Effects of data windows on the methods of surrogate data

Tomoya Suzuki,[1] Tohru Ikeguchi,[2] and Masuo Suzuki[1]
[1]*Graduate School of Science, Tokyo University of Science, 1-3 Kagurazaka, Shinjuku-ku, Tokyo 162-8601, Japan*
[2]*Graduate School of Science and Engineering, Saitama University, 225 Shimo-Ohkubo, Sakura-ku, Saitama-city 338-8570, Japan*
(Received 26 October 2004; published 27 May 2005)

To generate surrogate data in nonlinear time series analysis, the Fourier transform is generally used. In the calculation of the Fourier transform, the time series is assumed to be periodic. Because such an assumption does not always hold true, the estimation accuracy of the Fourier transformed data and thus the power spectra is reduced. Due to such an estimation error, it is also possible that the surrogate test will lead to a false conclusion; for example, that a linear time series is nonlinear. In this paper, we experimentally evaluated the effects of data windows from the viewpoint of false rejections with several types of surrogate data. Our results indicate that if the data length becomes shorter, the false rejections by the data windows are reduced to a greater extent. However, if the data length is sufficient, the use of data windows is not a viable option. In the worst possible case wherein the linear memory of the original data is very long as in the nonstationary case, the critical length of the data for which the data windows were effective was approximately 1000.

DOI: 10.1103/PhysRevE.71.056708 PACS number(s): 02.50.−r, 02.70.−c, 05.10.−a, 05.40.−a

I. INTRODUCTION

Nonlinear time series analysis often use the method of surrogate data [1] for statistical significance testing of the results. A surrogate time series is a reordered sequence that preserves the statistical properties of the original data, such as empirical histograms and power spectra. A comparison of the statistical properties of the original and surrogate data sets reveals whether or not the original data have nonlinear properties. Moreover, it is also possible to determine the essential properties of the original data by varying the preserved statistical properties.

To generate the surrogate data sets, the Fourier transform is generally used for estimating the power spectra. In the calculation of the transform, because the original data are assumed to be periodic, gap effects always pose a problem. Thus, the calculation accuracy of the Fourier transform (and the power spectrum) is reduced. It is widely acknowledged that the multiplication of a data window for estimating the power spectra reduces such artifacts and thus counters this problem. Several reports give the effects of a data window [2] for evaluating the power spectra of the observed time series.

However, no systematic results are available on the relation between the data-window effects and significance testings by the method of surrogate data. We expected that the data window could reduce false rejections. However, several issues, such as which window has a better effect and what relation exists between the effect and the number of data points, have yet to be addressed.

In this study, we performed surrogate tests in which we used the autoregressive (AR) model as the original data. We also used nearly nonstationary AR models with shorter lenghts to consider the case wherein false rejections occur easily. In these cases, if a false rejection occurs, the surrogate test concludes that the AR model has a nonlinear property. This false conclusion is a serious problem in the field of nonlinear time series analysis. Next, we also examined which data window is suitable for generating the surrogate data sets in order to reduce such misleading false rejections. Moreover, we determined the limit on the temporal epoch of the original data for which the data windows are effective.

II. THE METHOD OF SURROGATE DATA

The method of surrogate data [1] is frequently used because it is useful for obtaining reliable results in nonlinear time series analysis. We must avoid the spurious identification of deterministic chaos underlying the time series data. It is well known that statistical control, such as the method of surrogate data, is important to avoid such careless estimation of nonlinear indices, for example, fractal dimensions [3] or Lyapunov exponents [4]. In chaotic time series analysis, the surrogate data sets are constructed to satisfy a null hypothesis based on the existence of a linear stochastic process. By comparing the statistics of the given data and the surrogate data sets, it is possible to reject the null hypothesis if the statistics of the original data differ significantly from those of the surrogates. In other words, if the given data do not possess any nonlinear properties, the calculated statistics should be almost equal to those of the surrogates. On the other hand, if the given data are truly nonlinear, the statistics are different from those of the surrogates.

In this study, we used three well-known algorithms for generating the surrogate data: the Fourier transform (FT) [5,6], Fourier shuffle (FS) [7,8], and iterative amplitude adjusted Fourier transform (IAAFT) [9].

A. FT surrogate

The FT surrogate algorithm completely preserves the power spectrum of the original data completely; however, the empirical histogram is destroyed. The algorithm for generating the FT surrogate is as follows.

(1) The Fourier transform is applied to the original data to obtain the power spectrum.

(2) Preserving the power spectrum, the phase of the spec-

trum is randomized by Gaussian random numbers, and the randomized spectrum is symmetrized to obtain a real-time series.

(3) The inverse Fourier transform of the data generated during the second step produces a FT surrogate that preserves the power spectrum of the original time series.

B. FS surrogate

Since the FT surrogate does not preserve the empirical histogram, a different algorithm for generating the surrogate data was proposed [7,8]. The FS surrogate preserves not only the power spectrum but also the empirical histogram of the original data. The algorithm for generating the FS surrogate is as follows.

(1) We generate the FT surrogate data of the original time series.

(2) We shuffle the original data such that the rank order is identical to that of the FT surrogate data. Here, the rank order is the order of the state values in the time series.

Although the FS surrogate preserves the empirical histogram completely, it cannot preserve the power spectrum, as the FT surrogate does.

C. IAAFT surrogate

The IAAFT surrogate preserves the power spectrum more accurately than the FS surrogate. Moreover, the IAAFT surrogate completely preserves the empirical histogram of the original time series. The algorithm for generating the IAAFT surrogate is as follows.

(1) We estimate the power spectrum of the original time series by the Fourier transform.

(2) We generate the FS surrogate data of the original time series.

(3) The Fourier transform is applied to the FS surrogate. Here, the power spectrum is replaced by that of the original time series estimated in the first step; however, the phase of the spectrum is preserved.

(4) The inverse Fourier transform is applied to the data obtained in the third step. Although the generated time series exhibits the same power spectrum as the original time series, its empirical histogram is different from that of the original.

(5) To preserve the empirical histogram of the original data, the original time series is reordered to obtain the same rank order as that of the time series generated by the fourth step. The reordered time series is the IAAFT surrogate.

(6) If the discrepancy between the power spectrum of the original and the IAAFT surrogate obtained in the previous step is not sufficiently small, we repeat the above steps by replacing the FS surrogate with the IAAFT surrogate in the third step. By repeating these steps several times, the power spectrum of the original data can be preserved more accurately than with the FS surrogate.

III. IMPROVING THE CALCULATION OF THE FOURIER TRANSFORM BY APPLYING THE DATA WINDOW

During the process that generates the surrogate data, we calculate the Fourier transforms. This calculation becomes fast and accurate if we use the fast Fourier transform (FFT) [10]. However, there exists a possibility that its accuracy will worsen. For example, if the sampled original data gives a noiseless sinusoidal curve, the data must have a single frequency component and the calculated power spectrum must have only a single peak. However, if the length of the sampled original data does not completely match the period, several peaks leak from the true peak, and these leakages adversely affect the estimation of the power spectrum. The reason is that in the calculation of the Fourier transform, the original data are assumed to be periodic. If the sampled original data do not have the same period as the observation epoch, gaps exist on both sides; consequently, the calculation accuracy of the power spectrum worsens.

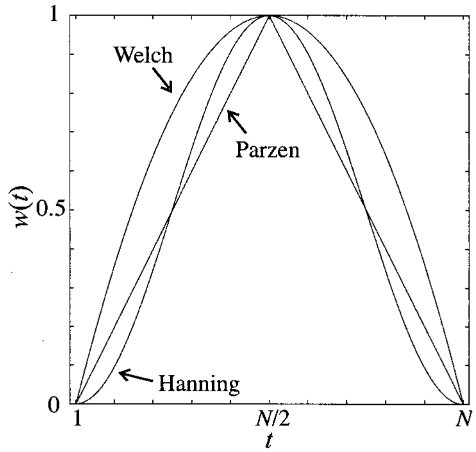

FIG. 1. Shapes of the data windows.

In most cases of real-time series analysis, the original data rarely have the same period as the observation epoch. Even if we use a part of the data by resampling several periods from the original data to reduce the gaps at both sides, it could eliminate essential information that the original data might contain.

Here, we introduced a transformation of the original data by data windows [10] to reduce both sides of the original data to zero or to reduce these gaps almost to zero. We used the data windows for generating each surrogate set and examined the usefulness of data windows for the surrogate tests [9].

For the calculation of the FFT in each process of generating the surrogate data, we used the following three data windows:

(1) the Parzen window

$$w_P(t) = 1 - \left| \frac{t - \frac{1}{2}(N+1)}{\frac{1}{2}(N-1)} \right|, \quad (1)$$

(2) the Hanning window

$$w_H(t) = \frac{1}{2}\left[1 - \cos\left(\frac{2\pi t}{N-1} \right) \right], \quad (2)$$

and

(3) the Welch window

$$w_W(t) = 1 - \left[\frac{t - \frac{1}{2}(N+1)}{\frac{1}{2}(N-1)} \right]^2, \tag{3}$$

where N is the data length. The shapes of these data windows are shown in Fig. 1.

Let us describe the discrete Fourier transform $\tilde{X}(k)$ of the original data $x(t)$ with a data window $w(t)$. Thus, we have

$$\begin{aligned}\tilde{X}(k+1) &= \sum_{t=0}^{N-1} x(t+1)w(t+1)e^{-j2\pi kt/N} \\ &= \sum_{t=0}^{N-1} \left[\frac{1}{N}\sum_{k'=0}^{N-1} X(k'+1)e^{j2\pi k't/N} \right] w(t+1)e^{-j2\pi kt/N} \\ &= \frac{1}{N}\sum_{k'=0}^{N-1} X(k'+1) \sum_{t=0}^{N-1} w(t+1)e^{-j2\pi(k-k')t/N} \\ &= \frac{1}{N}\sum_{k'=0}^{N-1} X(k'+1)W(k-k'+1),\end{aligned} \tag{4}$$

where $X(k+1)$ and $W(k)$ are the $(k+1)$th components of the Fourier transforms of $x(t)$ and $w(t)$, respectively; k is an index, $k=0,1,\ldots,N-1$; and j is the imaginary unit. As shown in Eq. (4), using data windows implies smoothing $X(k)$ by the weight $W(k)$. If the leakages of the power spectrum do not exist, using data windows would adversely affect the estimation accuracy of the power spectrum.

IV. NUMERICAL TESTS AND DISCUSSIONS OF FALSE REJECTION

In this section, to calculate the false rejection rates, we used the first-order autoregressive AR(1) model with $a=0.995$ as the original datum. AR(1) is defined by

$$x(t+1) = ax(t) + \eta(t), \tag{5}$$

where $\eta(t) \sim N(0,1), x(1) \sim N(0,1)$, and a is a parameter. When $a<1$, AR(1) is stationary and closely resembles white noise in the higher frequency regions. When $a=1$, AR(1) is not stationary; however, it is equivalent to the Brownian motion whose power spectrum is of the type $1/f^{\alpha}$. In this case, the rejection is true [11]. As shown in Fig. 2, we estimated the ensemble variance $\sigma^2_{\{x(t)\}}$ of $x(t)$ by varying the parameter a of AR(1). Because the condition of stationarity is that $\sigma^2_{\{x(t)\}}$ is temporally constant, we must extract a temporal subset of data for numerical experiments to generate artificial data that satisfies the condition of stationarity. In addition, the ensemble average of $x(t)$ must also be temporally constant. In the present paper, we used AR(1) with $a=0.995$, and we did not use the transient part of the data from $t=1$ to 1000 but the succeeding N points as data. Next, we varied the data length N: $N=64$, 128, 256, 512, 1024, 2048, and 4096. As the data length shortened, we assumed that the problem of the gap between the sampled data and its period becomes more serious, resulting in easy occurrence of a false rejection.

For each test, 19 surrogates were created by the FT, FS, and IAAFT algorithms with and without the data windows

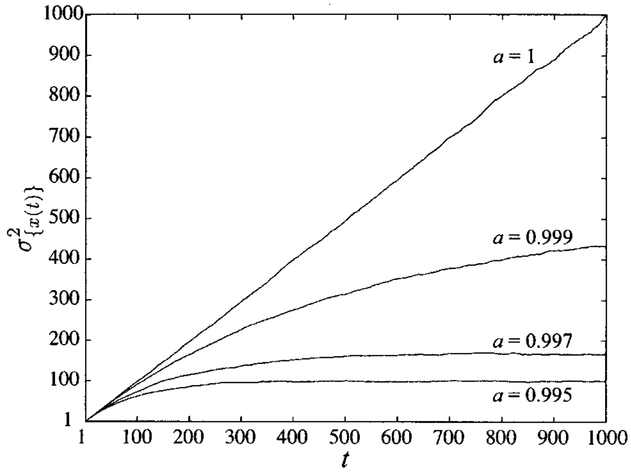

FIG. 2. Characteristics of the ensemble variance $\sigma^2_{\{x(t)\}}$ of the AR model $x(t)$ by varying the coefficient a of the AR model.

introduced in Sec. III. When we used a shorter data length, we were unable to reliably estimate the nonlinear indices such as the fractal dimensions and the Lyapunov exponents [12]. Instead, we used a local linear approximation prediction method [13,14] in a five-dimensional delay space [15,16] and the normalized mean square error [17] for estimating its prediction error. The null hypothesis is rejected at the 95% level of significance if the prediction error for the data is smaller than that of the 19 surrogates. The number of false rejections was estimated by performing 500 independent tests. Since the original data were generated by AR(1), any rejection is a false rejection.

Figure 3 shows the results of the false rejection rates with the FT surrogates. When data windows were not used (results are labeled as "No window"), the false rejection rates decreased as the data length N increased. The reason for this is as follows: If the number of sampled data increased, the data could contain several periods; thus, we required fewer data to arrange both sides of the sampled data. In other words, the sampling gap at both sides was reduced. Because the Fourier coefficients are decided by the inner products between the sampled data and the fundamental waves, the influence of the sampling gap was generally reduced when

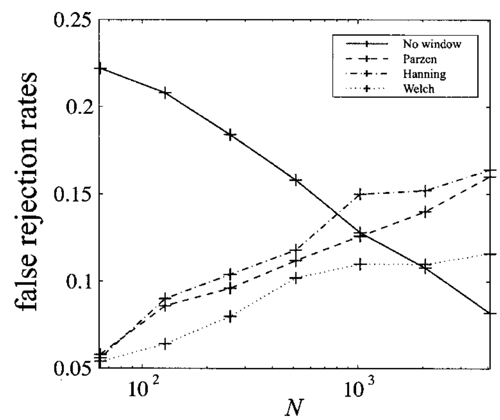

FIG. 3. Results of the false rejection rates with the FT surrogate data.

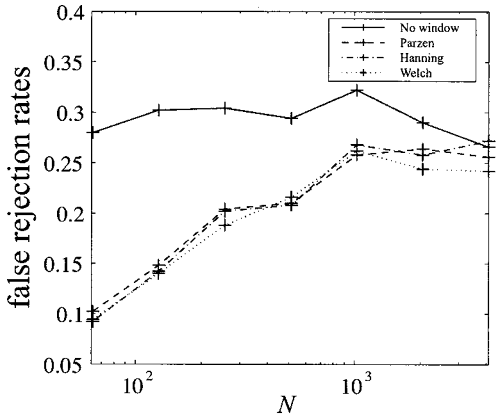

FIG. 4. Results of the false rejection rates with the FS surrogate data.

the sampled data included more periods. In the case of stationary data, the data have more periods as the data length is larger. Thus, because the data length N is large, the power spectrum is estimated more accurately and the false rejection rate is reduced. However, because we cannot reduce the influence of the sampling gap of the period, setting the sampling rate higher to increase the data length is futile.

Next, when data windows were used, they were very effective as the data length N was smaller. It should be noted that using data windows smoothes the power spectrum, as shown in Eq. (4). Thus, it is preferable to refrain from using the data windows in the case where N is large because the influence of the sampling gap is already smaller. Moreover, the wide variety of prediction accuracies of the surrogate sets is considered to be another reason for the reduction in false rejection rates at smaller values of N. Because the distribution of the statistics of the surrogates has a large tail, the rejection did not occur easily. However, in the case in which the data window was not used, the false rejection rate was very high. In other words, if N was small, the use of data windows afforded a considerable benefit. Moreover, when we used a data window with the FT surrogate, the Welch window was found to be the best choice.

Figure 4 shows the results of the false rejection rates with the FS surrogates. In all cases, the false rejection rates increased as compared to those observed in the case of the FT surrogates. This is because although the power spectrum was completely preserved by the FT surrogate, it was destroyed in the second step of generating the FS surrogate data. Further, along with the results of the FT surrogates, the false rejection rates were restricted by a trade-off between the sampling gap depending on N and the smoothing error of the data windows. When we used a data window with the FS surrogate, the Welch window was once more found to be the better choice.

Next, to modify the power spectrum destroyed by the FS surrogate, we used the IAAFT surrogate with and without the Welch window. We consider two methods for applying the data windows to the IAAFT surrogates as follows.

Method 1. We applied the Welch window in the first and second steps for generating the IAAFT surrogate.

Method 2. We applied the Welch window in the first, second, and third steps for generating the IAAFT surrogate.

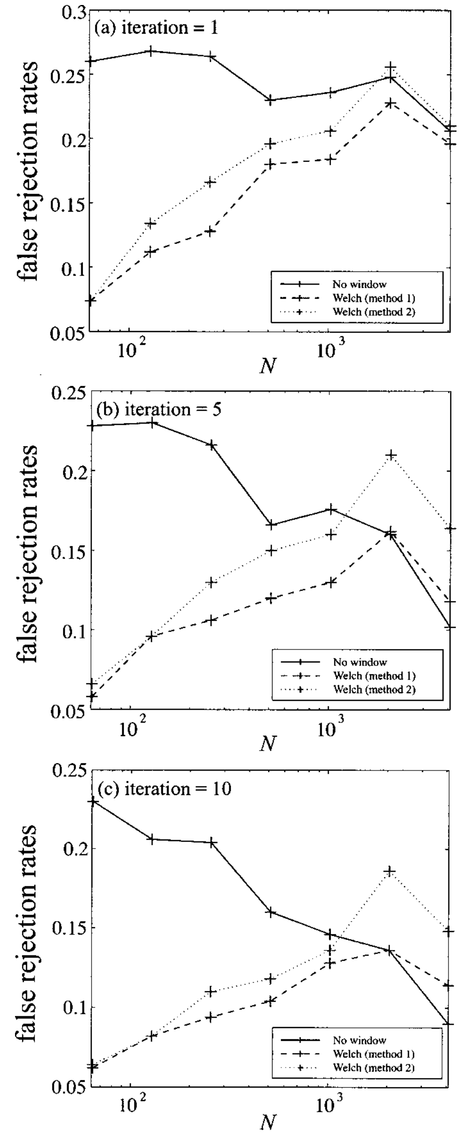

FIG. 5. Results of the false rejection rates with the IAAFT surrogate data.

We adopted method 2 because it is very important to investigate whether the use of data windows in the third step for generating the IAAFT surrogate improves the estimation accuracy of the power spectrum of the FS surrogate. Because both sides of the FS surrogate data were almost completely arranged, there existed a possibility that the smoothing error of the data windows adversely affected the estimation accuracy.

On the basis of Fig. 5 we confirm that method 1 was superior in almost all cases. In addition, method 1 was better than method 2 from the viewpoint of numerical costs. Further, we also confirmed that the IAAFT surrogates modified the power spectra destroyed by the FS surrogates by iterating the fifth step of the IAAFT surrogate generating algorithm since the false rejection rates were improved. Moreover, along with the results of the FT and FS surrogates, the false rejection rates were based on a trade-off between the sampling gap depending on N and the smoothing error of the data window.

On the basis of these numerical tests, we concluded that data windows, particularly the Welch window, reduced the false rejection rates in the case of a shorter data length, and the data windows were useful not only for the FT surrogate but also for the FS and IAAFT surrogates. In this paper, we tested for nonlinearity in the data by performing a local linear fit in a five-dimensional delay space. Even for clean chaotic data, such a fit would give a substantial amount of forecast errors if the data length were relatively shorter. In such a case, we cannot obtain any conclusive results. However, it is definitely important to test the reliability of methods for short time series. In our results, the applications of data windows remarkably reduced the false rejections of linear time series, even though there are limits of the applications.

Moreover, in the worst possible case wherein the linear memory of the original data is very long, as in the nonstationary case, the critical length of the original data for which the data windows were effective was approximately 1000.

V. CONCLUSIONS

In this paper, we examined the effects of data windows on the calculation of the Fourier transform for generating the FT, FS, and IAAFT surrogate data. We demonstrated that the use of data windows, particularly the Welch window, reduces the false rejection rates in the case of a shorter data length, and the data windows were useful not only for the FT surrogate but also for FS and IAAFT surrogates. We did not insist that the use of subsampled data was preferable for each surrogate test even if we applied a data window. By discussing the relation between the data length and the rejection ability of each surrogate test, we concluded that the false rejection rates were extremely low in the case of a shorter data length not only due to the effect of the data window but also due to the wide variety of prediction accuracies of the surrogate sets. Hence, we strongly recommend that data windows be used when the length of the sampled data is short.

In addition, when the data window is used, it is necessary to consider the trade-off between the sampling gap depending on the data length and the smoothing error of the data window. On the basis of our results, the balance point of the trade-off, i.e., the point at which the data windows were effective, corresponded to the point at which the original data length is approximately 1000 in the worst possible case wherein the linear memory of the original data is very long, as in the nonstationary case. If the linear memory is shorter, the data length at the critical point might be shorter since the effect of the sampling gap reduces. Thus, the trade-off between the sampling gap depending on N and the smoothing error of the data window may depend on the type of power spectrum of the original data. It is important to examine this trade-off in greater detail in future studies.

The research of T.I. was partially supported by Grant-in-Aids for Scientific Research (C), Grant No. 13831002 from the Japan Society for the Promotion of Science and for Scientific Research on Priority Areas, Grant No. 15014101 from the Ministry of Education, Culture, Sports, Science and Technology.

[1] J. Theiler *et al.*, Physica D **58**, 77 (1992).
[2] M. Hino, *Spectral Analysis* (Asakura-Shoten, Tokyo, 1977) (in Japanese).
[3] P. Grassberger and I. Procaccia, Physica D **9**, 189 (1983).
[4] J. P. Eckmann *et al.*, Phys. Rev. A **34**, 4971 (1986).
[5] D. T. Kaplan *et al.*, Circ. Res. **67**, 886 (1990).
[6] J. Theiler, Phys. Lett. A **155**, 480 (1991).
[7] T. Chang *et al.*, Biophys. J. **67**, 671 (1994).
[8] T. Chang *et al.*, Chaos **5**, 118 (1995).
[9] T. Schreiber and A. Schmitz, Phys. Rev. Lett. **77**, 635 (1996).
[10] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C* (Cambridge University Press, Cambridge, U.K., 1988).
[11] J. Timmer, Phys. Rev. E **58**, 5153 (1998).
[12] J. P. Eckmann and D. Ruelle, Physica D **56**, 185 (1992).
[13] E. N. Lorenz, J. Atmos. Sci. **26**, 636 (1969).
[14] G. Sugihara and R. M. May, Nature (London) **344**, 734 (1990).
[15] T. Sauer, J. A. Yorke, and M. Casdagli, J. Stat. Phys. **65**, 579 (1991).
[16] F. Takens, *Dynamical System and Turbulence*, edited by D. A. Rand and L. S. Young, Lecture Notes in Mathematics vol. 898, (Springer, Berlin, 1981), p. 366.
[17] T. Ikeguchi, T. Yamada, and M. Komuro, in *Fundamentals and Applications of Chaotic Time Series Analysis* edited by K. Aihara (Sangyo-Tosho, Tokyo, 2000) (in Japanese).