Applications of Double-Wayland Algorithm to Detect Anomalous Signals

Hiroki TAKADA[1]*, Takayuki MORIMOTO[1], Hitoshi TSUNASHIMA[2], Taishi YAMAZAKI[2], Hiroyuki HOSHINA[3] and Masaru MIYAO[4]

[1]*Department of Computational Science and Engineering, Nagoya University, Chikusa-ku, Nagoya 464-8603, Japan*
[2]*College of Industrial Technology, Nihon University, 1-2-1 Izumicho, Narashino-shi, Chiba 275-8575, Japan*
[3]*Risk & Profit Analysis, AXA Life Insurance Co. Ltd., NBF Platinum Tower, 1-17-3 Shirokane, Minato-ku, Tokyo 108-8020, Japan*
[4]*IT Center, Nagoya University, Chikusa-ku, Nagoya 464-8601, Japan*
**E-mail address: takada@u-gifu-ms.ac.jp*

(Received March 3, 2006; Accepted May 22, 2006)

Keywords: Translation Error, Embedding Space Corrugation, Price Earning Ratio (PER), Economic Shocks

Abstract. The Wayland algorithm has been improved in order to evaluate the degree of visible determinism for dynamics that generate a time series in a simple and accurate manner. Additionally, the Double-Wayland algorithm that we proposed can detect phase transitions among multi-states and non-stationarity in the dynamics. We are applying the Double-Wayland algorithm to detect anomalous signals in railways, stock prices, stabilometry and electrograms recorded by using mapping catheters. In this study, we reported the manner in which these anomalous signals can be detected; however, due to space limitations, we have not reported this data for applications in the field of medicine.

1. Introduction

It is known that attractors can be reconstructed by Dynamical equation systems (DESs) such as the Duffing equation, Henon map and Lorenz differential equation. It is very interesting to note that the structure of an attractor is also derived from time series data in an embedding space. The form of the structure is regarded as a subject of mathematical science and can be measured by using the Double-Wayland algorithm proposed by the present authors (TAKADA *et al*., 2005a).

The DESs were obtained as mathematical models that regenerated time series data. Anomalous signals can either be given by the degeneration of the potential function in the DESs or they essentially change, for instance, their degree of freedom increases or stochastic factors are added to them. The visible determinism in the latter case would be different from that in the case where random variables do not exist. There is no unified view of calculus to distinguish between deterministic and stochastic processes; however, the

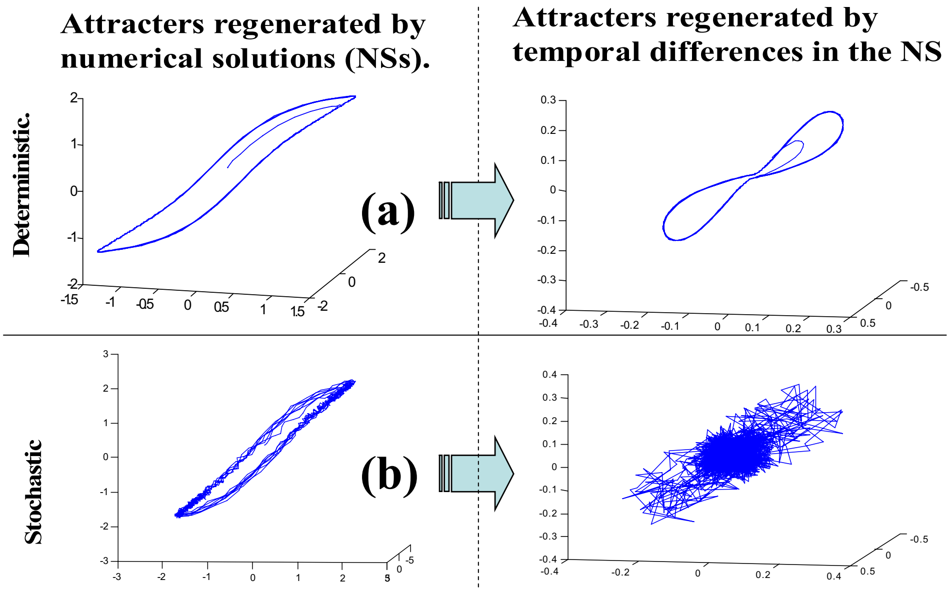

Fig. 1. Attracters regenerated by numerical solutions (NSs) to van der Pol equations without (a) and with random noise terms (b).

Grassberger-Procaccia and Wayland algorithms (WAYLAND *et al*., 1993) might be used as mathematical methods to evaluate their degree of freedom. The dimension of the DES can be estimated as a fractional number by using the Grassberger-Procaccia algorithm, which is considered to be beneficial for ensuring accuracy. However, it is difficult to examine whether a stochastic process is suitable for a mathematical model of time series data. It is not easy to complete the former computation because the fractal dimension of the DES is derived from the calculations of all the points in the embedding space. In contrast, visible determinism can be estimated statistically in the case of the latter algorithm. The use of statistics shortens the computation time. It is well known that 0.5 is the empirical threshold of the translation error to classify mathematical models as deterministic and stochastic generators; however, the translation error was generally not estimated at the same value as that in the case of a larger signal to noise (S/N) ratio (MATSUMOTO *et al*., 2002).

The authors compared the various translation errors involved in the time series, along with their differences (TAKADA *et al*., 2005a). It was supposed that this Double-Wayland algorithm can also detect phase transitions among multi-states and non-stationarity in the dynamics (TAKADA, 2005). The algorithm can be applied to various fields if the translation errors are estimated at the same value as that in the case of a larger S/N ratio. The translation error in the numerical solution to a one-dimensional Langevin equation $\dot{x} = f(x) + gw(t)$ was compared with those with other noise amplitudes . Consequently, the translation error estimated by the Double-Wayland algorithm did not depend on the noise amplitudes for $g < 6$ (TAKADA *et al*., 2005b).

Fig. 2. Corrugation.

In this research, the Double-Wayland algorithm was applied to the following engineering fields in order to detect anomalous signals:

1) Mechanical Engineering: Time series data of accelerations were measured for common vehicles, and we examined whether this algorithm could detect defects (corrugation) in railway tracks.

2) Financial Engineering: Time series data of price earning ratio (PER) were derived from the stock average, and we examined whether this algorithm could detect heavy falls or steep rises in prices.

2. Method

Delay coordinates $\{ \vec{x}_t \}$ can reconstruct a continuous trajectory without crossings in an embedding space that has a high dimension. The Wayland algorithm (Appendix) supposes that the difference vectors $\mathbf{v}_t = \vec{x}_{t+\tau} - \vec{x}_t$ in this space are approximated to temporal variations of the trajectories (Fig. 1) and estimates the translation error in an m-dimensional embedding space ($m = 1, 2, ..., 10$). This translation error is a statistical index that measures the complexity of the dynamics generating the time series. In addition, the randomness can be evaluated by the Double-Wayland algorithm (Appendix) by a comparison of the transition errors in the temporal differences of the time series with the results of the Wayland algorithm in each embedding space. Several applications of the Double-Wayland algorithm are presented in order to demonstrate its usefulness in this research.

1) There are several types of defects, such as irregularity of the cross level, alignment and levelling defects, improper distance between two rails, twist and corrugation (Fig. 2). Railway accidents might occur due to these defects, which can be identified either by conducting visual checks or by using expensive detectors in track inspection cars (see

Fig. 3. A track inspection car in Japan.

Fig. 3). We examine whether corrugation can be easily detected based on the measurements of the acceleration of common vehicles and apply the Double-Wayland algorithm to the time series data of the accelerations.

2) We examine whether economic shocks in the process of determining stock prices can be said to be a phase transition, although some statistical tests such as the CHOW test (CHOW, 1960) and the CUSUM test (BROWN *et al*., 1975) have been proposed to identify them. We analyse the Nikkei stock average y_t, which is the average stock price of 225 companies. This stock average can be monitored on the Web (Yahoo Finance[#]). The PER is defined as

$$x_t = \log\left(y_{t+1}/y_t\right).$$

The translation errors in the two intervals of the PER were evaluated by using the Double-Wayland algorithm. The first interval was set to be from 23 October 1997 to 24 October 2001, and the second interval was fixed as the period between 25 October 2001 and 21 November 2005. In the former interval, the World Trade Center was destroyed by terrorists (11 September 2001).

3. Results

1) We observed the horizontal/vertical accelerations of a common vehicle. Figure 4 shows the vertical acceleration that was measured during business hours. One of these

[#]http://table.yahoo.co.jp/t

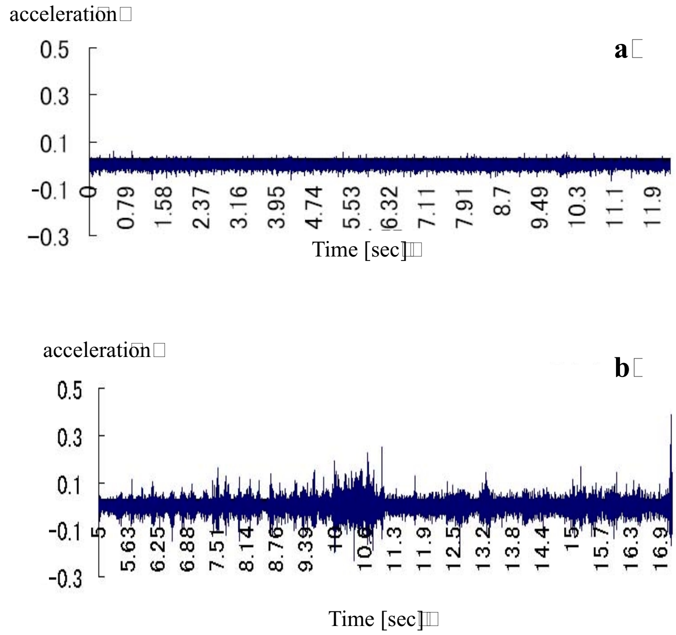

Fig. 4. Typical time series of vertical acceleration. The acceleration was observed in the normal section (a) and the anomalous section (b).

accelerations was recorded in the operating section; the visual checks revealed no corrugation (Fig. 4a). On the other hand, the corrugation was confirmed in another section (Fig. 4b). In this paper, the former section is termed as the normal section, while the latter is referred to as the anomalous section. The Double-Wayland algorithm was employed in every interval for 1.5 seconds; the total number of sequences in each interval was 3000. The translation errors in the normal section were estimated to be approximately 0.5 (Fig. 5). The translation error displayed a weak dependence on the embedding dimension and decreased with leniency (Fig. 5a). Moreover, these values were lower than any translation error values derived from the temporal differences of the time series in the accelerations. However, the order was reversed in the anomalous section (Fig. 5b).

2) The PER was measured from 7 January 1991 to 21 November 2005 (Fig. 6), and it included the two intervals mentioned above. In the first interval, there was a jump that appears to have been caused by the destruction of the World Trade Center. This could be regarded as an anomalous signal (see Fig. 6). According to the Double-Wayland algorithm, the translation errors abruptly decreased and had a minimal value in the three-dimensional embedding space (Fig. 7). The values of the translation errors derived from the time series of the PER were higher than those of their differences. The order was preserved despite the anomalous signal. However, the jump tended to lower translation errors in some embedding spaces (*$p < 0.05$, **$p < 0.01$).

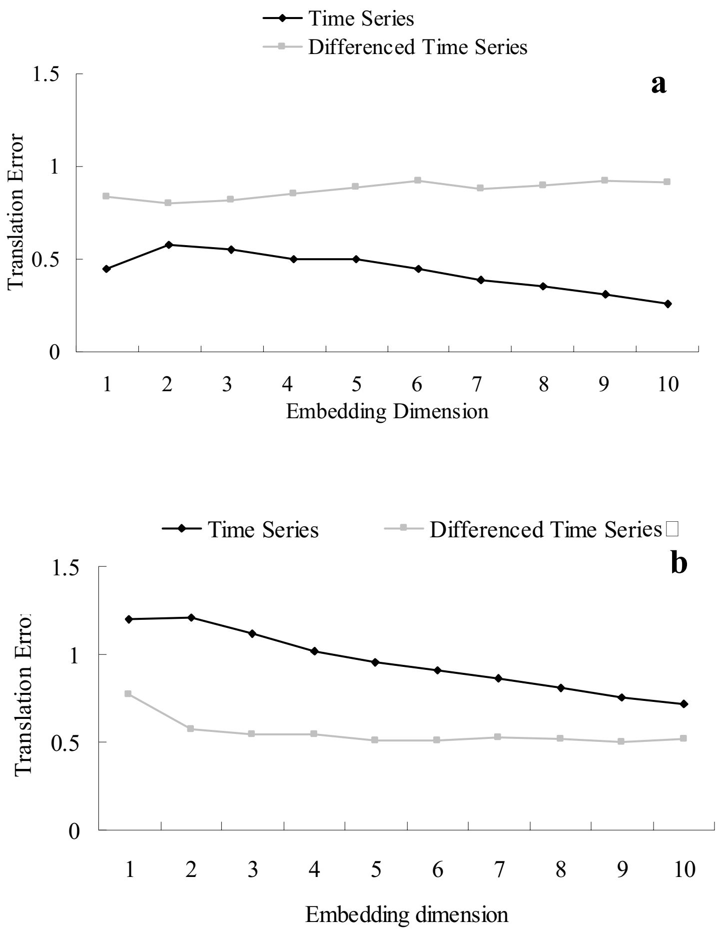

Fig. 5. Translation errors were derived from the time series of the acceleration in Fig. 4 for each embedding space. The accelerations were observed in the normal section (a) and the anomalous section (b).

4. Discussion

1) In the Double-Wayland algorithm, the process that generated the accelerations was generally regarded to be stochastic. Moreover, with regard to the anomalous section, the mathematical model for the accelerations was transformed into a deterministic one (Fig. 5). The regularity in acceleration can be physically enhanced by carrying out periodic structural changes in the railway tracks (Fig. 2). Thus, we recognized the order in the translation errors derived from the time series of the horizontal/vertical accelerations and the temporal differences. We successfully detected the anomalous sections in which corrugation was observed by conducting visual check. These sections were also detectable

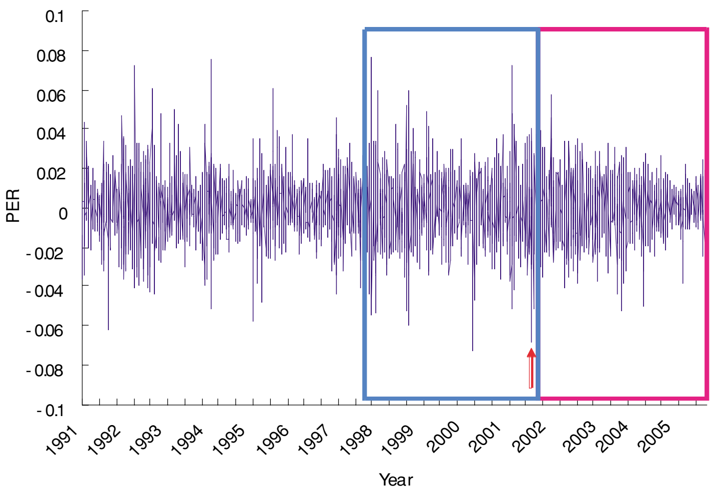

Fig. 6. Time series data of the PER in recent years. The number of sequences in each interval was 10^3.

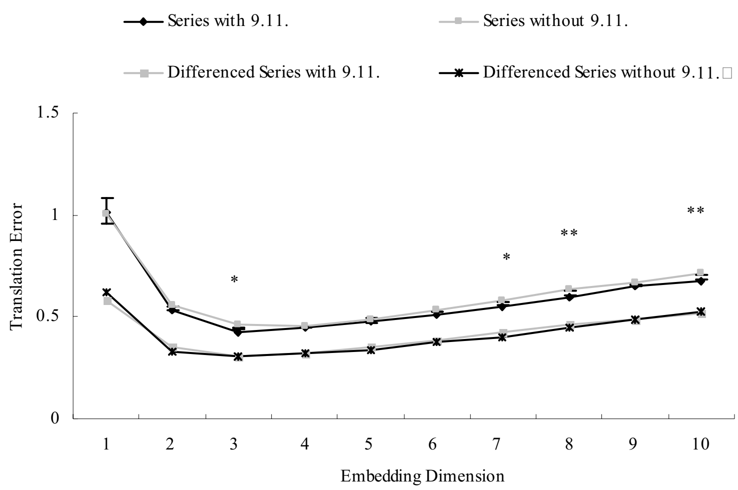

Fig. 7. Translation errors derived from the time series of the PER in Fig. 6.

by a complex multiresolution analysis, which indicated that the time series of the horizontal/vertical accelerations contained a certain frequency band (MATSUMOTO *et al.*, 2005).

This result suggests the applicability of a mathematical approach to the real-time detection of defects in railway tracks. The identification of defects is generally carried out by conducting visual checks or by using expensive detectors in track inspection cars (Fig.

3). However, due to hectic train schedules, it is not always possible to carry out this function. The proposed method of detecting defects based on the measurements of the accelerations of common vehicles is considered to be convenient since it enables the identification of defects irrespective of the situation.

2) Social events can be considered to have certain psychological effects on investors. Therefore, these events could act as a trigger in a market and result in drastic changes in stock prices. In the Double-Wayland algorithm, a DES without random variables was regarded as a generator of the PER, and the result did not depend on the intervals. There might be jumps that act as anomalous signals in both the intervals. Moreover, the minimal point shown in Fig. 7 was considered to provide information regarding the optimal embedding space in the DES as the mathematical model of the PER. Hence, three-dimensional differential equations were able to generate time series data regardless of the existence of the jump.

Currently, we are applying the Double-Wayland algorithm to the field of medicine. Anomalous signals might be produced in the anterior/posterior components of the body sway due to Meniere's disease. The Double-Wayland algorithm can be used to identify the sites of high-frequency activity that maintain atrial fibrillation in humans and evaluate the effect of ablation at these sites. These applications will be reported in the next paper.

A part of this study was supported by the 21st Century Center of Excellence (COE) Program for Frontiers of Computational Science, the Japan Science and Technology Agency (JST), and the Program for Promoting Fundamental Transport Technology Research from the Japan Railway Construction, Transport and Technology Agency (JRTT).

Appendix

The Wayland algorithm is composed of the following steps for each dimensional embedding space (WAYLAND *et al.*, 1993):

(i) A series of delay coordinate vectors $\{\vec{x}_t\}$ is embedded in each space.

(ii) M onset periods t_0 are chosen at random.

(iii) The values of

$$E_{\mathrm{trans}}(t_0) = \frac{1}{K+1}\sum_{i=0}^{K}\frac{|\mathbf{v}(t_i) - \bar{\mathbf{v}}|}{|\bar{\mathbf{v}}|} \tag{A1}$$

are standardized by the average of the difference vectors at $K + 1$ points $\{\vec{x}_{t_i}\}_{i=0}^{K}$.

$$\bar{\mathbf{v}} = \frac{1}{K+1}\sum_{i=0}^{K}\mathbf{v}(t_i) \tag{A2}$$

is obtained at every onset period, where K points nearest to $\vec{x}_{t_0}$ are selected as $\{\vec{x}_{t_i}\}_{i=1}^{K}$.

(iv) The median of the *M* values of Eq. (A1) is extracted.
(v) *Q* medians are obtained by repeating the above steps. The translation error is estimated by the expectation value of *Q* medians.
The Double-Wayland algorithm includes the following additional steps.
(vi) Translation errors are derived from the temporal differences of the time series by the Wayland algorithm outlined above.
(vii) If a DES that includes stochastic factors was the generator of the time series, the flow would not be smooth (Fig. 1). In such a case, a significantly higher number of translation errors would be estimated in the last step than in step (v). Based on this result, we decide the stochasticity of the mathematical model that describes the time series data.

REFERENCES

BROWN, R. L., Durbin, J. and Evans, J. M. (1975) Techniques for testing the constancy of regression relationships over time, Journal of the Royal Statistical Society, Series B, **37**, 149–172.

CHOW, G. C. (1960) Tests of equality between sets of coefficients in two linear regressions, *Econometrica*, **28**, 591–605.

MATSUMOTO, T., TOKUNAGA, R., MIYANO, T. and TOKUDA, I. (2002) *Chaos and Time Series*, Baihukan, Tokyo (in Japanese).

MATSUMOTO, A., OGATA, S., SATOU, Y., OUNO, H., TSUNASHIMA, H. and KOJIMA, T. (2005) The study of prove vehicles in railways (1), *Abstracts of National Traffic Safety and Environment Laboratory Annual Meeting in 2005*, 163–168 (in Japanese).

TAKADA, H. (2005) Evaluation in non-stationary processes by double-wayland test, *Bulletin of Society for Science on Form*, **20**, No. 1, 135–136.

TAKADA, H., SHIMIZU, Y., HOSHINA, H. and SHIOZAWA T. (2005a) Wayland tests for differenced time series could evaluate degrees of visible determinism, *Bulletin of Society for Science on Form*, **17**, No. 3, 301–310.

TAKADA, H., HOSHINA, H., TSUNASHIMA, H. and MIYAO, M. (2005b) Does translation error estimated by double-wayland test depend upon noise-amplitude?, in *Proceeding of Computational Science Symposium: Computational Science on Interaction*, ISPJ Symposium Series 2005 No. 11, pp. 57–64 (in Japanese).

WAYLAND, R., BROMLEY, D., PICKETT, D. and PASSAMANTE, A. (1993) Recognizing determinism in a time series, *Phys. Rev. Lett.*, **70**, 530–582.