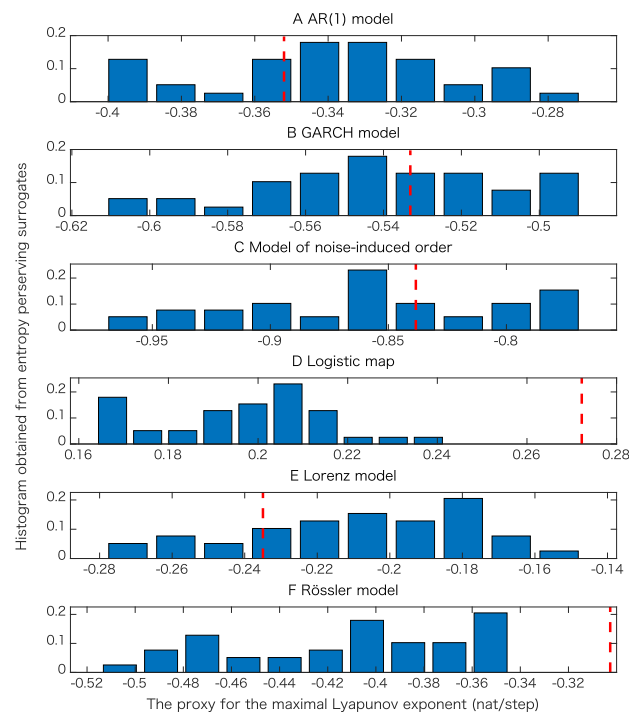


Figure 10. Examples of tests of determinism using the proposed entropy preserving surrogates for various models when 5% observational noise is added. (A) result for the AR(1) model; (B) result for the GARCH model; (C) result for the model of noise-induced order; (D) result for the logistic map; (E) result for the Lorenz model; (F) result for the Rössler model.

5.2. Real Data Example of the USD/JPY Market

We analyzed the dataset of the USD/JPY market compiled by the Thomson Reuters Cooperation. The record starts from 1 January 2006 and ends on 31 December 2015. We use the first 100,000 quotes for the analysis here. We divided the dataset by every 1000 quotes into 100 segments, and took inter-quote intervals for each segment.

For the first segment, one of generated entropy preserving surrogates looks as shown in Figures 11 and 12. We can see that typical characteristics for the time series as well as return plots are almost preserved.

The results are summarized as Table 3. Nonlinearity was detected in 24 out of 100 cases, while determinism beyond 30 steps was detected in 12 out of 100 cases. Because these numbers are significant from the viewpoint of the binomial distribution of 100 trials and the probability 0.05 for each test, namely judging from the facts that each test is 5% significant and each time segment is independent from each other, overall, the dataset of the USD/JPY market seems nonlinear with the determinism beyond 30 quotes.

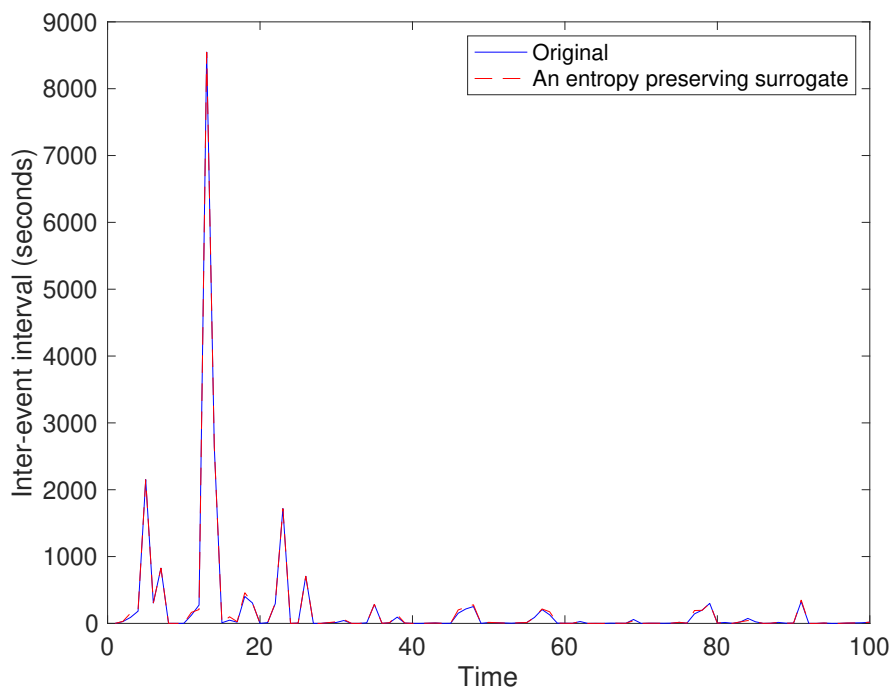


Figure 11. Example of an entropy preserving surrogate for a part of the USD/JPY data.

Table 3. Results of the USD/JPY data summarized as classifications. See the caption of Table 1 to interpret the results.

Property	Number of Time Segments
Nonlinearity with $E[x(t)^2x(t+1)^2]$	24
Determinism beyond pseudo-periodicity with correlation dimensions	0
Determinism beyond 30 steps with maximal Lyapunov exponent	12
Total	100

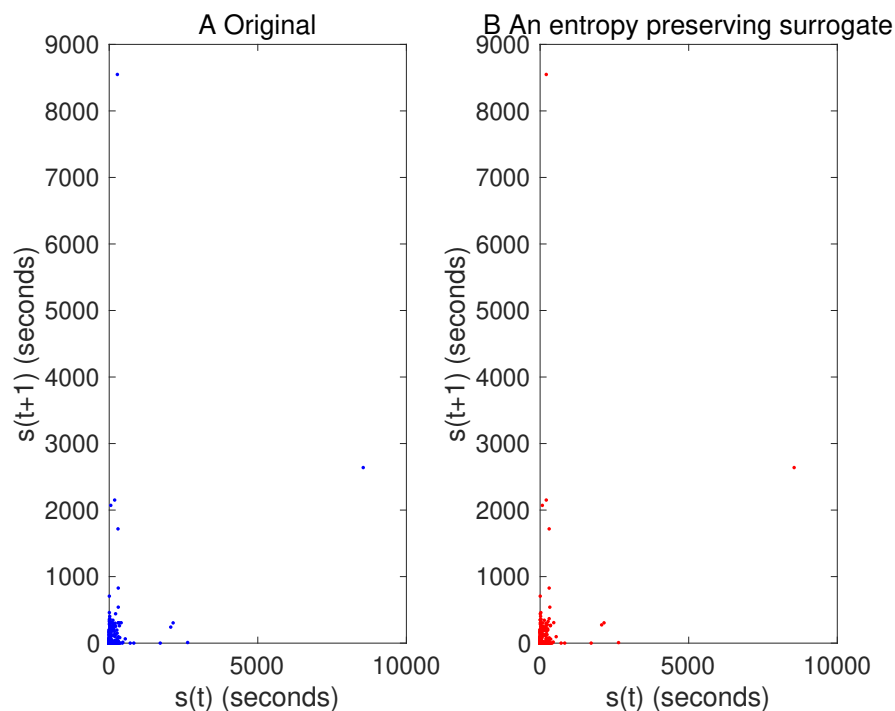


Figure 12. Return plot for the original time series of a USD/JPY data part (A) and that for one of its entropy preserving surrogates (B).

6. Discussions

Although we set $L = 30$ in this manuscript, we may vary the length L of permutations for elucidating the effect and the length of dynamical dependence. By choosing L , we can control the length of dependence which should have significant meaning. Thus, by varying L , we can narrow down the topical area of a target time series mostly into the intersection of nonlinear and deterministic regions, whose regions could be smaller than the region specified with pseudo-periodic surrogates as shown in Figure 13. Hence, together with the methods [5,19] in the existing literature, the proposed entropy preserving surrogate helps us to specify the assumptions of a model more finely when we try to construct a model based on a time series.

For pseudo-periodic surrogates, the length 1000 of time series might have been too short to show the determinism beyond pseudo-periodicity for the dataset of the USD/JPY data, while it was sufficient to show the determinism beyond 30 quotes using the proposed method. Thus, we would like to explore the effect for the length of time series in the future more deeply.

The proposed method preserves series of permutations for the original time series as well as that for its sub-sampled moving average. Thus, the proposed method of entropy preserving surrogates can be regarded as a constrained realization [10] rather than a typical realization. When we focus on surrogate data generated by permutations, there are methods such as those of References [28–30]. Because these methods are surrogate data as typical realizations, the proposed method is the first method generating surrogate data with permutations as a constrained realization. As a constrained realization, the proposed method can formally be used with a non-pivotal statistic [10], which does not have to provide a consistent value for a class of null models. Thus, we hope that the proposed method be powerful for investigating the deterministic properties beyond a pre-defined length for a given time series.

If there are a pair of time series and we generate entropy preserving surrogates for both, then we can also preserve symbolic transfer entropies [31] and transcripts [32]. Therefore, applying entropy preserving surrogates to multivariate data time series could be an interesting and open problem.

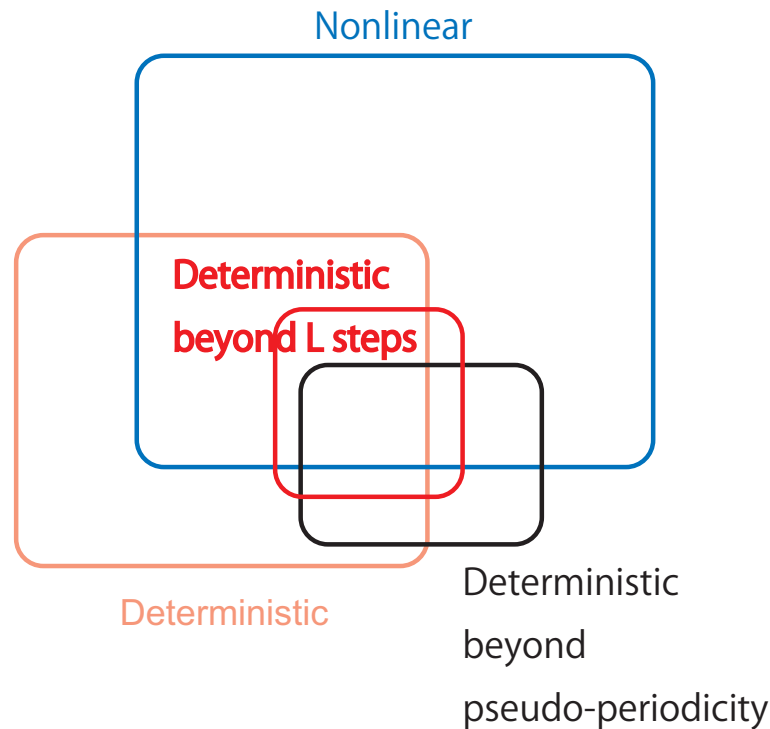


Figure 13. The Venn diagram describing the relationship among original properties for the underlying dynamics such as nonlinearity and determinism against properties we can identify with surrogate data such as determinism beyond pseudo-periodicity (pseudo-periodic surrogates [19]) and determinism beyond L steps (the proposed entropy preserving surrogates).

7. Conclusions

We have proposed a method for generating surrogate data such that all the properties of permutations up to a certain length are preserved. Such surrogate data look very similar to the original data as shown in Figures 2 and 11, but with dynamical noise especially demonstrated in Figure 3. By using the four toy models, we evaluated that the proposed method works finely. Then, we applied the proposed method to inter-quote interval data in the USD/JPY market and found that the market behaved in a nonlinear and deterministic manner, which is consistent with our previous findings [33].

Author Contributions: Conceptualization, Y.H., M.S. and J.M.A.; methodology, Y.H.; numerical experiments, Y.H.; writing—original draft Y.H.; writing—review and editing, Y.H., M.S. and J.M.A.; supervision, J.M.A.; funding acquisition, Y.H.

Funding: The research of Y.H. is supported by JSPS KAKENHI Grant Number JP18K11461.

Acknowledgments: We thank Michael Small (University of Western Australia) very much for his making his codes freely available, with which we generated pseudo-periodic surrogates [19] as well as calculated the correlation dimensions [26] throughout this paper.

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

Appendix A

Let $\{s(t) \in \mathbf{R} | t = 1, 2, \dots, T\}$ be a given time series. Then, its moving average of L consecutive points sub-sampled by every L time points can be defined by $\{\bar{s}(u) = \frac{1}{L} \sum_{i=1}^L s((u-1)L + i) | u = 1, 2, \dots, \lfloor T/L \rfloor\}$.

First, we convert the given time series $\{s(t)\}$ and its moving average $\{\bar{s}\}$ to the corresponding permutation series $\{\pi(\{s\}, t, L) | t = 1, 2, \dots, T - L + 1\}$ and $\{\pi(\{\bar{s}\}, u, L) | u = 1, 2, \dots, \lfloor T/L \rfloor - L + 1\}$.

Second, we initialize our simulated annealing algorithm by setting the current time series $\{c(t)\}$ to the original time series $\{s(t)\}$.

Third, we repeat the following process until the number of iterations reaches $(NS + 10) \times S$, where we set $NS = 39$, which is the number of surrogate data, and $S = 10,000$, which is the number of iterations we skip:

1. Increment the current number i of iterations by 1.
2. Prepare an attempt $a(t)$ for replacement by swapping two elements of $\{c(t)\}$.
3. Calculate $\{\pi(\{a\}, t, L) | t = 1, 2, \dots, T - L + 1\}$ and $\{\pi(\{\bar{a}\}, u, L) | u = 1, 2, \dots, \lfloor T/L \rfloor - L + 1\}$.
4. Calculate the number of differences between $\{\pi(\{s\}, t, L) | t = 1, 2, \dots, T - L + 1\}$, $\{\pi(\{\bar{s}\}, u, L) | u = 1, 2, \dots, \lfloor T/L \rfloor - L + 1\}$ and $\{\pi(\{a\}, t, L) | t = 1, 2, \dots, T - L + 1\}$, $\{\pi(\{\bar{a}\}, u, L) | u = 1, 2, \dots, \lfloor T/L \rfloor - L + 1\}$. Let $\#n$ denote this number.
5. Let p be the probability for accepting the attempt, which can be calculated as $\exp[-i\beta\#n]$.
6. Generate a uniform random number between 0 and 1. If the random number is less than p , then replace the current time series $\{c(t)\}$ by the attempt $\{a(t)\}$.
7. If i is a multiple of S and $i > 10S$, then record the current $\{c(t)\}$ as the $(i/S - 10)$ -th surrogate data.

References

1. Schreiber, T.; Schmitz, A. Improved surrogate data for nonlinearity tests. *Phys. Rev. Lett.* **1996**, *77*, 635–638. [[CrossRef](#)]
2. Theiler, J.; Eubank, S.; Longtin, A.; Galdrikian, B.; Farmer, J.D. Testing for nonlinearity in time series: The method of surrogate data. *Phys. D* **1992**, *58*, 77–94. [[CrossRef](#)]
3. Wayl, R.; Bromley, D.; Pickett, D.; Passamante, A. Recognizing determinism in a time series. *Phys. Rev. Lett.* **1993**, *70*, 580–582.
4. Hirata, Y.; Shiro, M. Detecting nonlinear stochastic systems using two independent hypothesis tests. *Phys. Rev. E* **2019**, in press.
5. Nakamura, T.; Small, M.; Hirata, Y. Testing for nonlinearity in irregular fluctuations with long-term trends. *Phys. Rev. E* **2006**, *74*, 026205. [[CrossRef](#)] [[PubMed](#)]
6. Bandt, C.; Pompe, B. Permutation entropy: A natural complexity measure for time series. *Phys. Rev. Lett.* **2002**, *88*, 174102. [[CrossRef](#)] [[PubMed](#)]
7. Amigó, J.M.; Kennel, M.B. Topological permutation entropy. *Phys. D* **2007**, *231*, 137–142. [[CrossRef](#)]
8. Amigó, J.M.; Zambrano, S.; Sanjuán, M.A.F. Combinatorial detection of determinism in noisy time series. *EPL* **2008**, *83*, 60005. [[CrossRef](#)]
9. Michalowicz, J.V.; Nichols, J.M.; Bucholtz, F.; Olson, C.C. An Isserlis' theorem for mixed Gaussian variables: Application to the auto-bispectral density. *J. Stat. Phys.* **2009**, *136*, 89–102. [[CrossRef](#)]
10. Theiler, J.; Prichard, D. Constrained-realization Monte-Carlo method for hypothesis testing. *Phys. D* **1996**, *94*, 221–235. [[CrossRef](#)]
11. Kaplan, D.T.; Glass, L. Direct test for determinism in a time series. *Phys. Rev. Lett.* **1992**, *68*, 427–430. [[CrossRef](#)] [[PubMed](#)]
12. Casdagli, M.C.; Weigend, A.S. Exploring the continuum between deterministic and stochastic modeling. In *Time Series Prediction: Forecasting the Future and Understanding the Past*; Weigend, A.S., Gershenfeld, N.A., Eds.; Westview Press: New York, NY, USA, 1993; pp. 347–366.
13. Amigó, J.M.; Kocarev, L.; Szczepanski, J. Order patterns and chaos. *Phys. Lett. A* **2006**, *355*, 27–36. [[CrossRef](#)]
14. Amigó, J.M.; Zambrano, S.; Sanjuán, M.A.F. Detecting determinism with ordinal patterns: A comparative study. *Int. J. Bifurcat. Chaos* **2010**, *20*, 2915–2924. [[CrossRef](#)]
15. Amigó, J.M.; Kennel, M.B.; Kocarev, L. The permutation entropy rate equals the metric entropy rate for ergodic information sources and ergodic dynamical systems. *Phys. D* **2005**, *210*, 77–95. [[CrossRef](#)]
16. Schreiber, T. Constrained randomization of time series data. *Phys. Rev. Lett.* **1998**, *80*, 2105–2108. [[CrossRef](#)]
17. Gershenfeld, N. *The Nature of Mathematical Modeling*; Cambridge University Press: Cambridge, UK, 1998.
18. Hirata, Y.; Takeuchi, T.; Horai, S.; Suzuki, H.; Aihara, K. Parsimonious description for predicting high-dimensional dynamics. *Sci. Rep.* **2015**, *5*, 15736. [[CrossRef](#)] [[PubMed](#)]
19. Small, M.; Yu, D.; Harrison, R.G. Surrogate test for pseudoperiodic time series data. *Phys. Rev. Lett.* **2001**, *87*, 188101. [[CrossRef](#)]
20. Hamilton, J.D. *Time Series Analysis*; Princeton University Press: Princeton, NJ, USA, 1994.

21. Lamoureux, C.G.; Lastrapes, W.D. Persistence in variance, structural change, and the GARCH model. *J. Bus. Econ. Stat.* **1990**, *8*, 225–234.
22. Matsumoto, K.; Tsuda, I. Noise-induced order. *J. Stat. Phys.* **1983**, *31*, 87–106. [[CrossRef](#)]
23. May, R.M. Simple mathematical models with very complicated dynamics. *Nature* **1976** *261*, 459–467. [[CrossRef](#)]
24. Lorenz, E.N. Deterministic nonperiodic flow. *J. Atmos. Sci.* **1963**, *20*, 130–141. [[CrossRef](#)]
25. Rössler, O.E. An equation for continuous chaos. *Phys. Lett.* **1976**, *57A*, 397–398. [[CrossRef](#)]
26. Yu, D.J.; Small, M.; Harrison, R.G.; Diks, C. Efficient implementation of the Gaussian kernel algorithm in estimating invariants and noise level from noisy time series data. *Phys. Rev. E* **2000**, *61*, 3750–3756. [[CrossRef](#)]
27. Schreiber, T.; Schmitz, A. Surrogate time series. *Phys. D* **2000**, *142*, 346–382. [[CrossRef](#)]
28. Hirata, Y.; Amigó, J.A.; Matsuzaka, Y.; Yokota, R.; Mushiake, H.; Aihara, K. Detecting causality by combined use of multiple methods: climate and brain examples. *PLoS ONE* **2016**, *11*, e0158572. [[CrossRef](#)] [[PubMed](#)]
29. McCullough, M.; Sakellariou, K.; Stemler, T.; Small, M. Regenerating time series from ordinal networks. *Chaos* **2017**, *27*, 035814. [[CrossRef](#)]
30. Small, M.; McCullough, M.; Sakellariou, K. Ordinal network measures: Quantifying determinism in data. In Proceedings of the 2018 IEEE International Symposium on Circuits and Systems (ISCAS), Florence, Italy, 27–30 May 2018.
31. Staniek, M.; Lehnertz, K. Symbolic transfer entropy. *Phys. Rev. Lett.* **2008**, *100*, 158101. [[CrossRef](#)] [[PubMed](#)]
32. Amigó, J.M.; Monetti, R.; Aschenbrenner, T.; Bunk, W. Transcripts: An algebraic approach to coupled time series. *Chaos* **2012**, *22*, 013105. [[CrossRef](#)]
33. Hirata, Y.; Aihara, K. Timing matters in foreign exchange markets. *Phys. A* **2012**, *391*, 760–766. [[CrossRef](#)]

Sample Availability: Matlab codes are available from the corresponding author's following website: <https://sites.google.com/view/yoshitohirata/home>.



© 2019 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/>).