Wayland test, noise, and surrogates

Yoshito Hirata[1,*], Shunsuke Horai[2], Hideyuki Suzuki[1], and Kazuyuki Aihara[3,1,2]
[1]Department of Mathematical Informatics, The University of Tokyo,
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan
[2] Aihara Complexity Modelling Project, ERATO, JST,
45-18 Oyama, Shibuya-ku, Tokyo 151-0065, Japan
[3] Institute of Industrial Science, The University of Tokyo,
4-6-1 Komaba, Meguro-ku, Tokyo, 153-8505, Japan
[*] yoshito@sat.t.u-tokyo.ac.jp

15 October 2004

Abstract

We show analytically and numerically that Wayland test can exceed one when a time series is contaminated with observational noise. Therefore, the test is not suitable for detecting the determinism from noisy experimental data. We also report a weird phenomenon: When the Wayland test is applied to random shuffle surrogates, they could show more "determinism" than the original data.

1 Introduction

Nonlinear time series analysis has been intensively investigated for the last two decades. One of the main aims is to give evidence that a given time series is generated from deterministic chaos. There are some key features of deterministic chaos. One of them is determinism. For detecting the determinism from time series, there are several methods proposed [1, 2]. Among them, a simple and easy-to-use one is Wayland test [2]. In the original paper, it was shown using examples that the statistic becomes close to 0 when a time series is deterministic, and it becomes about 1 when it is not.

In this paper, we report that the statistic of the Wayland test can be greater than 1 if a time series, generated by a deterministic system, is contaminated with observational noise. In Section 2, we introduce the Wayland test [2]. In Section 3, we argue analytically that the statistic of the Wayland test increases if the level of observational noise increases. In Section 4, we demonstrate the above analysis using some numerical examples. When the Wayland statistic is used in surrogate tests, a weird thing happens; randomly shuffled surrogates could exhibit more determinism than the original data. We report an example of this phenomenon in Section 5. We conclude the paper in Section 6.

2 Wayland test

In this section, we summarize a method proposed by Wayland *et al.* [2] for detecting the determinism in a time series. Suppose that a scalar time series $\{s_t\}_{t=1}^N$ is given. Let τ be the time lag, and m, the embedding dimension. Then we form delay coordinates x_t by $(s_t, s_{t-\tau}, \cdots, s_{t-(m-1)\tau})^T$. We call $\{x_t\}_{t=(m-1)\tau+1}^N$ the experimental attractor.

Choose the number l of integers between $(m-1)\tau+1$ and $N-1$. Call this set of integers T. For x_t of $t \in T$, find the number k of nearest neighbors from the experimental attractor. We label the i-th nearest neighbor for x_t as $x_{n_i(t)}$. For notational convenience, we define $n_0(t) = t$. For each $x_{n_i(t)}$, we look at its image $x_{n_i(t)+1}$ and take the translation vector,

$$v_i(t) = x_{n_i(t)+1} - x_{n_i(t)}. \quad (1)$$

To quantify this notion, let

$$v(t) = \frac{1}{k+1}\sum_{i=0}^{k} v_i(t). \quad (2)$$

This $v(t)$ shows the average of the translation vectors $v_i(t)$. Using it, we define the translation error $e(t)$ as

$$e(t) = \frac{1}{k+1}\sum_{i=0}^{k} \frac{\|v_i(t) - v(t)\|^2}{\|v(t)\|^2}. \quad (3)$$

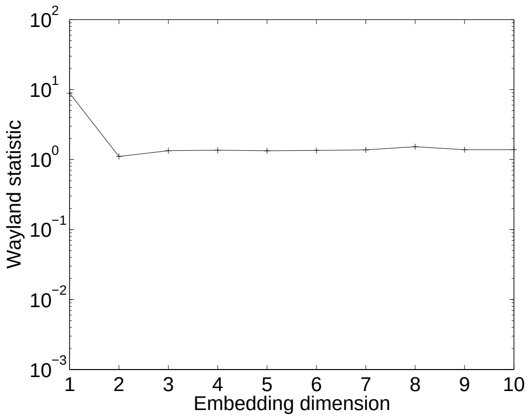

Figure 1: The result of the Wayland test, applied to a scalar time series of the Lorenz model.

The translation error $e(t)$ shows the fractional spread in the displacements of $x_{n_i(t)}$ relative to the average displacement $v(t)$.

We find the median of $e(t)$ over T, and declare it as the test statistic for the determinism. Wayland *et al.* showed using examples that the test statistic is close to 0 if a time series is deterministic and the test statistic is about 1 if a time series is random.

We applied the Wayland test to a time series generated by the Lorenz model [3]. The Lorenz model is defined as

$$\left\{ \begin{array}{rcl} \frac{dx}{dt} & = & -ax + ay \\ \frac{dy}{dt} & = & -xz + bx - y \\ \frac{dz}{dt} & = & xy - cz, \end{array} \right. \tag{4}$$

where $(a, b, c) = (10, 28, 8/3)$. We integrated the equation using ode45 of MATLAB. During integrating, we observed x-coordinate every 0.01 unit time and obtained a scalar time series of length 1000. Then we added to the data observational noise, which follows the Gaussian distribution $N(0, (0.05\sigma_o)^2)$ of the mean 0 and standard deviation $0.05\sigma_o$, where σ_o is the standard deviation of the original clean data.

The result is shown in Fig. 1. We used the parameters $\tau = 16$, $k = 4$, and $l = 100$. As the statistic exceeds 1 for all the tested dimensions, the time series should be interpreted as not deterministic. Although the Lorenz model itself is deterministic, the Wayland test did not detect its determinism from the noisy time series.

We remark that Wayland *et al.* [2] defined $\{s_t\}_{t=1}^N$ deterministic if x_t can be accurately modeled as the iteration of some continuous function. Therefore, according to this definition, the noisy time series tested here is not deterministic.

3 Analytical reasoning that Wayland statistic can be bigger than one

Why does the statistic of the Wayland test exceed 1? We argue in this section that there is a scaling law between the Wayland statistic and the level of observation noise.

Suppose that $\{y_t\}_{t=1}^N$ is a clean deterministic scalar time series and that $\{y_t\}_{t=1}^N$ is contaminated by observational noise $\{\eta_t\}_{t=1}^N$. Therefore, now the observed time series $\{s_t\}_{t=1}^N$ has the relation $s_t = y_t + \eta_t$ for each t. We assume that for each t, the noise η_t follows the Gaussian distribution $N(0, \sigma^2)$ of mean 0 and standard deviation σ.

Let

$$\tilde{v}_i(t) = \begin{pmatrix} y_{n_i(t)+1} - y_{n_i(t)} \\ y_{n_i(t)+1-\tau} - y_{n_i(t)-\tau} \\ \vdots \\ y_{n_i(t)+1-(m-1)\tau} - y_{n_i(t)-(m-1)\tau} \end{pmatrix}, \tag{5}$$

$$\tilde{v}(t) = \frac{1}{k+1} \sum_{i=0}^{k} \tilde{v}_i(t). \tag{6}$$

Then

$$v_i(t) = \tilde{v}_i(t) + \begin{pmatrix} \eta_{n_i(t)+1} - \eta_{n_i(t)} \\ \eta_{n_i(t)+1-\tau} - \eta_{n_i(t)-\tau} \\ \vdots \\ \eta_{n_i(t)+1-(m-1)\tau} - \eta_{n_i(t)-(m-1)\tau} \end{pmatrix}. \tag{7}$$

Hence, the average translation $v(t)$ is written as

$$\tilde{v}(t) + \begin{pmatrix} \frac{\sum_{i=0}^{k} (\eta_{n_i(t)+1} - \eta_{n_i(t)})}{k+1} \\ \frac{\sum_{i=0}^{k} (\eta_{n_i(t)+1-\tau} - \eta_{n_i(t)-\tau})}{k+1} \\ \vdots \\ \frac{\sum_{i=0}^{k} (\eta_{n_i(t)+1-(m-1)\tau} - \eta_{n_i(t)-(m-1)\tau})}{k+1} \end{pmatrix} \tag{8}$$

Assume that each $n_i(t)$ is far from the others in time. Then, for each j, the sum $\sum_{i=0}^{k}(\eta_{n_i(t)+1-j\tau} - \eta_{n_i(t)-j\tau})$ follows Gaussian distribution $N(0, 2(k+1)\sigma^2)$. Therefore $\frac{1}{k+1}\sum_{i=0}^{k}(\eta_{n_i(t)+1-j\tau} - \eta_{n_i(t)-j\tau})$ follows $N(0, \frac{2\sigma^2}{k+1})$. Because $2\sigma^2/(k+1)$ is small if k is big, we approximate it as

$$\frac{1}{k+1} \sum_{i=0}^{k} (\eta_{n_i(t)+1-j\tau} - \eta_{n_i(t)-j\tau}) \approx 0. \tag{9}$$

As a result, we have $v(t) \approx \tilde{v}(t)$.

Next we calculate $e(t)$. From its definition, we have

$$\begin{aligned} & e(t) \\ = \; & \frac{1}{k+1}\sum_{i=0}^{k}\frac{\|v_i(t)-v(t)\|^2}{\|v(t)\|^2} \\ = \; & \frac{1}{(k+1)\|v(t)\|^2}\sum_{i=0}^{k}\sum_{j=0}^{m-1}\left[([\tilde{v}_i(t)]_j-[v(t)]_j)^2\right. \\ & +2(\eta_{n_i(t)+1-j\tau}-\eta_{n_i(t)-j\tau})([\tilde{v}_i(t)]_j-[v(t)]_j) \\ & \left.+(\eta_{n_i(t)+1-j\tau}-\eta_{n_i(t)-j\tau})^2\right]. \end{aligned}$$

Taking the mean of $e(t)$ over all η_t's, we have

$$E[e(t)] \;=\; \frac{\sum_{i=0}^{k}(\|\tilde{v}_i(t)-v(t)\|^2+2m\sigma^2)}{(k+1)\|v(t)\|^2}.$$

Letting

$$\tilde{e}(t)=\frac{1}{k+1}\sum_{i=0}^{k}\frac{\|\tilde{v}_i(t)-\tilde{v}(t)\|^2}{\|\tilde{v}(t)\|^2}, \qquad (10)$$

we obtain

$$E[e(t)] \approx \tilde{e}(t)+\frac{2m\sigma^2}{\|v(t)\|^2}. \qquad (11)$$

The quantity $\tilde{e}(t)$ can be different from the statistic we obtain from a clean time series as the selected nearest neighbors can be different because of the observation noise. However, if the noise level is moderate, the selected points are still the neighbors of x_t, and the value of $\tilde{e}(t)$ is close to the one we obtain from a clean time series. Hence, we regard that $\tilde{e}(t)$ is approximately equal to the statistic we obtain from a clean time series as far as the noise level is moderate.

As $\frac{m}{\|v(t)\|^2} > 0$, the average of $e(t)$ will increase if the variation of the noise increases.

4 Numerical examples

To verify the formula in Eq. (11) numerically, we used a time series generated from the Hénon map [5]. The Hénon map is defined as $(u_{t+1}, v_{t+1}) = (1-a{u_t}^2+bv_t, u_t)$ where $(a,b) = (1.4, 0.3)$. We observed u_t and obtained a scalar time series of length 1000.

We tested the formula with several noise levels: $\sigma = 0, 0.01\sigma_o, 0.02\sigma_o, \cdots, 0.25\sigma_o$ where σ_o is the standard deviation of the original time series. For each noise level, we generated 100 realizations of noisy data and found the Wayland statistic for each of them. For calculating the Wayland statistic, we used $m = 2$, $\tau = 1$, $k = 4$ and $l = 100$. The result is shown in Fig. 2. We

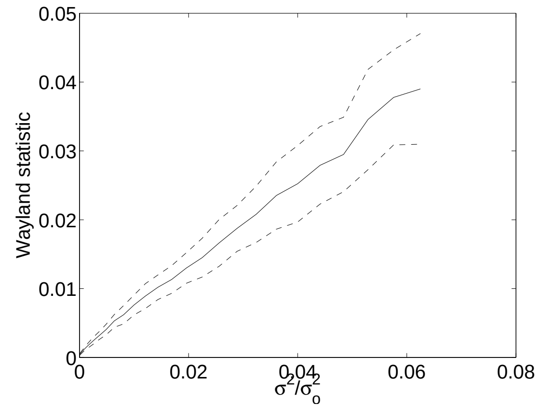

Figure 2: The relation between the Wayland statistic and the noise level σ in the example of the Hénon map. The solid line is the mean, the broken lines are the mean $\pm$ the standard deviation of 100 realizations.

observed the linear relationship between the Wayland statistic and σ^2, which was expected from Eq. (11).

We also tested this relation using the data of the Lorenz model. We observed the x-coordinate of the Lorenz model every 0.01 unit time and obtained a time series of length 1000. The tested noise levels σ are $0, 0.01\sigma_o, 0.02\sigma_o, \cdots, 1\sigma_0$ where σ_o is the standard deviation of the original clean time series. For finding the Wayland statistic, we used $m = 3$, $\tau = 16$, $k = 4$ and $l = 100$. The lag τ was chosen such that τ gives the first minimum of the mutual information [4]. The result is shown in Fig. 3. In this example, we also observed the linear relationship between the Wayland statistic and σ^2 when σ^2 is small. When the noise level is high, the average of the Wayland statistic is greater than one.

5 Wayland statistic and surrogates

When the Wayland statistic is used for a time series contaminated by observation noise with surrogate data analysis, we might observe a strange phenomenon: randomly shuffled surrogates show more "determinism" than the original time series.

We applied surrogate data analysis to the time series of the Lorenz model which is contaminated with 5% observation noise. The result is shown in Fig. 4. We observed that in most embedding dimensions, the Wayland statistic for the original time series is greater than the maximum of 39 randomly shuffled surrogates.

The following two reasons explain the phenomenon: When applied to a surrogate, or a dataset without any

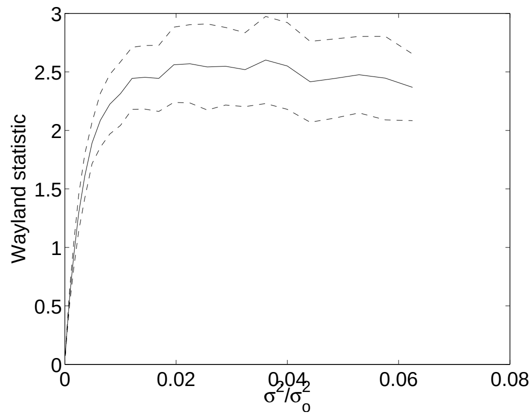

Figure 3: The relation between the Wayland statistic and the noise level σ of the observation noise, in the example of the Lorenz model. The solid line is the mean, the broken lines are the mean ± the standard deviation of 100 realizations.

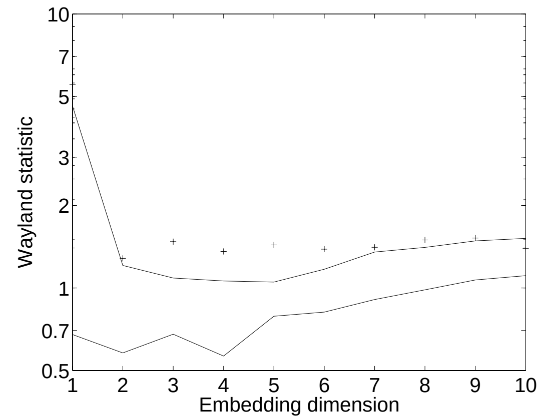

Figure 4: The result of surrogate data analysis on the data of the Lorenz model, contaminated with 5% observational noise. We generated 39 randomly shuffled surrogates from the original time series. For each embedding dimension, the sign + shows the value of the Wayland statistic obtained for the original noisy data, and the solid lines show the minimum and maximum of the Wayland statistic obtained from 39 surrogates.

temporal correlation, the Wayland statistic tends to be close to 1 [2]. However, because of the scaling law between the Wayland statistic and the noise level, the original data, or a dataset contaminated by observational noise, can yield a value greater than 1.

6 Conclusion

We showed that Wayland test is not appropriate for the test of the determinism in noisy experimental data because the statistic of the test can be greater than 1 even if there is a deterministic law behind the dynamics of a time series. The reason is that there is a scaling law between the statistic of the Wayland test and the level of observational noise. Although the original paper contains the analysis of observational noise, the result is too optimistic to be applied to noisy experimental data.

Because of the scaling law, when the statistic of the Wayland test is used for the test statistic of a surrogate data test, randomly shuffled surrogates can demonstrate more determinism than the original time series.

Acknowledgment

This study was supported by Industrial Technology Research Grant Program in 2003 from New Energy and Industrial Technology Development Organization (NEDO) of Japan.

References

[1] D. T. Kaplan and L. Glass, "Direct test for determinism in a time series," *Phys. Rev. Lett.*, vol.68, pp.427–430, 1992; D. T. Kaplan and L. Glass, "Coarse-grained embeddings of time series: random walks, Gaussian random processes, and deterministic chaos," *Physica D*, vol.64 pp.431–454, 1993; J. Bhattacharya and P. P. Kanjilal, "On the detection of determinism in a time series," *Physica D*, vol.132, pp.100–110, 1999; S. Horai, T. Yamada, and K. Aihara, "Determinism analysis with iso-directional recurrence plots," *T. IEE Japan*, vol.122, pp.141–147, 2002 (in Japanese).

[2] R. Wayland, D. Bromley, D. Dickett, and A. Passamante, "Recognizing determinism in a time series," *Phys. Rev. Lett.*, vol.70, pp.580–582, 1993.

[3] E. N. Lorenz, "Deterministic non-periodic flow," *J. Atmos. Sci.*, vol.20, pp.130–141, 1963.

[4] A. M. Fraser and H. L. Swinney, "Independent coordinates for strange attractors from mutual information," *Phys. Rev. A*, vol.33, pp.1134–1140, 1986.

[5] M. Hénon, "A two-dimensional mapping with a strange attractor," *Commun. Math. Phys.*, vol. 50, pp.69–77, 1976.