Detection of changes in non-linear dynamics for time series based on the theory of KM_2O -Langevin equations

Tetsuji Hidaka · Yasunori Okabe

Received: 16 August 2011 / Revised: 1 December 2011 / Published online: 28 December 2011
© The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract This paper presents a framework for detecting changes in non-linear dynamics for time series based on the theory of KM_2O -Langevin equations. This paper has two contributions one is to define the pseudo determinacy function that deduces the non-linear dynamics behind the subsequences of the time series data the other is to define the distance of dynamics between the subsequences by the pseudo determinacy function to understand how and when the non-linear dynamics changes. In contrast to the existing methods such as Singular-Spectrum Analysis and Recurrence Quantification Analysis the proposed method can distinguish the difference between a temporal outlier and a change in non-linear dynamics. Moreover, applying the proposed method to the real data, we found the indirect relation between the change in non-linear dynamics and some outer force (the actress in TV commercial or attack by competitors).

Keywords Change point detection · Non-linear dynamics · Time series · Self-organizing maps · The theory of KM_2O -Langevin equations

Mathematics Subject Classification (2000) 37M10 · 62M10 · 91B84

In 2011 Y. Okabe died prematurely. He was an excellent mathematician and he gave a substantial contribution to the present paper. This work is supported by the Meiji University Global COE Program “Formation and Development of Mathematical Sciences Based on Modeling and Analysis”.

T. Hidaka (✉)

Hakuhodo Inc., Akasaka Biz Tower, 5-3-1 Akasaka, Minatoku 107-6322, Japan
e-mail: TETSUJI.HIDAKA@hakuhodo.co.jp

T. Hidaka · Y. Okabe

Graduate School of Advanced Mathematical Sciences, Meiji University,
1-1-1 Higashimita, Tamaku, Kawasaki 214-8571, Japan

1 Introduction

Industrial firms struggle for making their sales performance better and struggle against competitors using various marketing efforts. They check the effects of the effort whether and how the sales have changed. Scientists search for interesting patterns in data observed in experiments or the environment. In these areas, the most common data form observed is time series. One of the important problems related to these issues is to detect a change point in characteristics of time series. This trend of research is often called “change-point detection”.

There are several reasons why the change-point detection is important. First, in general, one uses only one probabilistic mechanism of data generation. One can statistically estimate only those parameters that have not changed in the process of data acquisition. Therefore, prior to parameter estimation and model creation, one has to check the hypothesis of stochastic homogeneity of data obtained. If this hypothesis is rejected, then segments of stochastic homogeneity of data should be detected for the estimation of the parameters of the object in each segment separately. Second, disruption of the stochastic homogeneity of data might be a sign of certain changes (failures, trouble, etc.) in the process observed. Such changes require urgent decisions to avoid possible losses and casualties. Third, an intrinsic features are not observed at a glance, the superficial pattern of time series data is usually very complicated to find some structure behind the data. For these reason, the change-point detection is important for all who seek the way for finding an intrinsic non-linear model from the time series data.

The problem of the change-point detection has been studied in various fields. First, quality control or quality management is one of the first areas where change-point detection is utilized. In this area, Shewhart charts [1] have been standard procedure for 20–30 years as an indication whether or not the process is “in control”. The charts are the application of statistical theory to maintain the same condition including some stochastic variations that we cannot explain in terms of known causes, for example, random fluctuation such as humidity. Page [2] proposed stopping and declaring the process to be called “out of control”. This and related procedures are known as CUSUM (cumulative sum) procedure.

Second, the field of economist or econometricians including statisticians, has been trying to detect some structural change in economic value such as Gross Domestic Product. For example, Chow [3] proposed “Chow Test” for detecting the changes of the regression parameters. Takanami and Kitagawa [4] proposed the locally stationary AR model to estimate onset times of seismic waves.

Third, the change-point detection problem is sometimes called “time series clustering” among the data mining community who utilize various methods. Kresta et al. [5] and Dunia and Qin [6] adopted Principal Component Analysis. Mörchen [7] demonstrated the Discrete Wavelet Transform and the Discrete Fourier Transform for dimensionality reduction. Fu et al. [8] adopted self-organizing maps (SOM) for pattern discovery in temporal data sequences which are segmented using a continuous sliding window. Chantelou et al. [9] adopted SOM for the time series clustering. They applied SOM to a large number of load curves collected from each customer of Electricité de France to find a small number of classes. Wang et al. [10] proposed characteristic-based

clustering using 13 set of global measure including periodicity, non-linearity by neural network test, skewness, kurtosis, self-similarity by Hurst parameter and Lyapunov Exponent. They showed that their method performed better in detecting the global characteristics for the time series. But it is unclear whether the method performs better for change-point detection in non-linear dynamics. Desobry et al. [11] proposed the kernel method to be called Kernel Change Detection based on the Support Vector Machine.

Fourth, dynamical system community introduced another visualization method. Broomhead and King [12] proposed the application of Singular Spectrum Analysis (SSA) to the problems of dynamical systems theory. Vautard and Ghil [13] and Vautard et al. [14] first applied SSA to extract trends and harmonics and to eliminate noise from climate and geophysical time series. Wang et al. [15] were the first to apply SSA to the diagnosis of rotating machinery failures using vibration signals. Moskvina and Zhigljavsky [16] introduced the change-point detection algorithm by SSA. Vaisman et al. [17] applied this method to the electromyographic signal data. Eckmann et al. [18] introduced the Recurrence Plots to visualize the recurrence of states in a phase space. In order to describe the recurrence plots in a quantitative way, recurrence quantification analysis (RQA) which quantify the small-scale structures, such as nonlinear determinism, was proposed by [19–21].

Despite these rapid developments, the change-point detection remains challenging problems. The first reason is that it is usually difficult to find intrinsic features at a glance. Their method cannot find intrinsic changes that are not necessarily directly observed. The second reason is that the existing methods cannot detect how the non-linear dynamics changes after the change point. For example, the existing methods cannot distinguish the difference between a temporal outlier and a non-linear structural change. Such a temporal outlier occurs accidentally in a variety of time series, but we are not interested in the outlier and want to ignore this.

In this paper, we introduce the change point visualization method based on the theory of KM_2O -Langevin equations. The proposed method overcomes the problem mentioned above, because it can deal with complex non-linear profile of data without any assumption. The theory of KM_2O -Langevin equations originates from the investigation of the mathematical and technological structure of the fluctuation-dissipation theorem, which is thought to be one of the principles of non-equilibrium statistical physics. We developed our time series analysis so as to carry out data analysis “from data to model”. Our analysis requires no prior information about the data and does not define any parametric models before data analysis, but extracts explanatory models in the stochastic difference equations directly from the data by checking the stationarity of the time series [22].

As an important part of this method, we define the pseudo determinacy function that describes a profile of dynamics for time series. Since we can estimate the distance of non-linear dynamics between before and after the change point, we can distinguish the difference between a temporal outlier and a non-linear structural change. It is hard to understand without some reduction of dimension because the pseudo determinacy function has high dimensions (in this paper, we treat it with 171 dimensions). Therefore, we utilize SOM for visualization and interpretation of the pseudo determinacy function. Furthermore, we investigate the performance of the method with some artificial data and real data. By using some artificial data, we confirmed that the

proposed method can detect outlier, mean shift and a change in dynamics comparing with some existing methods. We use the point of sales (POS) data as the real data. By using the real data, we found the relation between the change in dynamics and the period of the exposure of the actress in TV commercial film.

This paper is organized as follows. In Sect. 2, we introduce a methodology for change point detection based on the theory of KM_2O -Langevin equations. In Sect. 3, we perform some empirical analysis of the proposed method with four kinds of artificial data, comparing to SSA [16] and RQA [19–21]. In Sect. 4, we confirm the usefulness of the proposed method by using some real data. In Sect. 5, we discuss about results of the analysis.

2 Method

In order to visualize the change point in non-linear dynamics behind the time series data, we introduce a method which consists of the theory of KM_2O -Langevin equations [24–28] and SOM [29] according to the following steps. First, we define the pseudo determinacy function related to the characteristics of the non-linear dynamics by extension of the determinacy function of type (6,2) by [30] which is localization of Test (D) [31–33]. The detail of the algorithm is shown in Appendix A. Second, we define the degree of difference between subsequences of time series extracted with a sliding window (to be called a “cut length” in this paper) as the distance between each pseudo determinacy function that relate to the characteristics of the dynamics in the each subsequence of time series data. Finally, we visualize the characteristics of non-linear dynamics with SOM. The algorithm for minimum KM_2O -Langevin system used for following steps is shown in Appendix A

2.1 The pseudo determinacy function of type (6, 2; j, k)

By applying the results in [30] to $\mathcal{X}_{jk}^{(t)}$ and $\mathcal{Y}_0^{(t)}$ defined by Step 4 in Appendix A, we define a function $D^p(\mathcal{Z})(t; j, k) = D_n^p(\mathcal{Z})(t; j, k) : \{0, 1, \dots, M_3(N_{jk}^{(t)})\} \rightarrow [0, 1]$ for each t ($\ell + L + 1 \leq t \leq r$) by

$$D_n^p(\mathcal{Z})(t; j, k) \equiv C_n \left(\mathcal{Y}_0^{(t)} | \mathcal{X}_{jk}^{(t)} \right), \quad (1)$$

and call it a pseudo determinacy function of type (6,2; j, k) at time t associated with the time series $\mathcal{Z}$, where $N_{jk}^{(t)} \equiv t - \max \left\{ \sigma_j^{(t)}, \sigma_k^{(t)} \right\}$ and $M_3(N_{jk}^{(t)}) = \left\lceil \sqrt{N_{jk}^{(t)} + 1} \right\rceil - 1$. According to [31,32,34,35], the pseudo determinacy function is calculated by

$$C_n(\mathcal{Y}_0^{(t)} | \mathcal{X}_{jk}^{(t)}) = \left\{ \sum_{i=0}^n C(n, i) V_+(\mathcal{X}_{jk}^{(t)})(i)^t C(n, i) \right\}^{1/2}, \quad (2)$$

and

$$C(n, i) \equiv \begin{cases} R(\mathcal{Y}_0^{(t)}, \mathcal{X}_{jk}^{(t)}(n)R((\mathcal{X}_{jk}^{(t)})(0))^{-1} & (i = 0), \\ R(\mathcal{Y}_0^{(t)}, \mathcal{X}_{jk}^{(t)}(n-i) \sum_{k=0}^{i-1} \left(R(\mathcal{Y}_0^{(t)}, \mathcal{X}_{jk}^{(t)}(n-k) \right. \\ \quad \times {}^t\gamma_+^0(\mathcal{X}_{jk}^{(t)})(i, k) \} V_+(\mathcal{X}_{jk}^{(t)}(i))^{-1} & \left. \left(1 \leq i \leq M_3 \left(N_{jk}^{(t)} \right) \right) \right). \end{cases} \quad (3)$$

Contrast to the determinacy function $D(\mathcal{Z}) = D(\mathcal{Z})(t) : \{\ell + L + 1 \leq t \leq r\} : \{\ell + L + 1, \ell + L + 2, \dots, r\}$ defined by

$$D(\mathcal{Z})(t) \equiv \max\{D_{M_3(N_{j'k'}^{(t)})}(\mathcal{Z})(t; j', k')\}, \quad (4)$$

where $j', k' \in \{(j', k')\}$; three-dimensional time series ${}^t(\mathcal{Y}_0^{(t)}, {}^t\mathcal{X}_{jk}^{(t)})$ satisfies stationarity condition } in [30] which is utilized for the model selection, the pseudo determinacy function $D_n^p(\mathcal{Z})(t; j, k) (0 \leq j < k \leq 18)$ does not depend on stationarity condition and remains information about non-linear property.

2.2 The pseudo determinacy function

We define a pseudo determinacy function $\mathbf{D}(\mathcal{Z}) : \{\ell + L + 1, \dots, r\} \rightarrow [0, 1]^{171}$ by

$$\mathbf{D}(\mathcal{Z})(t) \equiv {}^t \left(D_{M_3(N_{0,1}^{(t)})}^p(\mathcal{Z})(t, 0, 1), \dots, D_{M_3(N_{0,18}^{(t)})}^p(\mathcal{Z})(t, 0, 18), \right. \\ D_{M_3(N_{1,2}^{(t)})}^p(\mathcal{Z})(t, 1, 2), \dots, D_{M_3(N_{1,18}^{(t)})}^p(\mathcal{Z})(t, 1, 18), \\ \left. \dots, D_{M_3(N_{17,18}^{(t)})}^p(\mathcal{Z})(t, 17, 18) \right). \quad (5)$$

This 171-dimensional vector represents a profile of non-linear dynamics for time series $\mathcal{Z}$ at time t .

2.3 Distance between each non-linear dynamics

By extending the determinacy function $D(\mathcal{Z})$ to the pseudo determinacy function $\mathbf{D}(\mathcal{Z})$, we define the degree of difference in non-linear dynamics by the distance of the pseudo determinacy function associated with the each time series data $\mathcal{Z}^{(t)}$ by

$$d(t, t-1) \equiv \|\mathbf{D}(\mathcal{Z})(t) - \mathbf{D}(\mathcal{Z})(t-1)\|, \quad (6)$$

where $\|\cdot\|$ is an Euclidian norm on $[0, 1]^{171}$. This distance measure represents the similarity between two time at $t-1$ and t . For example, if $d(t, t-1)$ equals zero, we can regard that the non-linear dynamics at time t and $t-1$ are the same.

2.4 Visualization with SOM

Applying the pseudo determinacy function to SOM, we can visualize the distance of the pseudo determinacy function in order to find the change in non-linear dynamics.

Associated with each component in the SOM array, we define a node's model vector $m_i(\tau) = {}^t(\mu_{i1}(\tau), \mu_{i2}(\tau), \dots, \mu_{id}(\tau)) \in \mathbf{R}^d$ ($1 \leq i \leq N, 0 \leq \tau \leq T$). The *image* of an input vector $\mathbf{D}(\mathcal{Z})(t)$ on the SOM array is defined as the array component $m_i(\tau)$ that matches best with $\mathbf{D}(\mathcal{Z})(t)$, i.e., that has the index

$$c = \arg \min_i \{d(\mathbf{D}(\mathcal{Z})(t), m_i(\tau))\}, \quad (7)$$

where c is the index of closest model vector to $\mathbf{D}(\mathcal{Z})(t)$ in the space of input signals. The algorithm of SOM is shown in Appendix B

By using SOM, we can obtain the map which describes the distance of each time t of time series $\mathcal{Z}$. The nodes of map are classified into some groups by the distance of node. For example, if $\mathcal{Z}(t_1)$ and $\mathcal{Z}(t_2)$ belong to the same group, we can regard that $\mathcal{Z}(t)(t_1 \leq t \leq t_2)$ has the same dynamics property.

3 Empirical analysis with artificial data

The purpose of this section is to present empirical analysis of several types of numerical examples and to confirm the advantage of the proposed method comparing to SSA [16] and RQA [19–21]. First, we begin with an outlier in the random noise in order to confirm that the proposed method can distinguish the difference between a temporal outlier and a lasting change of means and variances. Second, we examine a mean shift in the random noise in order to confirm that the proposed method can detect the mean shift and distinguish the difference between a mean shift and a change in non-linear dynamics. Third, we examine a change from the random noise to the data generated by chaotic system in order to confirm that the proposed method can detect the change in non-linear dynamics. Fourth, we investigate the data which has deterministic characteristics changing from linearity to non-linearity. We expect that the fourth data will be the most challenging data for all detection methods. We execute these methods under the following conditions.

- Proposed method
 - cut length: $L + 1 = 100$.
 - topology for SOM: torus type and hexagonal.
 - the number of nodes for SOM: $N = R \cdot C = 30 \cdot 24 = 720$.
 - SOM is executed by “Mr.Torus Ver.1.0” [36] based on SOM_PAK [37]).
 - SOM parameter: $\alpha = 0.01$, $T = 100,000$, $r_0 = 30$ and random seed = 1, 2, 3, 30.
 - The brightness of the nodes shows the average distance among the nodes and nearest neighbors. The white color shows that the distance is very short, so white nodes are recognized as the cluster and black nodes are recognized as the boundary of each cluster.
 - The number on each node shows the time t .

- SSA
 - SSA is executed by “ChangePoint Ver.1.0” [38].
 - We adopt CUSUM type statistics for the analysis.
 - length of base sample : 50.
 - lag of base sample : 25.
 - parameter for test sample : from 27 to 75.
- RQA
 - RQA is executed by “Commandline Recurrence Plots Ver. 1.132” by [39].
 - cut length: $L + 1 = 100$.
 - step size: $s = 1$.
 - embedding dimension : 3.
 - embedding delay : 1.
 - threshold : 0.2.
 - We adopt %RR (recurrence rate) %DET (determinism) for the analysis.

3.1 Empirical analysis of artificial data with an outlier

We consider the artificial time series data including temporal outlier as the simplest case. The data is generated by the following way:

$$Z(t) = \begin{cases} e(t) & (t = 1, \dots, 200, 202, \dots, 400), \\ 5 & (t = 201), \end{cases} \quad (8)$$

where $e(t)$ are independent, identically distributed random variables (i.i.d.r.v.) having uniform distribution over $[0, 1]$ ($e(t) \sim U(0, 1)$). The raw data and the pseudo determinacy function is shown in Fig. 1. Each graph in Fig. 1 shows the degree of the components of the pseudo determinacy function of each 171 nonlinear model. If the pseudo determinacy function is the same between the time t and t' , we can regard that the non-linear dynamics is the same. Because the pseudo determinacy function consists of 171 components, the detection of the outlier is not clear. This is the reason why we apply SOM to the pseudo determinacy function. SOM of the pseudo determinacy function is shown in Fig. 2. Obviously, we can see the two clusters of nodes divided by the dark colored nodes. The summary of the transitions are shown in Table 1.

Figure 3 shows the typical node's model vectors m_i for each node i ($i = 1, 2, \dots, 720$) (see Appendix B) of SOM corresponds to each cluster. The typical nodes are selected around the center of each cluster. From Fig. 3, we can see that cluster A, B, D and E are more similar than the cluster C which has small degree of components of the pseudo determinacy functions. For this reason, we can detect that the change point is an outlier.

Figure 4 shows the result of SSA statistics [16]. From Fig. 4, we can see that SSA detect the time t of outlier but does not detect whether or not the change point is temporal outlier, because the statistics has not the information of dynamics. We cannot find also the outlier from the result of RQA shown in Fig. 5.

Fig. 1 **a** Artificial data with an outlier; **b** The pseudo determinacy function of the artificial data with an outlier

Fig. 2 SOM of the pseudo determinacy function of artificial data with an outlier. This map is composed of 720 hexagonal nodes with 30 rows and 24 columns. Each node has the model vector $m_i = {}^t(\mu_{i1}, \mu_{i2}, \dots, \mu_{i171}) \in [0, 1]^{171}$ ($i = 1, 2, \dots, 720$). Darkness of the node describes the mean distance from 6 neighbors. The nodes are divided by dark colored nodes into A, B, C, D and E. The number on each node represents time t . The arrow on the map represents the change-point from the clusters B to C

Table 1 Summary of the result

	Cluster or Interval	From	To
The number in the column "From" and "To" represents the time t	Cluster A & B	101	200
	Cluster C	201	299
	Cluster D & E	300	400

Fig. 3 The components of node's model vector m_i of typical node i of cluster A, B, C, D and E in Fig. 2

Fig. 4 **a** Artificial data with an outlier; **b** SSA Statistics of the artificial data with an outlier

3.2 Empirical analysis of artificial data with an abrupt mean shift

We consider the artificial time series data with an abrupt mean shift. The data is generated by the following way:

$$Z(t) = \begin{cases} e(t) & (t = 1, \dots, 500), \\ 2.5 + e(t) & (t = 501, \dots, 1000), \end{cases} \quad (9)$$

Fig. 5 **a** Artificial data with with an outlier; **b** %RR of the artificial data with an outlier; **c** %DET of RQA of the artificial data with an outlier

Fig. 6 **a** Artificial data with an abrupt mean shift; **b** The pseudo determinacy function of the artificial data with an abrupt mean shift

where the $e(t)$ are i.i.d.r.v., $e(t) \sim U(0, 1)$. The raw data and the pseudo determinacy function are shown in Fig. 6. From Fig. 6, we can see that the time around 500 is the change point because the degree of all components of the pseudo determinacy function becomes very high, and goes back to the same degree before the mean shift. From Fig. 7, we can see the cluster A and the cluster B of nodes divided by the dark colored nodes.

The cluster A is filled with white colored nodes, so we can consider that the time within the cluster A has the same dynamics and the cluster B has different dynamics

Fig. 7 SOM of the pseudo determinacy function of artificial data with an abrupt mean shift This map is composed of 720 hexagonal nodes with 60 rows and 24 columns. Each node has the model vector $m_i = {}^t(\mu_{i1}, \mu_{i2}, \dots, \mu_{i171}) \in [0, 1]^{171} (i = 1, 2, \dots, 720)$. Darkness of the node describes the mean distance from 6 neighbors. The nodes are divided by dark colored nodes into A and B. The number on each node represents time t

Table 2 Summary of the result

Cluster or Interval	From	To
Cluster A	101	500
Interval	501	507
Cluster B	508	588
Interval	589	596
Cluster A	597	1000

The number in the column “From” and “To” represents the time t

from the cluster A, because the cluster B is completely isolated in this map. The summary of the transitions are shown in Table 2. The detail of the pseudo determinacy function is shown in Fig. 8 by using node’s model vector of SOM.

Figure 9 shows the result of SSA statistics [16]. Also from Fig. 9, we can see that SSA detects the time t of the mean shift but does not detect whether or not the change point is temporal outlier, because the statistics has no information of dynamics. Figure 10 shows the result of RQA. From Fig. 10, we cannot find the mean shift.

3.3 Empirical analysis of artificial data with a change from random noise to chaos

We consider the artificial time series data with the change in non-linear dynamics. We use Logistic Map [40] as the non-linear dynamics for demonstration here. The data is generated by the following way:

Fig. 8 The components of node's model vector m_i of typical node i of cluster A and B in Fig. 7

Fig. 9 **a** Artificial data with an abrupt mean shift; **b** SSA Statistics of the artificial data with an abrupt mean shift

$$Z(t) = \begin{cases} e(t) & (t = 1, \dots, 200), \\ 4Z(t-1)(1 - Z(t-1)) & (t = 201, \dots, 400), \end{cases} \quad (10)$$

where the $e(t)$ are i.i.d.r.v., $e(t) \sim U(0, 1)$. The raw data and the pseudo determinacy function is shown in Fig. 11. From Fig. 11, we can see that the time around 200–300 is the change point, because the degree of some components of the pseudo determinacy function becomes very high. From Fig. 12, we can see that three clusters of the white colored nodes are divided by the dark colored nodes. The detail of the transitions is shown in Table 3. Figure 13 shows the typical pseudo determinacy function by using node's model vector of SOM. From Fig. 13, we can see that the components of the node's model vector of cluster A are much lower than that of cluster C.

Fig. 10 **a** Artificial data with an abrupt mean shift; **b** %RR of the artificial data with an abrupt mean shift; **c** %DET of RQA of the artificial data with an abrupt mean shift

Fig. 11 **a** Artificial data with a change from random noise to chaos; **b** The pseudo determinacy function of the artificial data with a change from random noise to chaos

Figure 14 shows the result of SSA statistics [16]. From Fig. 14, we can see that SSA detects the time of the change point but do not detect the change in non-linear dynamics, because this statistics has no information of dynamics. Figure 15 shows the result of RQA [19–21]. From Fig. 15, we can see that RQA detects a change in dynamics by the increase of the RQA measure %RR and %DET.

3.4 Empirical analysis of artificial data with a change from AR model to chaos

We consider the artificial time series data with the change from AR model to chaos. We use Logistic Map [40] as the nonlinear dynamics for demonstration here. The data

Fig. 12 SOM of the pseudo determinacy function of artificial data with a change from random noise to chaos. This map is composed of 720 hexagonal nodes with 30 rows and 24 columns. Each node has the model vector $m_i = {}^t(\mu_{i1}, \mu_{i2}, \dots, \mu_{i171}) \in [0, 1]^{171}$ ($i = 1, 2, \dots, 720$). Darkness of the node describes the mean distance from 6 neighbors. The nodes are divided by dark colored nodes into A, B and C. The number on each node represents time t . The arrow on the map represents the change-points from the cluster A to the cluster B and from B to C

Table 3 Summary of the result

	Cluster or Interval	From	To
The number in the column “From” and “To” represents the time t	Cluster A	101	234
	Cluster B	235	280
	Cluster C	281	400

Fig. 13 The components of node's model vector m_i of typical node i of cluster A, B and C in Fig. 12

Fig. 14 **a** Artificial data with a change from random noise to chaos; **b** SSA Statistics of the artificial data with a change from random noise to chaos

Fig. 15 **a** Artificial data with a change from random noise to chaos; **b** %RR of the artificial data with a change from random noise to chaos; **c** %DET of RQA of the artificial data with a change from random noise to chaos

is generated by the following way:

$$Z(t) = \begin{cases} 0.9Z(t-1) + 0.01Z(t-2) + 0.3e(t) & (t = 1, \dots, 500), \\ 4Z(t-1)(1-Z(t-1)) & (t = 501, \dots, 1000), \end{cases} \quad (11)$$

where the $e(t)$ are i.i.d.r.v., $e(t) \sim U(-0.5, 0.5)$. The raw data and the pseudo determinacy function $\mathbf{D}(\mathcal{Z})(t)$ are shown in Fig. 16. From Fig. 16, we can see that the time around 500 is the change point, because the degree of some components of the pseudo determinacy function becomes very high. From Fig. 17, we can see that two clusters

Fig. 16 **a** Artificial data with a change from AR to chaos; **b** The pseudo determinacy function of the artificial data with a change from AR to chaos

Fig. 17 SOM of the pseudo determinacy function of artificial data with a change from AR to chaos. This map is composed of 720 hexagonal nodes with 30 rows and 24 columns. Each node has the model vector $m_i = {}^t(\mu_{i1}, \mu_{i2}, \dots, \mu_{i171}) \in [0, 1]^{171}$ ($i = 1, 2, \dots, 720$). Darkness of the node describes the mean distance from 6 neighbors. The nodes are divided by dark colored nodes into A, A' and B. The number on each node represents time t . The arrow on the map represents the change-points from the cluster A to the cluster A' and from A' to B

of the white colored nodes are divided by the dark colored nodes. The summary of the transitions is shown in Table 4. Figure 18 shows the typical pseudo determinacy function by using node's model vector of SOM. Also from Fig. 18, we can see that the some components of model vector $m_{i \in \text{cluster A}}$ is higher than the other compo-

Table 4 Summary of the result

	Cluster or Interval	From	To
The number in the column “From” and “To” represents the time t	Cluster A	101	535
	Interval (A')	536	595
	Cluster B	596	1000

Fig. 18 The components of node's model vector m_i of typical node i of cluster A, A' and B in Fig. 17

nents. The components correspond to the two dimensional non-linear transformation $^t(\phi_j(\mathcal{Z})(n), \phi_k(\mathcal{Z})(n))$ described in (16) in Appendix A. On the other hand, the some components of the node's model vector of cluster C is quite high. The part of components of high degree correspond to the two dimensional non-linear transformation $^t(\phi_j(\mathcal{Z})(n), \phi_k(\mathcal{Z})(n))$ ($j = 1, k = 2, \dots, 18$) described in (19) in Appendix A. This suggests that the pseudo determinacy function has some information to derive non-linear dynamics from time series.

Figure 19 shows the result of SSA statistics [16]. From Fig. 19, we can see that SSA does not detect the time of the change point, because the statistics has no information about dynamics. Figure 20 shows the result of RQA [19–21]. From Fig. 20, we can see that %DET do not detect the change but %RR detect it.

4 Empirical analysis of real data

In this section, we analyze some real data by using the proposed method for several purposes. The first purpose is to confirm the performance of the proposed method. It is worth to examine the performance with the artificial data in Sect. 3, but it is not sufficient to make sure that the proposed method can detect the change-point in non-linear dynamics within the real data. The second purpose is to investigate an indirect relation between the change-point and some external forces. It is sometimes difficult to find such a causal relation because most of phenomena are too complex to derive a causal model. But it becomes easier for when they find some change-points in the

Fig. 19 **a** Artificial data with a change from AR to chaos; **b** SSA Statistics of the artificial data with a change from AR to chaos

Fig. 20 **a** Artificial data with a change from AR to chaos; **b** %RR of the artificial data with a change from AR to chaos; **c** %DET of RQA of the artificial data with a change from AR to chaos

data that they are interested in. In this section, we use weekly volume of sales data of some brands.

4.1 About data

We use SRI supplied by Intage Inc. that is one of major sales data database service in Japan. SRI is a service to track market trends based on data collected from 5,051

Fig. 21 Sales of Brand ESS in every week

Fig. 22 Sales of brand ASI in every week

retail outlets nationwide. Sales data are collected from general merchandise stores, supermarkets, convenience stores, pharmacies, drug stores, home centers and other stores in major retail formats using POS systems. This allows major consumer goods manufacturers to access information what is essential for developing market strategies: the products were purchased when, where, in what quantity, at what price, and at what types of retail outlets. We treat brand ESS and brand ASI in shampoo category from these databases. The data consists of weekly sales from the year 2003 to 2010. We denote the sales $\mathcal{L}(n)$. Figure 21 displays the sales of brand ESS, and Fig. 22 displays the sales of brand ASI. At a glance, the sales of brand ESS have a

mean shift around time $t = 190$, and sales ASI has a small mean shift around time $t = 325$. These are obvious change point, so our first purpose of the analysis is to check whether or not our method can detect the mean shift correctly. The second purpose is to check whether or not our method can find other non-linear change in dynamics.

4.2 Analysis of the sales data

We apply the method proposed in Sect. 2 to two shampoo brands in Japan, ESS and ASI. Brand ESS has an obvious change point around time $t = 190$. But brand ASI has not such an obvious change point. So we first examine brand ESS whether or not the obvious change point is detected. Next, we examine whether or not the change contains non-linear change in dynamics. Finally, we examine brand ASI whether some change point in non-linear dynamics is related to some external force.

4.2.1 Brand ESS

Figure 23 shows the sales data $\mathcal{Z}$ of brand ESS and the pseudo determinacy function $\mathbf{D}(\mathcal{Z})$. At a glance, the average degree of the components of the pseudo determinacy function arises around $t = 190$, then decreases around $t = 280$. Figure 24 shows the SOM of brand ESS.

Obviously, we can see that two large clusters of nodes are divided by the dark colored nodes and the clusters are divided into some small clusters. Figure 24 shows the cluster and the trace of time t of brand ESS on SOM. Table 5 shows the summary of the result of SOM. The detail of the pseudo determinacy function is shown in Fig. 25 by using node's model vector of SOM.

We find from Fig. 24 that the time index from 101 to 191 are placed on the nodes of the cluster A, B or C and the time index 191, 192, 193, 194 and 195 are placed on the right side of some dark colored nodes out of these clusters and time index 196 is placed on the node of the cluster D. So we can regard that the time $t = 191$ is the first big change-point. The time t ends near the start point (Cluster A). Since the profile of the pseudo determinacy function does not change after the change point, we conclude that the change point around $t = 191$ is not the change in non-linear dynamics. The summary of the classification of dynamics is shown in Fig. 26

4.2.2 Brand ASI

Figure 27 shows the sales data $\mathcal{Z}$ of brand ASI and the pseudo determinacy function $\mathbf{D}(\mathcal{Z})$. Differently from the brand ESS, there is not any obvious mean shift in $\mathbf{D}(\mathcal{Z})$ graph. So it is useful to investigate $\mathbf{D}(\mathcal{Z})$ by using the visualization by SOM. Figure 28 shows the SOM of brand ASI. Figure 28 also shows the cluster and the trace of time period of brand ASI on SOM as Table 6.

From Fig. 28, we can regard that the time index $t = 166, 195, 263$ and 326 is the change-point in non-linear dynamics. It is possible that the change point at the time index $t = 166$ is made by another brand which is launched at the time

Fig. 23 a Sales of brand ESS, b The pseudo determinacy function of brand ESS

Fig. 24 SOM of the pseudo determinacy function of brand ESS. This map is composed of 720 hexagonal nodes with 30 rows and 24 columns. Each node has the model vector $m_i = {}^t(\mu_{i1}, \mu_{i2}, \dots, \mu_{i171}) \in [0, 1]^{171} (i = 1, 2, \dots, 720)$. Darkness of the node describes the mean distance from 6 neighbors. The nodes are divided by dark colored nodes into A, B, C, D and E. The number on each node represents time t . The arrow on the map represents the change-points from the cluster A to the cluster B, from B to D, from D to E, from E to C and from C to A

$t = 166$. Moreover, from Fig. 28 we regard that the biggest change point is the time index $t = 326$ because the cluster D and E are divided by the dark colored nodes. At time $t = 326$, the volume of sales of brand ASI is increased by the brand renewal. The detail of the pseudo determinacy function is shown in Fig. 29

Table 5 Summary of the result

Cluster or Interval	From	To
Cluster A	101	124
Interval	125	133
Cluster B	134	190
Interval	191	195
Cluster D	196	228
Interval	229	245
Cluster E	246	275
Interval	276	287
Cluster C	288	323
Cluster A	324	387

The number in the column
“From” and “To” represents the
time t

Fig. 25 The components of node's model vector m_i of typical node i of cluster A, B, C, D and E in the SOM of brand ESS

by using node's model vector of SOM. Obviously, the cluster E is different from other clusters. The summary of the classification of non-linear dynamics is shown in Fig. 30.

5 Discussion and future research

In this paper, we proposed a method for visualization of change in non-linear dynamics by using self-organizing map based on the theory of KM_2O -Langevin equations. In order to check the feasibility of the method, we examined the proposed method by four types of artificial data and two real data from marketing field.

In Sect. 3, we compared the proposed method to two kinds of existing methods; SSA change point detection by [16] and RQA by [19–21]. The summary of the result

Fig. 26 **a** Sales of Brand ESS, **b** The pseudo determinacy function of brand ESS with the summary of the results of the analysis

Fig. 27 **a** Sales of brand ASI, **b** The pseudo determinacy function of brand ASI

is shown in Table 7. In this table, “ KM_2O ” stands for the proposed method. From Table 7, we confirmed the advantage of the proposed method. SSA detects a change point but does not detect a change in non-linear dynamics. RQA detects a change in non-linear dynamics, but does not detect a change in an outlier and a mean shift. The proposed method detects not only all types of change point, but also a change in non-linear dynamics. Moreover, we found that the proposed method can detect how

Fig. 28 SOM of the pseudo determinacy function of brand ASI. This map is composed of 720 hexagonal nodes with 30 rows and 24 columns. Each node has the model vector $m_i = {}^t(\mu_{i1}, \mu_{i2}, \dots, \mu_{i171}) \in [0, 1]^{171} (i = 1, 2, \dots, 720)$. Darkness of the node describes the mean distance from 6 neighbors. The nodes are divided by dark colored nodes into A, B, C, D and E. The number on each node represents time t . The arrow on the map represents the change-points from the cluster A to the cluster B, from B to C, from C to D and from D to E

Table 6 Summary of the result

Cluster or Interval	From	To
Cluster A	141	165
Interval	166	169
Cluster B	170	194
Interval	195	209
Cluster C	210	262
Interval	263	268
Cluster D	269	325
Interval	326	347
Cluster E	348	387

The number in the column
“From” and “To” represents the
time t

the dynamics changes. For example, Fig. 18 clearly shows that the dynamical system changes from linearity to non-linearity, quadratic and chaotic characteristics.

We found that the brand ESS has a mean shift of sales that can be misread as a structural change by other change point detection method. We found that the brand ASI has some change points in non-linear dynamics and the change points correspond to the change of the actress in TV commercial and the action of competitor brands.

The detail is shown in Table 8. It is to be noted that despite of the same actress, the Cluster A and B is divided because of the Launch of brand TSU. It is sometimes difficult to find such a causal relation because most of phenomena are too complex to derive a causal model. But this result shows that it will become easier when we find some change-points in the data that we are interested in.

Fig. 29 The components of node's model vector m_i of typical node i of cluster A, B, C, D and E in the SOM of brand ASI

Fig. 30 **a** Sales of brand ASI, **b** The pseudo determinacy function of brand ASI with the summary of the results of the analysis

Figure 29 shows that the dynamics of the cluster can be classified by the shape of these graph. For example, the shape of the graph of the cluster C has “liner-rich” (poor non-linearity) characteristics, because the first 18 components of the pseudo determinacy function, which include linear part, are higher than the other components, and the cluster E has “non-linear-rich” characteristics, because many components of the pseudo determinacy function are very high.

Table 7 Summary of the result of Sect. 3

Type of change	KM ₂ O	SSA	RQA
Outlier	GOOD	OK	NG
Mean shift	GOOD	OK	NG
Random to logistic	GOOD	OK	OK
AR to logistic	GOOD	NG	OK

Good detected both change point and change in dynamics, *OK* detected only change point, *NG* not detected both, *KM₂O* proposed method, *SSA* Singular Spectrum Analysis and *RQA* recurrence quantification analysis, respectively

Table 8 The result and the actress in TV commercial of brand ASI

Cluster or Interval	From	To	Actress or event
Cluster A	141	165	Actress V (141 to 213)
Interval	166	169	Launch of Brand TSU
Cluster B	170	194	Actress V (141 to 213)
Interval	195	209	Renewal of Brand ESS
Cluster C	210	262	Actress W (214 to 247)
Interval	263	268	
Cluster D	269	325	Actress X (248 to 321)
Interval	326	347	Renewal of Brand ASI
Cluster E	348	387	Actress Y (322 to 387)

The number of column “From” and “To” represents the time t . () in “Actress or Event” shows time range of exposure

In this paper, we estimated the pseudo determinacy function without checking the stationarity which is given based on the theory of KM₂O-Langevin equations. Since the sample covariance function satisfies the non-negative definite condition, we calculated the KM₂O-Langevin matrix functions according to the Fluctuation–Dissipation algorithm. If the pseudo determinacy function is stable, in other words, the sequence is in the same cluster in Figs. 25 and 29, we can see that the sequence satisfies a necessary condition of the Fluctuation–Dissipation theorem which is sufficient condition for stationarity. But there is the room of examination whether the detected time series clusters have stationarity condition. Once the data passes the stationarity test (Test(S)) [26,41], we can investigate details of non-linear dynamics behind each cut length by using the causal analysis [34,35].

Acknowledgments This work is supported by the Meiji University Global COE Program “Formation and Development of Mathematical Sciences Based on Modeling and Analysis”. We are grateful to Mr. Shimazaki for the SOM software, Dr. Moskvina and Prof. Zhigljavsky for the SSA software and Prof. Marwan for the RQA software.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

Appendix A: The sample determinacy function based on the theory of KM_2O -Langevin equations

We recall the determinacy function proposed by Nakano and Okabe [30] according to the following steps.

Let $\mathcal{Z} = (\mathcal{Z}(n); \ell \leq n \leq r)$ be any one-dimensional time series. Here, we make Test(D) proposed up to now [31–33] localize to define a determinacy function of type (6,2) which draws the behavior of non-linear determinacy values associated with the time series $\mathcal{Z}$.

Step 1 (cut length and shift time series)

Fix any two natural numbers L, t ($\ell + L + 1 \leq t \leq r$). We define a time series $\mathcal{Z}^{(t)} = (\mathcal{Z}^{(t)}(n); t - L \leq n \leq t)$ with the cut length $L + 1$ by $\mathcal{Z}^{(t)}(n) \equiv \mathcal{Z}(n)$. By shifting the time series $\mathcal{Z}^{(t)}$, we define two kinds of time series $\mathcal{Z}^{(t),sh:0} = (\mathcal{Z}^{(t),sh:0}(n); t - L \leq n \leq t)$ and $\mathcal{Z}^{(t),sh+1} = (\mathcal{Z}^{(t),sh:sh+1}(n); t - L \leq n \leq t)$ by $\mathcal{Z}^{(t),sh:0}(n) \equiv \mathcal{Z}^{(t)}(n - 1)$ and $\mathcal{Z}^{(t),sh+1}(n) \equiv \mathcal{Z}^{(t)}(n)$.

Step 2 (standardization)

We define a standardized time series $\tilde{\mathcal{Z}} = (\tilde{\mathcal{Z}}(n); \ell \leq n \leq r)$ of the time series $\mathcal{Z}$ by

$$\tilde{\mathcal{Z}}_i(n) \equiv \frac{\mathcal{Z}_i(n) - \mu_i(\mathcal{Z})}{\sqrt{v_{ii}(\mathcal{Z})}} \quad (1 \leq i \leq d), \quad (12)$$

where $\mu_i(\mathcal{Z})$ and $v_{ii}(\mathcal{Z})$ stand for the i th-component and (i, i) -component of the sample mean vector $\mu(\mathcal{Z})$ and the sample variance matrix $v(\mathcal{Z})$ of $\mathcal{Z}$, respectively, defined by

$$\mu(\mathcal{Z}) \equiv \frac{1}{r - \ell + 1} \sum_{n=\ell}^r \mathcal{Z}(n), \quad (13)$$

$$v(\mathcal{Z}) \equiv \frac{1}{r - \ell + 1} \sum_{n=\ell}^r (\mathcal{Z}(n) - \mu(\mathcal{Z}))^t (\mathcal{Z}(n) - \mu(\mathcal{Z})). \quad (14)$$

We assume that $v_{ii}(\mathcal{Z}) > 0$ ($1 \leq i \leq d$). It is to be noted that $\mu_i(\tilde{\mathcal{Z}}) = 0$ and $v_{ii}(\tilde{\mathcal{Z}}) = 1$. By applying this, we obtain $\tilde{\mathcal{Z}}^{(t),sh:0} = (\tilde{\mathcal{Z}}^{(t),sh:0}(n); t - L \leq n \leq t)$ and $\tilde{\mathcal{Z}}^{(t),sh+1} = (\tilde{\mathcal{Z}}^{(t),sh:sh+1}(n); t - L \leq n \leq t)$.

Step 3 (non-linear transformations up to rank 6)

By applying the non-linear 19 numbers transformations up to rank 6 to the time series $\tilde{\mathcal{Z}}^{(t),sh:0}$, we construct 19 one-dimensional time series $\tilde{\mathcal{Z}}_i^{(t),sh:0}$ ($0 \leq i \leq 18$) as follows:

$$\begin{aligned} \tilde{\mathcal{Z}}_i^{(t),sh:0} &= (\tilde{\mathcal{Z}}_i^{(t),sh:0}(n); t - L + \sigma(i) \leq n \leq t) \\ \text{and } \tilde{\mathcal{Z}}_i^{(t),sh:0}(n) &= \varphi_i(\tilde{\mathcal{Z}}^{(t),sh:0}(n)), \end{aligned} \quad (15)$$

where $\varphi_i(x)(n)$ are polynomials of variables $x = (x(n); t - L + \sigma(i) \leq n \leq t) (\in \mathbf{R}^{L+1-\sigma(i)})$ with degree p_i defined by

$$\left\{ \begin{array}{ll} \varphi_0(x)(n) = x(n) & (t - L + \sigma(0) \leq n \leq t), \\ \varphi_1(x)(n) = x(n)^2 & (t - L + \sigma(1) \leq n \leq t), \\ \varphi_2(x)(n) = x(n)^3 & (t - L + \sigma(2) \leq n \leq t), \\ \varphi_3(x)(n) = x(n)x(n-1) & (t - L + \sigma(3) \leq n \leq t), \\ \varphi_4(x)(n) = x(n)^4 & (t - L + \sigma(4) \leq n \leq t), \\ \varphi_5(x)(n) = x(n)^2x(n-1) & (t - L + \sigma(5) \leq n \leq t), \\ \varphi_6(x)(n) = x(n)x(n-2) & (t - L + \sigma(6) \leq n \leq t), \\ \varphi_7(x)(n) = x(n)^5 & (t - L + \sigma(7) \leq n \leq t), \\ \varphi_8(x)(n) = x(n)^3x(n-1) & (t - L + \sigma(8) \leq n \leq t), \\ \varphi_9(x)(n) = x(n)^2x(n-2) & (t - L + \sigma(9) \leq n \leq t), \\ \varphi_{10}(x)(n) = x(n)x(n-1)^2 & (t - L + \sigma(10) \leq n \leq t), \\ \varphi_{11}(x)(n) = x(n)x(n-3) & (t - L + \sigma(11) \leq n \leq t), \\ \varphi_{12}(x)(n) = x(n)^6 & (t - L + \sigma(12) \leq n \leq t), \\ \varphi_{13}(x)(n) = x(n)^4x(n-1) & (t - L + \sigma(13) \leq n \leq t), \\ \varphi_{14}(x)(n) = x(n)^3x(n-2) & (t - L + \sigma(14) \leq n \leq t), \\ \varphi_{15}(x)(n) = x(n)^2x(n-1)^2 & (t - L + \sigma(15) \leq n \leq t), \\ \varphi_{16}(x)(n) = x(n)^2x(n-3) & (t - L + \sigma(16) \leq n \leq t), \\ \varphi_{17}(x)(n) = x(n)x(n-1)x(n-2) & (t - L + \sigma(17) \leq n \leq t), \\ \varphi_{18}(x)(n) = x(n)x(n-4) & (t - L + \sigma(18) \leq n \leq t) \end{array} \right. \quad (16)$$

and $\sigma(i)$, p_i are given by

$$\left\{ \begin{array}{l} \sigma(0) = \sigma(1) = \sigma(2) = \sigma(4) = \sigma(7) = \sigma(12) = 0, \\ \sigma(3) = \sigma(5) = \sigma(8) = \sigma(10) = \sigma(13) = \sigma(15) = 1, \\ \sigma(6) = \sigma(9) = \sigma(14) = \sigma(17) = 2, \\ \sigma(11) = \sigma(16) = 3, \\ \sigma(18) = 4, \end{array} \right. \quad (17)$$

$$\left\{ \begin{array}{l} p_0 = 1, \\ p_1 = p_3 = p_6 = p_{11} = p_{18} = 2, \\ p_2 = p_5 = p_9 = p_{10} = p_{16} = p_{17} = 3, \\ p_4 = p_8 = p_{14} = p_{15} = 4, \\ p_7 = p_{13} = 5, \\ p_{12} = 6. \end{array} \right. \quad (18)$$

For any fixed integers j, k ($0 \leq j < k \leq 18$), we define a two-dimensional time series $\tilde{\mathcal{Z}}_{jk}^{(t), sh:0} = (\tilde{\mathcal{Z}}_{jk}^{(t), sh:0}(n); \max\{t - L + \sigma(j), t - L + \sigma(k)\} \leq n \leq t)$ by

$$\tilde{\mathcal{Z}}_{jk}^{(t), sh:0}(n) \equiv {}^t \left(\tilde{\mathcal{Z}}_j^{(t), sh:0}(n), \tilde{\mathcal{Z}}_k^{(t), sh:0}(n) \right). \quad (19)$$

Moreover, we denote the standardized time series $\mathcal{X}_{jk}^{(t)}(n) \equiv \tilde{\mathcal{Z}}_{jk}^{(t), sh:0}(n)$ and $\mathcal{Y}_0^{(t)}(n) \equiv \tilde{\mathcal{Z}}^{(t), sh+1}(n)$.

Step 4 (sample covariance matrix function)

We define a sample covariance matrix function $R(\tilde{\mathcal{Z}}) = (R(\tilde{\mathcal{Z}})(n); |n| \leq M_d(r - \ell))$ of the time series $\tilde{\mathcal{Z}}$ by

$$\begin{cases} R(\tilde{\mathcal{Z}})(n) \equiv \frac{1}{r - \ell + 1} \sum_{k=0}^{r-\ell-n} \tilde{\mathcal{Z}}(n + \ell + k)^t \tilde{\mathcal{Z}}(\ell + k) & (0 \leq n \leq M_d(r - \ell)), \\ R(\tilde{\mathcal{Z}})(-n) \equiv {}^t R(\tilde{\mathcal{Z}})(n) & (0 \leq n \leq M_d(r - \ell)), \end{cases} \quad (20)$$

where $M_d(r - \ell) \equiv 3[\sqrt{r - \ell + 1}/d] - 1$. We note from certain experience rule in time series analysis that the number $M_d(r - \ell) + 1$ is equal to the number of n ($0 \leq n \leq r - \ell$) such that the value $R(\tilde{\mathcal{Z}})(n)$ is reliable. By applying this definition to the time series ${}^t(\mathcal{Y}_0^{(t)}, {}^t\mathcal{X}_{jk}^{(t)})$, we calculate the sample covariance matrix functions $R(\mathcal{Y}_0^{(t)})(n)$, $R(\mathcal{X}_{jk}^{(t)})(n)$ and $R(\mathcal{Y}_0^{(t)}, \mathcal{X}_{jk}^{(t)})(n)$. Note that the sample $R(\mathcal{Y}_0^{(t)})(n)$, $R(\mathcal{X}_{jk}^{(t)})(n)$ and $R(\mathcal{Y}_0^{(t)}, \mathcal{X}_{jk}^{(t)})(n)$ are non-negative definite.

Step 5 (minimum KM₂O-Langevin system with fluctuation dissipation algorithm)

Using the fluctuation-dissipation algorithm [42–45, 31, 32, 23], we construct the minimum KM₂O-Langevin system $\mathcal{LM}(R(\mathcal{X}_{jk}^{(t)}))$ associated with the matrix function $R(\mathcal{X}_{jk}^{(t)})$:

$$\mathcal{LM}(R(\mathcal{X}_{jk}^{(t)})) \equiv \{\gamma_{\pm}^0(\mathcal{X}_{jk}^{(t)})(n, k), V_{\pm}(\mathcal{X}_{jk}^{(t)})(m); 0 \leq k < n \leq M_d(r - \ell), 0 \leq m \leq M_d(r - \ell)\}. \quad (21)$$

It is to be noted that all components of the system $\mathcal{LM}(R(\mathcal{X}_{jk}^{(t)}))$ are $d \times d$ matrices. Since the sample covariance matrix function $R(\mathcal{X}_{jk}^{(t)})$ is non-negative definite, we can obtain the system $\mathcal{LM}(R(\mathcal{X}_{jk}^{(t)}))$ more effectively according to the following algorithm.

[FDA Step 0]

$$V_{\pm}(\mathcal{X}_{jk}^{(t)})(0) \equiv R(\mathcal{X}_{jk}^{(t)})(0) \quad (22)$$

[FDA Step 1]

$$\begin{cases} \delta_+(\mathcal{X}_{jk}^{(t)})(1) \equiv -R(\mathcal{X}_{jk}^{(t)})(1)V_-(\mathcal{X}_{jk}^{(t)})(0)^{-1}, \\ \delta_-(\mathcal{X}_{jk}^{(t)})(1) \equiv -{}^tR(\mathcal{X}_{jk}^{(t)})(1)V_+(\mathcal{X}_{jk}^{(t)})(0)^{-1}, \\ \gamma_+(\mathcal{X}_{jk}^{(t)})(1, 0) \equiv \delta_+(\mathcal{X}_{jk}^{(t)})(1), \\ \gamma_-(\mathcal{X}_{jk}^{(t)})(1, 0) \equiv \delta_-(\mathcal{X}_{jk}^{(t)})(1), \\ V_+(\mathcal{X}_{jk}^{(t)})(1) \equiv (I - \delta_+(\mathcal{X}_{jk}^{(t)})(1)\delta_-(\mathcal{X}_{jk}^{(t)})(1))V_+(\mathcal{X}_{jk}^{(t)})(0), \\ V_-(\mathcal{X}_{jk}^{(t)})(1) \equiv (I - \delta_-(\mathcal{X}_{jk}^{(t)})(1)\delta_+(\mathcal{X}_{jk}^{(t)})(1))V_-(\mathcal{X}_{jk}^{(t)})(0) \end{cases} \quad (23)$$

[FDA Step m]

$$\begin{cases} \delta_+(\mathcal{X}_{jk}^{(t)})(m) \equiv -\{R(\mathcal{X}_{jk}^{(t)})(m) \\ \quad + \sum_{k=0}^{m-2} \gamma_+(\mathcal{X}_{jk}^{(t)})(m-1, k)R(\mathcal{X}_{jk}^{(t)})(k+1)\}V_-(\mathcal{X}_{jk}^{(t)})(m-1)^{-1}, \\ \delta_-(\mathcal{X}_{jk}^{(t)})(m) \equiv -\{{}^tR(\mathcal{X}_{jk}^{(t)})(m) \\ \quad + \sum_{k=0}^{m-2} \gamma_-(\mathcal{X}_{jk}^{(t)})(m-1, k){}^tR(\mathcal{X}_{jk}^{(t)})(k+1)\}V_+(\mathcal{X}_{jk}^{(t)})(m-1)^{-1}, \\ \gamma_+(\mathcal{X}_{jk}^{(t)})(m, k) \equiv \gamma_+(\mathcal{X}_{jk}^{(t)})(m-1, k-1) \\ \quad + \delta_+(\mathcal{X}_{jk}^{(t)})(m)\gamma_-(\mathcal{X}_{jk}^{(t)})(m-1, m-k-1) \quad (1 \leq k \leq m-1), \\ \gamma_-(\mathcal{X}_{jk}^{(t)})(m, k) \equiv \gamma_-(\mathcal{X}_{jk}^{(t)})(m-1, k-1) \\ \quad + \delta_-(\mathcal{X}_{jk}^{(t)})(m)\gamma_+(\mathcal{X}_{jk}^{(t)})(m-1, m-k-1) \quad (1 \leq k \leq m-1), \\ \gamma_+(\mathcal{X}_{jk}^{(t)})(m, 0) \equiv \delta_+(\mathcal{X}_{jk}^{(t)})(m), \\ \gamma_-(\mathcal{X}_{jk}^{(t)})(m, 0) \equiv \delta_-(\mathcal{X}_{jk}^{(t)})(m), \\ V_+(\mathcal{X}_{jk}^{(t)})(m) \equiv (I - \delta_+(\mathcal{X}_{jk}^{(t)})(m)\delta_-(\mathcal{X}_{jk}^{(t)})(m))V_+(\mathcal{X}_{jk}^{(t)})(m-1), \\ V_-(\mathcal{X}_{jk}^{(t)})(m) \equiv (I - \delta_-(\mathcal{X}_{jk}^{(t)})(m)\delta_+(\mathcal{X}_{jk}^{(t)})(m))V_-(\mathcal{X}_{jk}^{(t)})(m-1). \end{cases} \quad (24)$$

Appendix B: Algorithm of SOM

We recall an algorithm by Kohonen [29] as follows.

1. Define the pseudo determinacy function $\mathbf{D}(\mathcal{X}) = (\mathbf{D}(\mathcal{X})(t); \ell + L + 1 \leq t \leq r)$ as input vectors.
2. Define the map's rows (R) and columns (C) number where $R \cdot C = N$.
3. Define the map's topology. In this paper, we used the torus structure.
4. Randomize the map's nodes' model vectors m_i to initialize the all $m_i(0)$.
5. Select an input vector $x(0)$.
6. Find $m_c(0)$.
7. Update $m_i(0)$ to $m_i(1)$ by

$$m_i(1) = m_i(0) + h_{ci}(0)(x(0) - m_i(0)). \quad (25)$$

Here $h_{ci}(\tau)$ is defined by

$$h_{ci}(\tau) = \alpha(\tau) \exp\left(-\frac{(m_c(\tau) - m_i(\tau))^2}{2r(\tau)^2}\right) \quad (26)$$

and r defines the width of the kernel and α is learning rate factor defined by

$$\alpha(\tau) = \alpha \left(1 - \frac{\tau}{T}\right) \quad (27)$$

$$r(\tau) = r_0 \left(1 - \frac{\tau}{T}\right) \quad (28)$$

where α and r_0 is a learning parameter and a width parameter we decide, for example, $\alpha = 0.9$ and $r_0 = 30$.

8. If $t < r$, grab an input vector $x(t + 1)$.
9. Find $m_c(\tau)$.
10. Update $m_i(\tau)$ to $m_i(\tau + 1)$ by

$$m_i(\tau + 1) = m_i(\tau) + h_{ci}(\tau)(x(t) - m_i(\tau)). \quad (29)$$

11. If $\tau < T$, repeat (5)–(7).

Then, we can visualize the distance of each pseudo determinacy function $\mathbf{D}(\mathcal{Z})(t)$. For example, if the $\mathbf{D}(\mathcal{Z})(t)$ is mapped on node i and $\mathbf{D}(\mathcal{Z})(t - 1)$ is mapped on node i' , the distance of these two vectors are visualized as the distance of the node i and i' which is defined by $d(m_i, m_{i'})$.

References

1. Shewhart, W.: Economic Control of Quality of Manufactured Product. Bull. Amer. Soc. Qual. Control, New York (1931)
2. Page, E.: Continuous inspection schemes. *Biometrika* **41**(1–2), 100–115 (1954)
3. Chow, G.: Tests of equality between sets of coefficients in two linear regressions. *Econometrica* **28**(3), 591–605 (1960)
4. Takanami, T., Kitagawa, G.: A new efficient procedure for the estimation of onset times of seismic waves. *J. Phys. Earth* **36**(6), 267–290 (1988)
5. Kresta, J., Macgregor, J., Marlin, T.: Multivariate statistical monitoring of process operating performance. *Can. J. Chem. Eng.* **69**(1), 35–47 (1991)
6. Dunia, R., Joe Qin, S.: Subspace approach to multidimensional fault identification and reconstruction. *AIChE J.* **44**(8), 1813–1831 (1998)
7. Mörchen, F.: Time series feature extraction for data mining using DWT and DFT, Technical Report No. 33, Department of Mathematics and Computer Science Philipps-University Marburg (2003)
8. Fu, T., Chung, F., Ng, V., Luk, R.: Pattern discovery from stock time series using self-organizing maps. In: *Work. Notes of Temp. Data Min. 7th ACM SIGKDD Intl. Conf.*, Aug. 27–37 (2001)
9. Chantelou, D., Hébrail, G., Muller, C.: Visualizing 2665 electric power load curves on a single A4 sheet of paper. In: *Proceedings on Intell. Syst. Appl. Power Syst.*, pp. 126–132 (1996)
10. Wang, X., Smith, K., Hyndman, R.: Characteristic-based clustering for time series data. *Data Min. Knowl. Discov.* **13**(3), 335–364 (2006)
11. Desobry, F., Davy, M., Doncarli, C.: An online kernel change detection algorithm. *IEEE Trans. Signal Process.* **53**(8), 2961–2974 (2005)

12. Broomhead, D., King, G.: Extracting qualitative dynamics from experimental data. *Physica D: Non-linear Phenom.* **20**(2–3), 217–236 (1986)
13. Vautard, R., Ghil, M.: Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series. *Physica D: Nonlinear Phenom.* **35**(3), 395–424 (1989)
14. Vautard, R., Yiou, P., Ghil, M.: Singular-spectrum analysis: a toolkit for short, noisy chaotic signals. *Physica D: Nonlinear Phenom.* **58**(1–4), 95–126 (1992)
15. Wang, W., Chen, J., Wu, X., Wu, Z.: The application of some non-linear methods in rotating machinery fault diagnosis. *Mech. Syst. Signal Process.* **15**(4), 697–705 (2001)
16. Moskvina, V., Zhigljavsky, A.: An algorithm based on singular spectrum analysis for change-point detection. *Commun. Stat. Simul. Comput.* **32**(2), 319–352 (2003)
17. Vaisman, L., Zariffa, J., Popovic, M.: Application of singular spectrum-based change-point analysis to EMG-onset detection. *J. Electromyogr. Kinesiol.* **20**(4), 750–760 (2010)
18. Eckmann, J., Kamphorst, S., Ruelle, D.: Recurrence plots of dynamical systems. *Europhys. Lett.* **4**(9), 973–977 (1987)
19. Zbilut, J., Webber, C.L. Jr.: Embeddings and delays as derived from quantification of recurrence plots. *Phys. Lett. A* **171**(3–4), 199–203 (1992)
20. Webber, C.L. Jr., Zbilut, J.: Dynamical assessment of physiological systems and states using recurrence plot strategies. *J. Appl. Physiol.* **76**(2), 965–973 (1994)
21. Marwan, N., Wessel, N., Meyerfeldt, U., Schirdewan, A., Kurths, J.: Recurrence-plot-based measures of complexity and their application to heart-rate-variability data. *Phys. Rev. E* **66**(2), 26702 (2002)
22. Takeo, M., Ueda, H., Okabe, Y., Matsuura, M., Klimek, M.: Waveform characteristics of deep low-frequency earthquakes: time-series evolution based on the theory of the KM_2O -Langevin equation. *Geophys. J. Int.* **165**(1), 87–107 (2006)
23. Okabe, Y., Matsuura, M., Klimek, M.: On a method for detecting certain signs of stock market crashes by non-linear stationarity tests. *Int. J. Pure Appl. Math.* **3**(4), 443–484 (2002)
24. Okabe, Y.: A generalized fluctuation–dissipation theorem for the one-dimensional diffusion process. *Commun. Math. Phys.* **98**(4), 449–468 (1985)
25. Okabe, Y.: On the Theory of Brownian motion with the Alder–Wainwright effect. *J. Stat. Phys.* **45**(5–6), 953–981 (1986)
26. Okabe, Y., Nakano, Y.: The theory of KM_2O -Langevin equations and its applications to data analysis (I): stationary analysis. *Hokkaido Math. J.* **20**(1), 45–90 (1991)
27. Okabe, Y.: Application of the theory of KM_2O -Langevin equations to the linear prediction problem for the multi-dimensional weakly stationary time series. *J. Math. Soc. Jpn.* **45**(2), 277–294 (1993)
28. Okabe, Y.: A new algorithm derived from the view-point of the fluctuation-dissipation principle in the theory of KM_2O -Langevin equations. *Hokkaido Math. J.* **22**(3), 199–209 (1993)
29. Kohonen, T.: Self-organized formation of topologically correct feature maps. *Biol. Cybern.* **43**(1), 59–69 (1982)
30. Nakano, Y., Okabe, Y.: A time series analysis of economical phenomena in japan’s lost decade (1): determinacy property of the velocity of money and equilibrium solution. *J. Econ* (in press)
31. Okabe, Y.: Fluctuation–dissipation principle in time series analysis and experimental mathematics (in Japanese edition). Nihon Hyoronsha Tokyo (2002)
32. Okabe, Y.: Experimental mathematics: time series analysis of Aurora, Seismic waves, brain waves and sound (in Japanese edition). Asakura Syoten, Tokyo (2005)
33. Okabe, Y., Yamane, T.: The theory of KM_2O -Langevin equations and its applications to data analysis (III): deterministic analysis. *Nagoya Math. J.* **152**, 175–201 (1998)
34. Okabe, Y.: Langevin equations and causal analysis. *Transl. Am. Math. Soc.* **161**, 19–50 (1994)
35. Okabe, Y., Inoue, A.: The Theory of KM_2O -Langevin equations and applications to data analysis (II): causal analysis (1). *Nagoya Math. J.* **134**(1), 1–28 (1994)
36. Tokutaka, H., Ohkita, M., Fujimura, K.: Self-organizing maps and its application (in Japanese edition). Springer, Tokyo (2007)
37. Kohonen, T., Hynninen, J., Kangas, J., Laaksonen, J.: SOM_PAK AND LVQ_PAK. <http://www.cardiff.ac.uk/maths/subsites/stats/changepoint/> (1996). Accessed 10 June 2009
38. Moskvina, V., Zhigljavsky, A.: Change-point detection in time series. <http://www.cardiff.ac.uk/maths/subsites/stats/changepoint/> (2003). Accessed 20 July 2009
39. Marwan, N.: Commandline recurrence plots. <http://www.agnld.uni-potsdam.de/~marwan/6.download/rp.php> (2006). Accessed 20 July 2009
40. May, J., May, R.: The ecology of the ecological literature. *Nature* **259**, 446–447 (1976)

41. Miyano, T., Kimoto, S., Shibuta, H., Nakashima, K., Ikenaga, Y., Aihara, K.: Time series analysis and prediction on complex dynamical behavior observed in a blast furnace. *Physica D: Nonl. Phenom.* **135**(3–4), 305–330 (2000)
42. Matsuura, M., Okabe, Y.: On a non-linear prediction problem for one-dimensional stochastic processes. *Jpn. J. Math. New Series* **27**(1), 51–112 (2001)
43. Okabe, Y.: On the theory of KM_2O -Langevin equations for stationary flows (1). *J. Math. Soc. Jpn.* **51**(4), 817–841 (1999)
44. Okabe, Y.: On the theory of KM_2O -Langevin equations for stationary flows (2): construction theorem. *Acta Applicandae Mathematicae: An International Surv. J. Appl. Math. Math. Appl.* **63**(1), 307–322 (2000)
45. Okabe, Y., Matsuura, M.: On the theory of KM_2O -Langevin equations for stationary flows (3): extension theorem. *Hokkaido Math. J.* **29**(2), 369–382 (2000)