Forecasting abrupt changes in foreign exchange markets: method using dynamical network marker

Shunsuke Oya1, Kazuyuki Aihara1,2 and Yoshito Hirata1,2

1 Department of Mathematical Informatics, Graduate School of Information Science and Technology, The University of Tokyo, Bunkyo-ku, Tokyo 113-8656, Japan
2 Institute of Industrial Science, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan
E-mail: yoshito@sat.t.u-tokyo.ac.jp

Received 4 June 2014, revised 10 August 2014
Accepted for publication 2 October 2014
Published 20 November 2014
New Journal of Physics 16 (2014) 115015
doi:10.1088/1367-2630/16/11/115015

Abstract
We apply the idea of dynamical network markers (Chen et al 2012 Sci. Rep. 2 342) to foreign exchange markets so that early warning signals can be provided for any abrupt changes. The dynamical network marker constructed achieves a high odds ratio for forecasting these sudden changes. In addition, we also extend the notion of the dynamical network marker by using recurrence plots so that the notion can be applied to delay coordinates and point processes. Thus, the dynamical network marker is useful in a variety of contexts in science, technology, and society.

Keywords: early warning signals, econophysics, foreign exchange markets, delay coordinates, recurrence plots

1. Introduction
There are many situations where it is desirable to forecast when sudden changes in foreign exchange markets might occur because their changes influence the fundamentals of economies for many countries. Although methods of providing early warning signals for the markets have been proposed [1, 2], these methods have some limitations because they assume that their models remain valid even when we face sudden changes. However, these assumptions are too
strong because sudden changes might be due to bifurcations that change the qualitative characteristics of the underlying economic systems.

Chen et al [3] proposed an index, called a dynamical network marker, for providing early warning signals that forecast qualitative changes accompanying bifurcations. In this paper, we modify the construction of the dynamical network marker to create an index for a dynamical network marker from a single time series. Then we apply the dynamical network marker to a time series of foreign exchange markets. Moreover, we extend the dynamical network marker using recurrence plots [4, 5] so that the dynamical network marker can be applied to a set of delay coordinates and point processes.

The rest of this paper is organized in as follows. In section 2, we briefly review the dynamical network marker. In section 3, we modify the construction of the dynamical network marker so that we can apply the index for a dynamical network marker to a single time series. In section 4, we extend the dynamical network marker using recurrence plots with delay coordinates [6, 7] or distances for point processes [8, 9]. In section 5, we discuss the results and conclude this paper.

2. Dynamical network marker

The dynamical network marker proposed by Chen et al [3] provides an early warning signal for a change in a network system within which a bifurcation is involved. Let $G = (V, E)$ represent a graph with a set of nodes $V$ and a set of edges $E$, for which we now consider all-to-all connections between every two nodes of $V$. We divide nodes into two groups: $V_l \subset V$ is a subset of nodes which we call the leading network; the remaining subset $V_o = V \setminus V_l$ are the nodes that are outside of the leading network. If a network system is getting close to a bifurcation point, then the standard deviations for the temporal changes of the nodes in $V_l$ increase; the correlation coefficients between the two nodes in $V_l$ also increase; each correlation coefficient between a node in $V_l$ and another node in $V_o$ decreases to 0. Let $\sigma$ be the mean for the standard deviations for the temporal changes of the nodes in the leading network. Denote by $C_l$ the mean of the correlation coefficients between the two nodes in $V_l$. We also use $C_o$ for the mean of the correlation coefficients between the nodes in $V_l$ and the nodes in $V_o$. Then, an index for the dynamical network marker is defined as

$$I = \frac{\sigma C_l}{C_o}.$$  

In the paper by Chen et al [3], the authors calculated $\sigma$, $C_l$, and $C_o$ for a set of samples for each of the three conditions, normal, pre-disease, and disease states, and showed that $I$ is significantly largest for the pre-disease state, where a transition related to a bifurcation is going on.

3. Modification for a single time series

We modify the index $I$ so that we can obtain an index for a dynamical network maker from a single time series generated from a network system, especially foreign exchange markets. Suppose that a $g$-dimensional time series $x(s)$ of the observed part of the state is given and that each variable $x_m(s)$ of time series $x(s)$ corresponds to a price for a currency against another at
time $s$. In addition, we suppose that there is the unobserved part $\bar{x}(s)$ of the state. Then, the system we consider can be written as

$$
\begin{align*}
\dot{x}(s + 1) &= f(x(s), \bar{x}(s)) + W(s), \\
\dot{\bar{x}}(s + 1) &= \hat{f}(x(s), \bar{x}(s)) + \hat{W}(s),
\end{align*}
$$

(2)

where $W(s)$ and $\hat{W}(s)$ are dynamical noise, either of which might be zero. We regard the unobserved part $\bar{x}(s)$ as a bifurcation parameter and apply the idea of the dynamical network biomarker.

Let $w$ be the window size. We split a time series $\{x(s): s = 1, 2, \ldots, 2wt\}$ of observables into two: the first part $\{x(s): s = 1, 2, \ldots, wt\}$ is used for defining the leading network that can detect the early warning signals, and the second part $\{x(s): s = wt + 1, wt + 2, \ldots, 2wt\}$ is used for forecasting abrupt changes. Then, we define the size $c_t$ of changes for the observables during the time period of $(w(t - 1), wt]$ as

$$
c_t = \max_m \max \left\{ \left[ \frac{\max_{s \in (w(t-1), wt]} x_m(s) - \min_{s \in (w(t-2), w(t-1)]} x_m(s)}{\min_{s \in (w(t-1), w(t-1)]} x_m(s)} \right], \left[ \frac{\min_{s \in (w(t-1), wt]} x_m(s) - \max_{s \in (w(t-2), w(t-1)]} x_m(s)}{\max_{s \in (w(t-2), w(t-1)]} x_m(s)} \right] \right\}.
$$

Given a leading network $V_l$, the index $I$ is also calculated for the window $(w(t - 1), wt]$, which we denote by $I_t(V_l)$. Namely, for a leading network $V_l$, we define $I_t(V_l)$ mathematically as follows:

$$
I_t(V_l) = \frac{\sigma_t(V_l) C_{l,t}(V_l)}{C_{o,t}(V_l)},
$$

$$
\bar{x}_d(t) = \frac{1}{w} \sum_{s \in (w(t-1), wt]} x_d(s),
$$

$$
\sigma(t, d) = \frac{1}{w} \sum_{s \in (w(t-1), wt]} (x_d(s) - \bar{x}_d(t))^2,
$$

$$
\sigma_t(V_l) = \frac{1}{|V_l|} \sum_{d \in V_l} \sigma(t, d),
$$

$$
C_{l,t}(V_l) = \frac{1}{|V_l|^2 - |V_l|} \sum_{(d, e) \in V_l \times V_l, d \neq e} \left| \frac{1}{w \sigma(t, d) \sigma(t, e)} \right| \times \sum_{s \in (w(t-1), wt]} (x_d(s) - \bar{x}_d(t))(x_e(s) - \bar{x}_e(t)).
$$
and

\[
C_{o,1}(V_t) = \frac{1}{|V_t|} \sum_{(d,e) \in \mathcal{V}_c \times \mathcal{V}_c} \left| \frac{1}{w d(t, e) \sigma(t, e)} \sum_{\tau \in (w(t-1), w(t)]} (x_d(\tau) - \bar{x}_d(t))(x_e(\tau) - \bar{x}_e(t)) \right|. \tag{3}
\]

Then we use threshold \( \theta_l(V_t) \) for \( I_l(V_t) \) and threshold \( \theta_c \) for \( c_t \) to forecast whether or not the change \( c_t \) is larger, depending on whether or not \( I_l(V_t) \) is larger. We restrict \( \theta_l(V_t) \) so that it is greater than or equal to the median for the values for the first half of each dataset. We count the numbers of incidences for \( (\theta_l < I_l(V_t) \text{ and } \theta_c < c_t) \), \( (\theta_l < I_l(V_t) \text{ and } \theta_c \geq c_t) \), \( (\theta_l \geq I_l(V_t) \text{ and } \theta_c < c_t) \), and \( (\theta_l \geq I_l(V_t) \text{ and } \theta_c \geq c_t) \) for the first half of each dataset. Let these numbers be \( b_{1,1}(V_t, \theta_l(V_t), \theta_c) \), \( b_{1,2}(V_t, \theta_l(V_t), \theta_c) \), \( b_{2,1}(V_t, \theta_l(V_t), \theta_c) \), and \( b_{2,2}(V_t, \theta_l(V_t), \theta_c) \), respectively. The leading network is selected in this paper so that the following chi-squared value \( \chi^2(V_t, \theta_l(V_t), \theta_c) \) for the first half of the dataset is maximized in terms of parameters \( \theta_l(V_t) \) and \( \theta_c \) over a set of leading networks which consist of all the subnetworks containing two or three nodes and the edges connecting them:

\[
\chi^2(V_t, \theta_l(V_t), \theta_c) = \sum_{i,j=1}^{2} \frac{b_{i,j}(V_t, \theta_l(V_t), \theta_c)^2}{p_i(V_t, \theta_l(V_t), \theta_c) \times q_j(V_t, \theta_l(V_t), \theta_c) \times n(V_t, \theta_l(V_t), \theta_c)}.
\]

\[
p_i(V_t, \theta_l(V_t), \theta_c) = \sum_{j=1}^{2} \frac{b_{i,j}(V_t, \theta_l(V_t), \theta_c)}{n(V_t, \theta_l(V_t), \theta_c)},
\]

\[
q_j(V_t) = \sum_{i=1}^{2} \frac{b_{i,j}(V_t, \theta_l(V_t), \theta_c)}{n(V_t, \theta_l(V_t), \theta_c)},
\]

\[
n(V_t, \theta_l(V_t), \theta_c) = \sum_{i,j=1}^{2} b_{i,j}(V_t, \theta_l(V_t), \theta_c),
\]

\[
\chi^2(V_t) = \max_{\theta_l(V_t), \theta_c} \chi^2(V_t, \theta_l(V_t), \theta_c). \tag{4}
\]
We use the most significant leading network $\hat{V}_t$ and the threshold values $\hat{\theta}_I(\hat{V}_t)$ and $\hat{I}_c(\hat{V}_t)$ for forecasting the abrupt changes that happen in the second part of the time series. We write the chi-squared value obtained for the second part of the time series using $\hat{V}_t$, $\hat{\theta}_I(\hat{V}_t)$ and $\hat{I}_c(\hat{V}_t)$ by $\chi^2$.

First, we tested the index $I$ with the following coupled map lattice [10]:

$$y_n(s + 1) = (1 - 2\varepsilon)\left(ay_n(s)(1 - y_n(s))\right) + (\varepsilon - \eta)\left(ay_{n+1}(s)(1 - y_{n+1}(s))\right) + (\varepsilon + \eta)\left(ay_{n-1}(s)(1 - y_{n-1}(s))\right),$$

where we set $a = 3.8$, $\varepsilon = 0.05$, $\eta = 0.01$, and $N = 100$. We generated a series of length 10 000. A part of the time series is shown in figure 1. We assumed that we could observe $y_n(s)$ ($n = 1, \ldots, 5$). We calculated the index $I_t$ every 10 points to forecast the abrupt changes.
within the next 10 points \( (w = 10) \). From the first half of the dataset, we found that the leading network containing the nodes of \( y_1, y_2, y_3 \) is most significant (figure 2).

We present the scatter plot between \( \log_{10}(I_t + 1) \) and \( \log_{10}(c_{t+1} + 1) \) in figure 3 because \( I_t \) and \( c_{t+1} \) were distributed in wide ranges. These two quantities are correlated with the correlation coefficient of 0.60 and a \( p \)-value of \( 9.8 \times 10^{-50} \). The forecasting results shown in table 1 mean that we could construct the forecasts of abrupt changes where the odds ratio defined by \( (b_{1,1}b_{2,2})/(b_{1,2}b_{2,1}) \) was 47.43. Thus, equation (5) provided statistically significant classifications.

Second, we applied the index \( I_t \) to the following Lorenz’96 model [11, 12] in a 240 dimensional phase space:
where we applied the following periodic boundary conditions:

\[
I_t < \hat{\theta}_l (\hat{V}_t) \\
I_t \geq \hat{\theta}_l (\hat{V}_t)
\]

\[
\begin{align*}
\dot{u}_k &= u_{k-1}(u_{k+1} - u_{k-2}) - u_k + \frac{h_u c}{b} \sum_{j=1}^{J} v_{j,k}, \\
\dot{v}_{j,k} &= c b v_{j+1,k} (v_{j-1,k} - v_{j+2,k}) - c v_{j,k} + \frac{h_v c}{b} u_k,
\end{align*}
\]

where \( \theta \) is the parameter.

We set \( K = 40, J = 5, F = 8, b = 10, c = 10, h_u = 1, \) and \( h_v = 1. \) Here, we assume that only \( v_{j,1} (j = 1, \ldots, 5) \) can be observed every 0.01 unit times to generate a time series of 20000 time points and that we forecast their abrupt changes for the following time period of 0.1 unit times every 0.1 unit times using the idea of the dynamical network marker (\( w = 10 \)).

The example of this time series is shown in figure 4. By calculating the maximum \( \chi^2 (V_t, \theta_l (V_t), \theta_c) \) for each potential leading network over \( \theta_l (V_t) \) and \( \theta_c \), we obtain the most significant leading network of nodes \( v_{4,1} \) and \( v_{5,1} \) (see figure 5). Then, we forecasted the abrupt changes for the second half of the time series. The scatter plot between \( \log_{10} (I_t + 1) \) and \( \log_{10} (c_{t+1} + 1) \) is shown in figure 6. These quantities were correlated with a correlation coefficient of 0.13 and a \( p \)-value of \( \times 10^{-5} \). The forecasting results are summarized in table 2. We can see that the value of \( c_{t+1} \) tends to be higher when \( I_t \) is higher. These forecasts correlated well with the outcomes of the abrupt changes.

We set \( K = 40, J = 5, F = 8, b = 10, c = 10, h_u = 1, \) and \( h_v = 1. \) Here, we assume that only \( v_{j,1} (j = 1, \ldots, 5) \) can be observed every 0.01 unit times to generate a time series of 20000 time points and that we forecast their abrupt changes for the following time period of 0.1 unit times every 0.1 unit times using the idea of the dynamical network marker (\( w = 10 \)).

The example of this time series is shown in figure 4. By calculating the maximum \( \chi^2 (V_t, \theta_l (V_t), \theta_c) \) for each potential leading network over \( \theta_l (V_t) \) and \( \theta_c \), we obtain the most significant leading network of nodes \( v_{4,1} \) and \( v_{5,1} \) (see figure 5). Then, we forecasted the abrupt changes for the second half of the time series. The scatter plot between \( \log_{10} (I_t + 1) \) and \( \log_{10} (c_{t+1} + 1) \) is shown in figure 6. These quantities were correlated with a correlation coefficient of 0.13 and a \( p \)-value of \( \times 10^{-5} \). The forecasting results are summarized in table 2. We can see that the value of \( c_{t+1} \) tends to be higher when \( I_t \) is higher. These forecasts correlated well with the outcomes of the abrupt changes.

Third, we tested our method using the datasets of foreign exchange markets. The time period was 88 weeks between 7 May 2007 and 28 December 2008. We used the pairs of the euro (EUR)/United States dollar (USD), USD/Japanese yen (JPY), USD/Swiss franc (CHF), EUR/JPY, and EUR/CHF. We took the moving average of price for the exchanges every 5 minutes for each pair. We truncated weekends and connected 12 am on Saturdays with 1 am on the following Mondays. Thus, there are 125 664 time points on this dataset.

The example of this time series is shown in figure 7. We calculated the index \( I_t \) every one hour to predict the abrupt changes that might happen in the following one hour (\( w = 12 \)). Hence there are 5235 instances to be predicted. By calculating the maximum of \( \chi^2 (V_t, \theta_l (V_t), \theta_c) \), we obtained the most significant leading network comprising the USD/CHF, EUR/JPY, and EUR/CHF pairs (figure 8). Then we tried to forecast the abrupt changes for the following part of the datasets. The quantities \( \log_{10} (I_t + 1) \) and \( \log_{10} (c_{t+1} + 1) \) correlated well with the correlation coefficient 0.74 and a \( p \)-value of less than \( 10^{-300} \) (figure 9). We found that we tended to see a
Figure 4. Time series data for the Lorenz’96 model.

Figure 5. Leading network $V_i$ versus the value of $\chi^2(V_i)$ for the example of the Lorenz’96 model using equation (4). Labels on the horizontal axis show sets of nodes for leading networks. For example, ‘1, 2’ corresponds to the leading network composed of $v_{1,1}$ and $v_{2,1}$. In this example, the most significant leading network was the one with $v_{4,1}$ and $v_{5,1}$. 
Figure 6. The scatter plot between $\log_{10}(I_t + 1)$ and $\log_{10}(c_{t+1} + 1)$ for the example of Lorenz’96 model. The gray dash-dotted line is the linear fitting for the scatter plot.

Table 2. The results of forecasting abrupt changes for the Lorenz’96 model using equation (5). See the caption of table 1 to interpret the results. The odds ratio was 2.30. The $p$-value was $3.6 \times 10^{-7}$.

|               | $c_{t+1} < \hat{\theta}_f(V_t)$ | $c_{t+1} \geq \hat{\theta}_f(V_t)$ | Total |
|---------------|---------------------------------|------------------------------------|-------|
| $I_t < \hat{\theta}_f(V_t)$ | 150                              | 403                                | 553   |
| $I_t \geq \hat{\theta}_f(V_t)$ | 62                               | 384                                | 446   |
| Total         | 212                              | 787                                | 999   |

Figure 7. Time series data for foreign exchange markets.
Figure 8. Leading network $V_l$ versus the value of $\chi^2(V_l)$ for the example of foreign exchange markets using equation (4). The numbers 1, 2, 3, 4, and 5 in the horizontal axis correspond to EUR/USD, USD/JPY, USD/CHF, EUR/JPY, and EUR/CHF, respectively. In this example, the most significant leading network was the one with nodes of USD/CHF, EUR/JPY, and EUR/CHF, respectively.

Figure 9. The scatter plot for the $\log_{10}(I_l + 1)$ and $\log_{10}(c_{l+1} + 1)$ for the foreign exchange markets. The gray dash-dotted line is the linear fitting for the scatter plot.

Table 3. The results of forecasting abrupt changes for foreign exchange markets using equation (5). See the caption of table 1 to interpret the results. The odds ratio was 13.29. The $p$-value was less than $2.2 \times 10^{-16}$.

|          | $c_{l+1} < \hat{\theta}_t(V_l)$ | $c_{l+1} \geq \hat{\theta}_t(V_l)$ | Total |
|----------|---------------------------------|------------------------------------|-------|
| $I_l < \hat{\theta}_t(V_l)$ | 1886                            | 662                                | 2548  |
| $I_l \geq \hat{\theta}_t(V_l)$ | 425                             | 1984                               | 2409  |
| Total    | 2311                            | 2646                               | 4957  |
larger change in $c_{t+1}$ when $I_t$ was larger (table 3). Thus, these forecasts correlated well with the outcomes of abrupt changes.

4. Extension to recurrence plots

We translate equation (5) so that we can use a similar notion for delay coordinates and point processes using recurrence plots [4, 5].

A recurrence plot is originally a two dimensional visualization of time series data. Both axes show the same time axis. For a pair of times, we calculate a distance and plot a point if the distance is smaller than a threshold value. Otherwise, we do not plot a point there. Mathematically, given a time series of $\{x_m(l): l = -E + 1, -E + 2, \ldots, L\}$ and threshold $r_m$ for system $m$, a recurrence plot for system $m$ can be defined as

$$R_m(i, j, r_m) = \begin{cases} 1, & \text{if } d(z_m(i), z_m(j)) < r_m, \\ 0, & \text{otherwise.} \end{cases} \quad (8)$$

Here we define $z_m(i)$ using delay coordinates [6, 7] as $z_m(i) = (x_m(i - E + 1), x_m(i - E + 2), \ldots, x_m(i))$ and $E$ is the embedding dimension.

This simple graph can show a lot of things, for example, we can calculate correlation entropy and correlation dimension from recurrence plots [13, 14]. We also recently learned that a recurrence plot contains almost all the information for the underlying dynamics because we can reproduce the rough shape of the original time series if the original time series is given as a time series with a fixed sampling frequency [15, 16]. Thus, recurrence plots have been used in a variety of contexts in science and technology including climate [17, 18], medicine [19], and economics [20, 21], to name a few. An important quantity for a recurrence plot is the recurrence rate [5], which is the proportion of plotted places. We control each threshold $r_m$ so that the recurrence rate becomes 0.2. Intuitively, such $r_m$ becomes larger when the standard deviation is larger.

A joint recurrence plot [22] is an extension of the recurrence plot towards bivariate analysis. A joint recurrence plot for systems $m_1$ and $m_2$ is defined as the intersection of two recurrence plots for the corresponding systems, namely,

$$J_{m_1, m_2}(i, j, r_{m_1}, r_{m_2}) = R_{m_1}(i, j, r_{m_1})R_{m_2}(i, j, r_{m_2}). \quad (9)$$

The important quantity for the joint recurrence plot is the joint recurrence rate [22], which is the proportion of plotted places in the joint recurrence plot. Similarly to the absolute value for the correlation coefficient, the joint recurrence rate for $J_{m_1, m_2}(i, j, r_{m_1}, r_{m_2})$ becomes larger if there is correlation between systems $m_1$ and $m_2$.

We modify equation (5) using recurrence plots. Let $\bar{r}_i(V_i)$ be the mean of threshold value $r_i(d)$ for a node $d$ included in the leading network $V_l$ such that the recurrence rates become 0.2 for the time period of $(w(t-1), wt]$. Let $JRR_{l, l}(V_l)$ be the mean for the joint recurrence rates $JRR_l(d, e)$ between two nodes $d, e \in V_l(d \neq e)$ within the leading network for the time period of $(w(t-1), wt]$. Let $JRR_{o, l}(V_l)$ be the mean for the joint recurrence rates $JRR_l(d, e)$ between a node $d \in V_l$ within the leading network and a node $e \in V_o$ that is outside of the leading network for the time period of $(w(t-1), wt]$. Mathematically, these variables are written as follows:
Then, the new index \( I'_t(V_l) \) for the dynamical network marker, given a leading network \( V_l \), is defined as

\[
I'_t(V_l) = \frac{\bar{r}_t(V_l) \times JRR_{r,t}(V_l)}{JRR_{o,t}(V_l)}.
\]

As we discussed above, \( \bar{r}_t(V_l) \), \( JRR_{r,t}(V_l) \), and \( JRR_{o,t}(V_l) \) behave similarly to \( \sigma_t(V_l) \), \( C_{r,t}(V_l) \), and \( C_{o,t}(V_l) \), respectively, while these new values can be obtained for delay coordinates and point processes as well. Therefore, we can expect that equation (10) can also detect early warning signals for a variety of contexts.

Using the first part of the time series, we define the most significant leading network for the above extension similarly to the case of equation (5). Namely, let \( b^{l,1}_{i,1}(V_l, \theta'_i(V_l), \theta'_c) \), \( b^{l,2}_{i,2}(V_l, \theta'_i(V_l), \theta'_c) \), \( b^{l,1,1}_{i,1}(V_l, \theta'_i(V_l), \theta'_c) \), and \( b^{l,2,2}_{i,2}(V_l, \theta'_i(V_l), \theta'_c) \) be the numbers of \( (I'_t(V_l) < \theta'_i(V_l) \) and \( c_{i+1} < \theta_c) \), \( (I'_t(V_l) < \theta'_i(V_l) \) and \( c_{i+1} \geq \theta_c) \), \( (I'_t(V_l) \geq \theta'_i(V_l) \) and \( c_{i+1} < \theta_c) \), and \( (I'_t(V_l) \geq \theta'_i(V_l) \) and \( c_{i+1} \geq \theta_c) \) for the first part of the time series, respectively. We define \( \chi^2(V_l, \theta'_i(V_l), \theta_c) \) as follows:

\[
\chi^2(V_l, \theta'_i(V_l), \theta_c) = \sum_{i,j=1}^2 -\frac{b^{l,1}_{i,j}(V_l, \theta'_i(V_l), \theta_c)}{\left( p'_i(V_l, \theta'_i(V_l), \theta_c) \right)^2} \times \frac{q'_j(V_l, \theta'_i(V_l), \theta_c) n'(V_l, \theta'_i(V_l), \theta_c)}{p'_i(V_l, \theta'_i(V_l), \theta_c) \times q'_j(V_l, \theta'_i(V_l), \theta_c) \times n'(V_l, \theta'_i(V_l), \theta_c)}.
\]
\[ p'_i(V_i, \theta'_i(V_i), \theta_c) = \sum_{j=1}^{n'} b'_{i,j}(V_i, \theta'_i(V_i), \theta_c), \]
\[ q'_i(V_i) = \frac{\sum_{j=1}^{n'} b'_{i,j}(V_i, \theta'_i(V_i), \theta_c)}{n'(V_i, \theta'_i(V_i), \theta_c)}, \]
\[ n'(V_i, \theta'_i(V_i), \theta_c) = \sum_{i,j=1}^{2} b'_{i,j}(V_i, \theta'_i(V_i), \theta_c). \]

We define the most significant leading network by maximizing \( \chi^2(V_i, \theta'_i(V_i), \theta_c) \) over the set of leading networks which contain two or three nodes and the threshold values \( \theta'_i(V_i) \) and \( \theta_c \):

\[ \chi^2(V_i) = \max_{\theta'_i(V_i), \theta_c} \chi^2(V_i, \theta'_i(V_i), \theta_c), \tag{11} \]

\[ \hat{V}'_i = \text{arg max}_{V_i \in \{V \times V, V \times V \times V\}} \chi^2(V_i), \]
\[ \hat{\theta}'_i(V'_i) = \text{arg max}_{\theta_c} \chi^2(V'_i, \theta'_i(V'_i), \theta_c), \]
\[ \hat{\theta}'_c(V'_i) = \text{arg max}_{\theta_c} \chi^2(V'_i, \theta'_i(V'_i), \theta_c). \]
\[ I'_i = I'_i(V'_i). \tag{12} \]

We use \( \hat{V}'_i, \hat{\theta}'_i(V'_i), \) and \( \hat{\theta}'_c(V'_i) \) for forecasting the abrupt changes for the second part of the time series. We denote by \( \chi^2 \) the chi-squared value obtained from the second part using \( \hat{V}'_i, \hat{\theta}'_i(V'_i), \) and \( \hat{\theta}'_c(V'_i) \).

Here are examples for the new index. First, we tested equation (12) with the example of the coupled map lattice above. We set \( E = 2 \) for simplicity. When we evaluated the leading networks, we found that the most significant leading network was the one containing \( y_4 \) and \( y_5 \) (figure 10). The two quantities \( \log_{10}(I'_i + 1) \) and \( \log_{10}(c_{i+1} + 1) \) were correlated with the correlation coefficient of 0.48 and a \( p \)-value of \( 4.1 \times 10^{-30} \) (figure 11). This leading network achieved 100\% success in forecasting \( c_{i+1} \geq \theta_c \) when classified as \( I'_i \geq \theta_i \) for the second half of the dataset (see table 4). Thus, the odds ratio was higher when using equation (12) than when using equation (5).

The second example is that of the Lorenz’96 model above. In this example, we set \( E = 2 \). When we applied equation (12), we found that the leading network containing \( v_{4,1} \) and \( v_{5,1} \) was most significant (figure 12). The two quantities \( \log_{10}(I'_i + 1) \) and \( \log_{10}(C_{i+1} + 1) \) were correlated with the correlation coefficient 0.13 and a \( p \)-value of \( 7.5 \times 10^{-5} \) (figure 13). When we forecasted the abrupt changes for the second half of the dataset, we obtained the odds ratio of 2.59 (see table 5), which was larger than that of table 2 obtained using equation (5). Therefore, our extension using equation (12) yielded better forecasts than equation (5).
Our third example is the dataset of foreign exchange markets. In this application, we set \( E = 2 \). Namely, the size of delay coordinates is 10 min. We found, whilst using the first half of the dataset, that the leading network containing USD/JPY, USD/CHF and EUR/JPY was most significant (figure 14). The two quantities \( \log_{10}(I'_t + 1) \) and \( \log_{10}(C_{t+1} + 1) \) were correlated

![Figure 10](image1.png)

**Figure 10.** Leading network \( V_l \) versus the value of \( \chi^2(V_l) \) for the example of the coupled map lattice using equation (11). See the caption of figure 2 to interpret the figure. The most significant leading network in this example was the one with nodes \( y_4 \) and \( y_5 \).

![Figure 11](image2.png)

**Figure 11.** The scatter plot between \( \log_{10}(I'_t + 1) \) and \( \log_{10}(C_{t+1} + 1) \) for the coupled map lattice. The gray dash-dotted line is the linear fitting for the scatter plot.

**Table 4.** The results of forecasting abrupt changes for the coupled map lattice using equation (12). See the caption of table 1 to interpret the results. The odds ratio was infinity because there was no event satisfying \( I'_t \geq \hat{\theta}_l(V_l) \) and \( C_{t+1} < \hat{\theta}_l(V_l) \). The \( p \)-value was less than \( 2.2 \times 10^{-16} \).

| \( c_{t+1} \) \( < \hat{\theta}_l(V_l) \) | \( c_{t+1} \) \( \geq \hat{\theta}_l(V_l) \) | Total |
|-------------------------------|-------------------------------|-------|
| \( I'_t < \hat{\theta}_l(V_l) \) | 56                           | 139   | 195   |
| \( I'_t \geq \hat{\theta}_l(V_l) \) | 0                            | 304   | 304   |
| Total                         | 56                           | 443   | 499   |

Our third example is the dataset of foreign exchange markets. In this application, we set \( E = 2 \). Namely, the size of delay coordinates is 10 min. We found, whilst using the first half of the dataset, that the leading network containing USD/JPY, USD/CHF and EUR/JPY was most significant (figure 14). The two quantities \( \log_{10}(I'_t + 1) \) and \( \log_{10}(C_{t+1} + 1) \) were correlated.
with the correlation coefficient 0.73 and a \( p \)-value of less than \( 10^{-300} \) (figure 15). When we forecasted the abrupt changes in the second half of the dataset, we achieved the odds ratio of 33.34 (see table 6). The value is also larger than the one in table 3 obtained by equation (5).

5. Discussions

We checked the validity of the proposed indices of equations (5) and (12) for the real dataset of the foreign exchange markets by using surrogate data analysis [23–25]. First, we applied random shuffle surrogates [23] to generate 200 data that do not have temporal correlation but preserve the spatial distribution. We found that the chi-squared values obtained by equations (5) and (12) from the original time series were larger than those obtained from random shuffle

![Figure 12](image_url)

**Figure 12.** Leading network \( V_l \) versus the value of \( \chi'^2(V_l) \) for the example of the Lorenz’96 model using equation (11). See the caption of figure 5 to interpret the results. The most significant leading network in this example was the one with \( v_{4,1} \) and \( v_{5,1} \).

![Figure 13](image_url)

**Figure 13.** The scatter plot between \( \log_{10}(I'_l + 1) \) and \( \log_{10}(C_{l+1} + 1) \) for the Lorenz’96 model. The gray dash-dotted line shows the linear fitting for the scatter plot.

with the correlation coefficient 0.73 and a \( p \)-value of less than \( 10^{-300} \) (figure 15). When we forecasted the abrupt changes in the second half of the dataset, we achieved the odds ratio of 33.34 (see table 6). The value is also larger than the one in table 3 obtained by equation (5).
Therefore, we can reject the null hypothesis that there is no temporal correlation (empirical \( p \)-values: 0.01 and 0.01, respectively, two-sided test). Second, we applied the multivariate version of iterative adjusted Fourier transform surrogates \cite{24} to the datasets, testing the null hypothesis that the time series was generated from linear noise with monotonic nonlinear transformations. The results presented in figures 18 and 19 show that we

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|}
\hline
\( I' \) & \( \hat{\theta}_1(\hat{V}_t) \) & \( \hat{\theta}_{c+1}(\hat{V}_t) \) & Total \\
\hline
\( I' < \hat{\theta}_1(\hat{V}_t) \) & 154 & 398 & 552 \\
\( I' \geq \hat{\theta}_1(\hat{V}_t) \) & 58 & 389 & 447 \\
\hline
Total & 212 & 787 & 999 \\
\hline
\end{tabular}
\caption{The results of forecasting abrupt changes for the Lorenz’96 model using equation (12). See the caption of table 1 to interpret the results. The odds ratio was 2.59. The \( p \)-value was \( 6.6 \times 10^{-9} \).
}
\end{table}

Figure 14. Leading network \( V_l \) versus the value of \( \chi^2(\hat{V}_t) \) for the example of foreign exchange markets using equation (11). See the caption of figure 8 to interpret the results. The most significant leading network was the one with USD/JPY, USD/CHF, and EUR/JPY.

Figure 15. The scatter plot between \( \log_{10}(I' + 1) \) and \( \log_{10}(C_{r+1} + 1) \) for the foreign exchange markets. The gray dash-dotted line shows the linear fitting for the scatter plot.

surrogates (see figures 16 and 17). Therefore, we can reject the null hypothesis that there is no temporal correlation (empirical \( p \)-values: 0.01 and 0.01, respectively, two-sided test). Second, we applied the multivariate version of iterative adjusted Fourier transform surrogates \cite{24} to the datasets, testing the null hypothesis that the time series was generated from linear noise with monotonic nonlinear transformations. The results presented in figures 18 and 19 show that we
cannot reject the null hypothesis (the empirical p-values: 0.76 and 0.70, respectively). We also took into account the trends of the time series and generated surrogate data by combining the method of [25] with that of [24], because it is known that the nonstationarity sometimes influences the surrogate data analysis [26]. We chose to randomize only the phases for the top 50 high-frequency components. However, the results were still the same (see figures 20 and 21):

Table 6. The results of forecasting abrupt changes for foreign change markets using equation (12). See the caption of table 1 to interpret the results. The odds ratio was 33.34. The p-value was less than $2.2 \times 10^{-16}$.

|                | $c_{t+1} < \hat{\theta}_c (\hat{V}_t)$ | $c_{t+1} \geq \hat{\theta}_c (\hat{V}_t)$ | Total |
|----------------|----------------------------------------|------------------------------------------|-------|
| $I_t' < \hat{\theta}_I (\hat{V}_t)$ | 649                                    | 763                                      | 1412  |
| $I_t' \geq \hat{\theta}_I (\hat{V}_t)$ | 95                                     | 3728                                     | 3823  |
| Total          | 744                                    | 4491                                     | 5235  |

Figure 16. Analysis using random shuffle surrogates [23] with equation (5), for the foreign exchange markets. All the observables were shuffled simultaneously so that the ‘spatial’ correlations were preserved. The solid thick line shows the value for the original data and the histogram shows the values for the random shuffle surrogates.

Figure 17. Analysis using random shuffle surrogates [23] with equation (12), for the foreign exchange markets. Here all the observables were shuffled simultaneously. The solid thick line shows the value for the original data and the histogram shows those for the random shuffle surrogates.
we could not reject the null hypothesis of linear noise (the empirical \( p \)-values: 0.66 and 0.61, respectively). There are two possibilities: the first possibility is that equations (5) and (12) are not sensitive to the potential nonlinear characteristics; the second possibility is that the markets are governed by linear dynamics. To differentiate these two possibilities, we used the entropy obtained by the joint permutations \([27, 28]\), which is an extension of the permutation entropy \([29]\). We used the permutations of length three and combined five permutations for each price time series to construct a joint permutation. In this case, the null hypothesis of linear noise was rejected by both the surrogate data of \([24]\) (figure 22, the empirical \( p \)-value: 0.01) and those of \([25]\) extended to multivariate time series based on \([24]\) (figure 23, the empirical \( p \)-value: 0.01). Therefore, it is likely that the temporal evolution of foreign exchange markets was governed by nonlinear dynamics. Simultaneously, the results of figures 18–23 mean that the proposed quantities from equations (5) and (12) mainly reflect linear characteristics of the underlying market dynamics.

**Figure 18.** Analysis using multivariate iterative amplitude adjusted Fourier transform surrogates \([24]\) with equation (5), for the foreign exchange markets. The thick line shows the value for the original data and the histogram shows the values for the surrogate data.

**Figure 19.** Analysis using multivariate iterative amplitude adjusted Fourier transform surrogates \([24]\) with equation (12), for the foreign exchange markets. The thick solid line shows the value for the original data and the histogram shows the values for the surrogate data.
Next, we randomized the ‘spatial correlation’, or correlation among different price time series components by introducing random delays to each of the five components. First, we used delays of multiples of an hour, up to 1000 h. We generated 200 combinations of such data. Then, when we compared their chi-squared values with those obtained by the same simultaneous delays to all the five components, the simultaneous delays tended to achieve larger chi-squared values than those of the random delays (see figures 24(a) and (b) for equations (5) and (12), respectively; we used 50 different same simultaneous delays to obtain the values). We also applied the random delays of multiples of a week, up to 44 weeks. But the results were similar: the same simultaneous delays tended to achieve larger chi-squared values than those of random delays temporarily shifting different components (see figures 24(c) and (d) for equations (5) and (12), respectively). Thus, the spatial correlation also exists in the dataset of the foreign exchange markets.

The overall results agreed well with the common hypotheses of econophysics that the markets are not random walk [30, 31] and are correlated with each other [30], and the results
were consistent with our previous work [20]; the underlying dynamics is possibly of deterministic chaos.

When we partially observe a high-dimensional system and apply an index for the dynamical network marker, we eventually regard the unobserved part of the high-dimensional system as a bifurcation parameter, as we discussed in section 3. Therefore, when we try to forecast abrupt changes in the partial observations for a high-dimensional system, the notion of the dynamical network marker is still valid.

The method has a limited tolerance to observational noise. We present the results with 1% Gaussian observational noise in figures 25–32. By comparing them with figures 3, 6, 11, and 13, we found that both equations (5) and (12) were robust in terms of observational noise.

Equation (12) is more robust than equation (5) when one of the time series does not change and is constant. When a time series does not change, then $C_{lt}$ and $C_{ot}$ become 0 and we cannot

**Figure 22.** Analysis using the multivariate amplitude adjusted Fourier transform surrogates [24] with the joint permutation entropy of length 3 as a test statistic. The thick solid line shows the value obtained from the original data and the histogram shows the values obtained from the surrogate data.

**Figure 23.** Analysis using the multivariate version of the truncated Fourier transform surrogates [25] with the joint permutation entropy of length 3 as a test statistic. The thick solid line shows the value obtained from the original data, and the histogram shows the values obtained from the surrogate data.
Figure 24. Randomizing spatial correlation of the foreign exchange markets by introducing a temporal delay to each of the five components. (a) and (b) show the results applying the delays of multiples of one hour. (c) and (d) show the results applying the delays of multiples of one week. In (a) and (c), we used equation (5). In (b) and (d), we used equation (12). In each panel, the range of $\chi^2$ or $\chi'^2$ is shown by a box plot: the range between 25% and 75% is shown by the box; the median is shown by the red horizontal line; the outliers are shown by plusses. ‘Simul’ and ‘Random’ correspond to the cases of applying the same simultaneous delays and random delays to different price time series components, respectively. The $p$-values obtained from the rank-sum tests were $8.1 \times 10^{-28}$, $1.3 \times 10^{-26}$, $3.6 \times 10^{-8}$, and $2.7 \times 10^{-20}$, respectively. Namely, the original time series consistently showed higher significance levels than those of different temporal shifts applied to different components for the foreign exchange markets.

Figure 25. Leading network $V_l$ versus the value of $\chi^2 (V_l)$ for the example of the coupled map lattice with 1% observational noise using equation (4). See the caption of figure 2 to interpret the results. The most significant leading network was the one with $y_1$, $y_2$, and $y_5$. 
define the index $I_t$ properly. Such events actually happened 278 times in the time series of foreign exchange markets. From this viewpoint, equation (12) is more desirable.

The results are not sensitive to the selection of threshold values. When we used 100 initial conditions to evaluate the effects of the size of threshold values, we found that the $p$-value based on equation (12) was smaller than the $p$-value obtained based on equation (5) in 91, 96, 83, and 74 out of 100 cases for the example of the coupled map lattice, and in 76, 76, 78, and 63 out of 100 cases for the example of the Lorenz’96 model, when the threshold values were controlled so that the recurrence rates became 0.05, 0.1, 0.2, and 0.3, respectively. These numbers also show the tendency for the proposed equation (12) to be more effective than the original equation (5).

Figure 26. The scatter plot between $\log_{10}(I_{t+1})$ and $\log_{10}(C_{t+1} + 1)$ for the coupled map lattice under 1% observational noise. The gray dash-dotted line shows the linear fitting for the scatter plot.

Figure 27. Leading network $V_l$ versus the value of $\chi^2 (V_l)$ for the example of the coupled map lattice with 1% observational noise using equation (11). See the caption of figure 2 to interpret the figure. The most significant leading network was the one with $y_4$ and $y_5$. 

New J. Phys. 16 (2014) 115015 S Oya et al
The superiority of equation (12) to equation (5) implies that delay coordinates help to reconstruct the information of the hidden part of a high-dimensional system including dynamical noise [32], deterministic driving force [33], and genuine time-varying parameters [34], so that we can forecast the future values for the observed part more accurately. When we use delay coordinates, we still have the important problem of how we should choose embedding dimensions. Possible methods include predictions [35, 36] and false nearest neighbors [37, 38]. We might be able to optimize the embedding dimension using the first half of the dataset as well, while we simply set $E = 2$ throughout the paper. This part is a topic for future research.

Although our examples here are limited to the ones in foreign exchange markets, the proposed methods themselves are general, and thus can be applied to wider contexts including the prediction of abrupt changes for renewable energy outputs [39–41]. These applications are...
in progress and we will write about them on another occasion; they should make it possible to introduce more renewable energy resources by keeping power grid systems stable.

In summary, we extended and applied the dynamical network marker proposed by Chen et al [3] to a single time series. We achieved statistically significant odds ratios. In addition, we extended the index of [3] by using recurrence plots so that we can apply the idea of the dynamical network marker to delay coordinates [6, 7] and point processes [8, 9]. The imaginary examples and the real example of foreign exchange markets showed that we tend to achieve higher probabilistic gains by using our extension. The proposed new index defined in equation (12) can provide better early warning signals given a stream of high-dimensional time series data.

**Figure 30.** The scatter plot between $\log_{10}(I_{t+1})$ and $\log_{10}(C_{t+1} + 1)$ for the Lorenz’96 model under 1% observational noise. The gray dash-dotted line shows the linear fitting for the scatter plot.

**Figure 31.** Leading network $V_l$ versus the value of $\chi^2(V_l)$ for the example of the Lorenz’96 model with 1% observational noise using equation (11). See the caption of figure 5 to interpret the results. The most significant leading network was the one with $v_{4,1}$ and $v_{5,1}$.
Acknowledgments

We thank the EBS Service Company Limited for providing the datasets of foreign exchange markets used in this study. The datasets are commercially available from the company. This research was partially supported by the Aihara Innovative Mathematical Modelling Project, the Japan Society for the Promotion of Science (JSPS) through its ‘Funding Program for World-Leading Innovative R&D on Science and Technology (FIRST Program)’, initiated by the Council for Science and Technology Policy (CSTP). This research was also supported by Core Research for Evolutional Science and Technology (CREST), Japan Science and Technology Agency (JST), and Platform for Dynamic Approaches to Living System from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

References

[1] Ahn J J, Oh K J, Kim T Y and Kim D H 2011 Expert Syst. Appl. 38 2966–73
[2] El-Shazly A 2011 Appl. Econ. 43 1817–28
[3] Chen L, Liu R, Liu Z-P, Li M and Aihara K 2012 Sci. Rep. 2 342
[4] Eckmann J-P, Kamphorst S O and Ruelle D 1987 Europhys. Lett. 4 973–7
[5] Marwan N, Romano M C, Thiel M and Kurths J 2007 Phys. Rep. 438 237–329
[6] Takens F 1981 Lecture Notes in Mathematics 898 366–81
[7] Sauer T, Yorke J A and Casdagli M 1991 J. Stat. Phys. 65 579–616
[8] Victor J D and Purpura K P 1997 Network 8 127–64
[9] Suzuki S, Hirata Y and Aihara K 2010 Int. J. Bifurcat. Chaos 20 3699–708
[10] Kaneko K 1984 Prog. Theor. Phys. 72 480–6
[11] Lorenz E N 1996 Proc. of Seminar on Predictability, ECMWF 1 1–18
[12] Hansen J A and Smith L A 2000 J. Atmos. Sci. 57 2859–71
[13] Faure P and Korn H 1998 Physica D 122 265–79
[14] Thiel M, Romano M C, Read P L and Kurths J 2004 Chaos 14 234–43
[15] Thiel M, Romano M C and Kurths J 2004 Phys. Lett. A 330 343–9
[16] Hirata Y, Horai S and Aihara K 2008 Eur. Phys. J. Spec. Top. 164 13–22
[17] Feldhoff J H, Donner R, Donges J F, Marwan N and Kurths J 2012 Phys. Lett. A 376 3504–13
[18] Goswami B, Marwan N, Feulner G and Kurths J 2013 Eur. Phys. J. Spec. Top. 222 861–73

Figure 32. The scatter plot between $\log_{10}(I_0 + 1)$ and $\log_{10}(C_{t+1} + 1)$ for the Lorenz'96 model under 1% observational noise. The gray dash-dotted line shows the linear fitting for the scatter plot.
[19] Ouyang G X, Zhu X Y, Ju Z J and Liu H H 2014 IEEE J. Biomed. Health Inform. 18 257–65
[20] Hirata Y and Aihara K 2012 Physica A 391 760–6
[21] Addo P M, Billio M and Guegan D 2013 North American Journal of Economics and Finance 26 416–35
[22] Romano M C, Thiel M, Kurths J and von Bloh W 2004 Phys. Lett. A 330 214–23
[23] Theiler J, Eubank S, Longtin A, Galdrickian B and Farmer J D 1992 Physica D 58 77–94
[24] Schreiber T 2000 Physica D 142 346–82
[25] Nakamura T, Small M and Hirata Y 2006 Phys. Rev. E 74 026205
[26] Timmer J 1998 Phys. Rev. E 58 5153–6
[27] Amigó J M, Monetti R, Aschenbrenner T and Bunk W 2012 Chaos 22 013105
[28] Bravo E P, Aihara K and Hirata Y 2013 Chaos 23 043104
[29] Bandt C and Pompe B 2002 Phys. Rev. Lett. 88 174102
[30] Mantegna R N, Palágyi Z and Stanley H E 1999 Physica A 274 216–21
[31] Ohira T, Sazuka N, Marumo K, Shimizu T, Takayasu M and Takayasu H 2002 Physica A 308 368–74
[32] Muldoon M R, Broomhead D S, Huke J P and Hgger R 1998 Dyn. Stab. Syst. 13 175–86
[33] Stark J 1999 J. Nonlinear Sci. 9 255–332
[34] Hegger R, Kantz H, Matassini L and Schreiber T 2000 Phys. Rev. Lett. 84 4092–5
[35] Cao L, Mees A and Judd K 1998 Physica D 121 75–88
[36] Hirata Y, Suzuki H and Aihara K 2006 Phys. Rev. E 74 026202
[37] Boccaletti S, Valladares D L, Pecora L M, Geffert H P and Carroll T 2002 Phys. Rev. E 65 035204
[38] Garcia S P and Almeida J S 2005 Phys. Rev. E 72 027205
[39] Hirata Y, Yamada T, Takahashi J, Aihara K and Suzuki H 2014 Renew. Energy 67 35–9
[40] Hirata Y 2014 Phys. Rev. E 89 052916
[41] Hirata Y, Aihara K and Suzuki H 2014 Eur. Phys. J. Spec. Top. in press doi:10.1140/epjst/e2014-02210-3