Invited Paper

Network structure detection using convergent cross mapping on multivariate time series

Kazuya Sawada¹ᵃ⁻, Yutaka Shimada², and Tohru Ikeguchi¹,³

¹ Department of Management Science, Graduate School of Engineering, Tokyo University of Science, 6-3-1 Niijuku, Katsushika, Tokyo 125-8585, Japan
² Department of Information and Computer Sciences, Graduate School of Science and Engineering, Saitama University, 255 Shimo-Okubo, Sakuraku, Saitama-shi, Saitama 338-8570, Japan
³ Department of Information and Computer Technology, Faculty of Engineering, Tokyo University of Science, 6-3-1 Niijuku, Katsushika, Tokyo 125-8585, Japan

ᵃ) sawada@hisenkei.net

Received January 31, 2020; Revised May 26, 2020; Published October 1, 2020

Abstract: The recent developments in measurement techniques have allowed us to observe multidimensional time series data in various fields. Thus, detecting causal relations between elements from multidimensional time series data is useful for prediction, model generation, and system control. In addition, causal relations between elements can be detected to estimate network connections. In other words, the network structure can be estimated from multidimensional time series data by using causality detection methods. Among the several methods for detecting causality, Granger causality is a well-known method that is widely used for causal estimation between time series. On the other hand, a method called convergent cross mapping (CCM) has been proposed, which can distinguish causality from pseudo-correlation. It is important to investigate whether CCM will be effective with increased number of elements, even though the evaluation of its performance with a few elements has been reported in the literature. Moreover, it is important to evaluate the performance with changing dynamics of the elements. In this study, to estimate the connectivity between elements in complex networks, we apply CCM to mathematical models of complex networks, or the Watts-Strogatz model. In particular, we investigate how complex network structures affect causal estimation, by applying CCM to multidimensional time series data produced from complex networks. According to the results, we find the connectivity estimation accuracies in the regular ring-lattice network to be slightly higher than those in random networks. Furthermore, we reveal that it is easier to perform connectivity estimation for a network with a community structure than a random structure.

Key Words: nonlinear time series analysis, complex network, causality, connectivity
1. Introduction

Causal relations exist between various events, and thus, it is important to identify them. In recent years, with the advances in observation technology and information science, considerable time series data can be measured easily. Thus, various methods have been proposed for estimating and detecting causality for time series data. For example, Granger causality is a commonly used method in causal estimation between time series signals [1]. However, several problems remain. Although Granger causality works well for a certain type of time series, it might not be effective for cases in which time series signals are generated from nonlinear deterministic dynamical systems.

On the other hand, a method called convergent cross mapping (CCM) [2] has been proposed as an effective method for estimating the causal relation between nonlinear deterministic time series signals. In [2], it was shown that CCM is effective for detecting the causality for a simple network with relatively few vertices. However, in the real world, the causality between events might be produced from complex networks and real-world networks are known to have a complex network structure [3, 4], for example, neural networks, power grids, the Internet, friendships, the spread of disease, and the spread of information. Estimating these complex connections is generally a difficult problem, but an important topic. That is, we need to evaluate the CCM performance for complex networks with relatively more vertices. Then, our purpose is two-fold. The first one is to investigate the ability of CCM to perform causality detection and coupling estimation in complex networks with relatively more vertices. The second one is to investigate whether the performance of coupling estimation by CCM depends on the network structure. Therefore, this study investigates how the causality detection accuracy resulting from the difference in the network structure appears when component elements in a network are connected by a complex network model.

In addition, coupling estimation using CCM has some problems. In CCM, after we mutually predict time series in-sample through cross mapping, we need to manually decide the threshold that determines whether a coupling exists. Therefore, we propose the use of CCM in combination with the Otsu method [5]. The proposed method can automatically decide the threshold for coupling estimation.

2. Method

Let us refer to a time series that influences as “the time series of cause,” and one that is influenced as “the time series of result.” The information in the time series of cause can be accumulated in the time series of result, and as the length of the time series increases, the amount of information of the time series of cause accumulated in the time series of result increases. In fact, such a concept is used in this paper as a common basis for causality estimation methods of CCM [2]. In this section, we review CCM, or a causality estimation method, based on the nonlinear dynamical systems theory.

The basic idea of CCM is that the higher the mutual prediction accuracy of a time series, the higher the possibility that causality exists. First, we consider a dynamical system comprising two variables, \( x_t \) and \( y_t \), where \( t \) is time. Let \( x_t \) be a cause of \( y_t \). Originally, CCM estimates the causality between multidimensional time series data, but for simplicity, we use two variables here.

When applying CCM to two time series data \( x_t \) and \( y_t \), we first reconstruct the attractors using a time-delay coordinate system [6] from the original time series \( x_t \) and \( y_t \). Embedding can be achieved by transforming an observed time series into a time-delay coordinate with a sufficiently large dimension [7]. Even if we cannot observe all variables, we can have the information on the original dynamical system only from an observed time series by reconstructing the attractors.

Let us define a reconstructed attractor from the time series of \( x_t \) by a time-delay vector \( \mathbf{x}(t) \), as shown in Eq. (1).

\[
\mathbf{x}(t) = \left( x_t, x_{t+\tau}, x_{t+2\tau}, \cdots, x_{t+(m-1)\tau} \right),
\]

where \( \tau \) is the time delay and \( m \) is the dimension of the reconstructed state space. Similarly, \( \mathbf{y}(t) \) is defined by

\[
\mathbf{y}(t) = \left( y_t, y_{t+\tau}, y_{t+2\tau}, \cdots, y_{t+(m-1)\tau} \right).
\]
Next, to predict the original time series \( x_t \) and \( y_t \), we use the above-mentioned time-delay vectors \( x(t) \) and \( y(t) \). This method is called cross mapping because \( y_t \) is predicted using \( x(t) \) and \( x_t \) is predicted using \( y(t) \). Now, we define \( x_\hat{t} \) and \( y_\hat{t} \) as the predicted time series. In the cross mapping, \( y_\hat{t} \) is calculated by

\[
\hat{y}_t = \sum_{i=1}^{K} w_{i\ell} \ y_{\ell}.
\]

Here, \( K \) is the number of neighbors for predicting the time series. We take the \( K \) neighbors of \( x(t) \), which are \( x(t^1), x(t^2), \ldots, x(t^K) \), where \( x(t^i) = (x_{t^i}, x_{t^i+\tau}, x_{t^i+2\tau}, \ldots, x_{t^i+(m-1)\tau}) \). Next, we take the corresponding \( y(t^1), y(t^2), \ldots, y(t^K) \), where \( y(t^i) = (y_{t^i}, y_{t^i+\tau}, y_{t^i+2\tau}, \ldots, y_{t^i+(m-1)\tau}) \) and predict \( \hat{y}_t \) by Eq. (3). In what follows, we basically use \( K = m + 1 \) in numerical experiments. However, when investigating the effects of the number of neighbors \( K \) on predictability in numerical experiments, we use \( K = m + 1 \) or \( 2m \). In Eq. (4), \( w_{i\ell} \) is defined by

\[
w_{i\ell} = \frac{u_{i\ell}}{\sum_{j=1}^{K} u_{j\ell}},
\]

where \( w_{i\ell} \) is the weight calculated from the time-delay vector of the \( i \)th nearest neighbor \( x(t^i) \) in the reconstructed state space \( x(t) \). In Eq. (4), \( u_{i\ell} \) is defined by

\[
u_{i\ell} = \exp \left\{ -\frac{d[x(t), x(t^i)]}{d[x(t), x(t^i)]} \right\},
\]

where \( d[x(t), x(s)] \) is the Euclidean distance between \( x(t) \) and \( x(s) \) and \( x(t^i) \) is the \( i \)th nearest neighbor of the time-delay vector \( x(t) \) in the reconstructed state space. When \( d[x(t), x(t^i)] = 0 \), \( u_{i\ell} \) is defined by

\[
u_{i\ell} = \begin{cases} 1 & (i = 1), \\ 0 & (i \neq 1). \end{cases}
\]

Then, using Eqs. (3)–(6), the predicted state \( \hat{y}_t \) is obtained:

\[
\hat{y}_t = \sum_{i=1}^{K} \sum_{j=1}^{K} \exp \left\{ -d[x(t), x(t^i)]/d[x(t), x(t^i)] \right\} y_{\ell} \cdot
\]

The accuracy of cross mapping is evaluated by the correlation coefficient \( \rho_{xy} \) between the actual time series \( y_t \) and \( \hat{y}_t \) obtained by Eq. (7). Namely, \( \rho_{xy} \) is defined by

\[
\rho_{xy} = \frac{\sum_{i=1}^{T} (y_i - \bar{y})(y_\hat{i} - \bar{y}_\hat{i})}{\sqrt{\sum_{i=1}^{T} (y_i - \bar{y})^2 \sqrt{\sum_{i=1}^{T} (y_i - \bar{y}_\hat{i})^2}}.
\]

where \( T \) is the length of time series and \( \bar{y} \) is the average of \( y_t \). Here, let us assume that \( y_t \) is affected by \( x_t \) and \( x_t \) is not affected by \( y_t \) (\( x_t \) is cause and \( y_t \) is result). In this case, the correlation coefficient \( \rho_{xy} \) between \( y_t \) and \( \hat{y}_t \) predicted from \( x(t) \) is expected to take a low value. On the other hand, the correlation coefficient \( \rho_{yx} \) between \( x_t \) and \( \hat{x}_t \) predicted from \( y(t) \) is expected to be a high value. This is because \( x_t \) does not contain the information of \( y_t \) but \( y_t \) contains the information of \( x_t \). From this fact, we estimate the existence and direction of causality from \( y_t \) to \( x_t \). The prediction accuracy of the time series increases and converges when the time series length is increased while performing the cross mapping. Therefore, it is called convergent cross mapping [2].
3. Network models and time series models

3.1 Network models

In this study, we investigate how complex network structures affect the accuracy of connectivity estimation by CCM. Then, we use two complex network structures.

The first one is generated from the Watts–Strogatz model [8], which is a network model that can easily reproduce the small-world structure [8]. We use a network with 20 and 100 vertices generated from the WS model in numerical experiments. The reason why we use the WS model is that it is one of the most famous network models that can change the network randomness by changing the rewiring probability. We first create a regular ring-lattice network of arbitrary average degree when creating a network using the WS model. Here, we set the average degree to four in case of 20 vertices, and six in case of 100 vertices. Next, we rewire the edges of a regular ring-lattice network by changing the rewiring probability \( p \). When \( p = 1 \), it is a random network.

The second one is a network with community structure. As shown in Fig. 1, we create two networks: a network with a community structure (Fig. 1(a)) and that with a random structure (Fig. 1(b)). Here, we describe how to create a network with a community structure. When we create such a network, we first create two complete graphs of 11 vertices. Next, we select two edges from each complete graph and create a network with a community structure by rewiring these edges. In addition, we use the following procedure to create a fixed-degree random network. We first create a regular ring-lattice network with 22 vertices and an average degree of 10, using the WS model [8]. Next, we rewire the edges randomly so that the degree of each vertex of the created network does not change [9].

![Fig. 1. Networks with (a) community structure and (b) random structure.](image)

Creating these two complex network structures, we introduce two nonlinear dynamical systems for the vertex dynamics in the networks. The first nonlinear dynamical system is a coupled logistic map described in Sec 3.2 and the second one is a chaotic neural network described in Sec 3.3.

The reason why we use these dynamical systems is two-fold. The first reason is that [2] used a coupled logistic map consisting of two logistic maps to evaluate the CCM. That is, [2] did not try more than two logistic maps to evaluate the CCM. However, as it is well-known that the relation among more than three elements is complex, it is important to evaluate the CCM on more than three vertices with the same dynamics. The second reason is that functional shapes of elements used in these dynamical systems are different: the logistic map is a unimodal map, while a chaotic neuron is a bimodal map. Namely, it is important to evaluate the performance of CCM when the dynamics changes.

3.2 Coupled logistic map

The coupled logistic map [10, 11] is defined by Eq. (9), where \( z_i(t) \) is the state value of the vertex \( i \) at time \( t \), \( \varepsilon \) is the coupling strength between vertices, \( k_i \) is the degree of the vertex \( i \), \( A_{ij} \) is the \((i, j)\)th component of the adjacency matrix obtained from the WS model, and \( N \) is the number of vertices in the coupled logistic map. The function \( f \) describes a logistic map: \( f(\varphi) = a\varphi(1 - \varphi) \), where \( a \) is a parameter. In this paper, we use \( a = 4.0 \) and \( \varepsilon = 0.1 \).
\[ z_i(t + 1) = (1 - \varepsilon) f(z_i(t)) + \frac{\varepsilon}{k_i} \sum_{j=1}^{N} A_{ij} f(z_j(t)). \] (9)

The reason why we set \( \varepsilon = 0.1 \) is that the state values could be a periodic solution depending on the value of \( \varepsilon \). If the state value is a periodic solution, the prediction obtained with the cross mapping is almost unity, even if there is no causality between the vertices. Here, a bifurcation diagram visualizing the relation between \( \varepsilon \) and the state values when we use the WS model on \( p = 0 \) is shown in Fig. 2.

![Fig. 2. Bifurcation diagram of the coupled logistic map \((a = 4.0)\) on the WS model with the rewiring probability \( p = 0 \).](image)

Figure 2 shows the values of \( z_i(1 \leq i \leq 20) \) as the \( y \)-axis; we exclude 100,000 transient states and use 100 as the time series length. Figure 2 shows that the state value partially becomes a periodic solution when \( \varepsilon \) is more than 0.5. Therefore, we use the parameter \( \varepsilon = 0.1 \) with which the state value does not become a periodic solution.

### 3.3 Chaotic neural network

The chaotic neural network [12] is defined by Eq. (10), where \( v_i(t) \) is a state value of the chaotic neuron \( i \) at time \( t \), \( \kappa_i \) is the refractory time decay constant, \( \alpha_i \) is a coefficient for refractory terms, \( a_i \) is a bias, \( N \) is the number of neurons in the network, and \( w_{ij} \) is the coupling strength between the neurons \( i \) and \( j \) \( (i, j = 1, 2, \cdots, N) \). The function \( g \) is a sigmoid function \( g(\psi) = \frac{1}{1 + \exp(-\psi/\varepsilon')} \), where \( \varepsilon' \) is a parameter. In this paper, we use \( \kappa_i = 0.8, \alpha_i = 1.0, a_i = 0.2, \varepsilon' = 0.04, \) and \( w_{ij} = 0.2 \) \( (i \neq j) \). When \( i = j, w_{ij} = 0 \).

\[ v_i(t + 1) = \kappa_i v_i(t) - \alpha_i g(v_i(t)) + a_i + \sum_{j=1}^{N} w_{ij} g(v_j(t)) \] (10)

The reason why we set \( w_{ij} = 0.2 \) is that the state value could be a periodic solution depending on the value of \( w_{ij} \). Here, a bifurcation diagram visualizing the relation between \( w_{ij} \) and the state values when we use the WS model of \( p = 0 \) is shown in Fig. 3.

![Fig. 3. Bifurcation diagram of the chaotic neural network.](image)
In Fig. 3, we show the values of \( v_i \) (1 \( \leq \) \( i \) \( \leq \) 20) as the \( y \)-axis, exclude 100,000 transient states and use 100 as the time series length. This figure indicates that the state value becomes a periodic solution when the \( w_{ij} \) is greater than 0.2. Therefore, we use a parameter whose state value does not become a periodic solution, \( w_{ij} = 0.2 \).

4. Numerical experiments

4.1 Estimation of connectivity for the coupled logistic map and the chaotic neural network

In this section, we explain the results of the numerical experiments. First, we use the coupled logistic map of Eq. (9) as the dynamics of each vertex of the WS model. Then, we change the edge rewiring probability \( p \) in the WS model and generate ten networks for each edge rewiring probability. We vary \( p \) from 0.0 to 1.0, in steps of 0.05. Next, we apply CCM to the time series generated from the network and estimate the connectivity between the vertices. To evaluate the accuracy of connectivity estimation, we calculate the average values of accuracies for 10 networks. Then, we can obtain 400 prediction accuracies for a single network because it has 20 vertices, and we predict the time series mutually in CCM. With sufficient data points in the time series, we obtain relatively high prediction accuracies in the case where vertices are connected and low accuracy in the case where the vertices are not connected. In other words, when a sufficient length of the time series is used, the distribution of the prediction accuracies is expected to be bimodal.

In connectivity estimation, the vertices are found to be connected if the prediction accuracy is greater than the threshold, and not connected if the prediction accuracy is less than the threshold. Then, we use the Otsu method for deciding the threshold of the prediction accuracies. The Otsu method is an effective method for deciding the threshold in cases where the distribution is bimodal. Normally, we can obtain an appropriate threshold of connectivity estimation using the Otsu method. In the Appendix, we describe the algorithm of the Otsu method.

In the following numerical experiments, the length of the time series used in the CCM is set to 4,000, the dimension \( m \) of the reconstructed state space is two, and \( \tau \) is unity. Figure 4 shows the relation between the edge rewiring probability and the average of the connectivity estimation accuracy in a network generated from the WS model. In Fig. 4, the purple line with circles indicates the average values of the estimation accuracy, while the gray-shaded area indicates 95% confidence interval when we use the \( t \)-distribution.

![Fig. 4. Relation between the edge rewiring probability \( p \) and estimation accuracy of connectivity when dynamics of vertices are the coupled logistic map (Eq. (9)) in case of 20 vertices. The gray-shaded area indicates 95% confidence interval when we use the \( t \)-distribution.](image)

Figure 4 shows that the estimation accuracies of connectivity with varying edge rewiring probability take higher values than 0.8. In addition, the connectivity estimation accuracies in a regular ring-lattice network are relatively higher than those in random networks. However, the estimation accuracies of connectivity take almost the same value even when the edge rewiring probability in the WS model is changed. The confidence intervals for the random network are wider than those for regular ring-lattice networks, and the variance in the accuracy of connectivity estimation is large.
Next, we perform a similar numerical experiment for the 100-vertex case. The average degree is set to six and the other parameters are the same as those in the case of 20 vertices. Figure 5 shows the relation between the edge rewiring probability and the average of the connectivity estimation accuracy. Figure 5 shows that the accuracy of connectivity estimation is better than about 0.8 for all edge rewiring probabilities. In addition, the overall connectivity estimation accuracy is slightly lower for random networks than for regular networks, and the confidence intervals are wider. This may be because as \( p \) increases, high-degree vertices emerged. Furthermore, the characteristics of the results are not much different from those of the case with 20 vertices.

![Fig. 5. Relation between the edge rewiring probability \( p \) and estimation accuracy of connectivity when the dynamics of the vertices are the coupled logistic map (Eq. (9)) in case of 100 vertices. The gray-shaded area indicates 95% confidence interval when we use the \( t \)-distribution.](image)

Next, we apply the CCM to the chaotic neural network (Eq. (10)). Figure 6 shows the results of the numerical experiments conducted for the chaotic neural network. As shown in Fig. 6, the prediction accuracies take values between 0.6 to 0.9, which indicates that the accuracies are totally lower than the results of the logistic map. However, as shown in Fig. 4, the connectivity estimation accuracies in the regular ring-lattice network are relatively higher than those in the random networks and the confidence intervals for a random network are wider than for regular ring-lattice networks. These tendencies are almost the same as the case of the coupled logistic map shown in Fig. 4.

![Fig. 6. Relation between the edge rewiring probability \( p \) and estimation accuracy of connectivity when the dynamics of vertices are the chaotic neural network (Eq. (10)). The gray-shaded area indicates 95% confidence interval when we use the \( t \)-distribution.](image)

4.2 Effects of goodness of predictors and edge degrees

In causal estimation by CCM, we detect causality based on mutual prediction accuracy. In other words, the higher the prediction accuracy of the time series, the higher is the possibility that causality exists which means that it is important to create a good predictor. In particular, when we predict the time series using CCM in this paper, we calculate the weights using the distance information in the reconstructed state space (Eq. (4)). If we use this method for calculating the weights, the prediction point should be surrounded by neighboring points. Thus, if the prediction accuracy increases with
the number of neighbors used for the prediction, the prediction point might not be surrounded by neighboring points. On the other hand, the prediction point is surrounded by neighboring points with a high probability when the prediction accuracy does not increase even with an increase in the number of neighboring points used for prediction. Therefore, we investigate the mutual prediction accuracies by changing the number of neighboring points used for time series prediction, and compared their accuracies.

In addition, it is highly likely that the mutual prediction accuracy decreases in the case where the vertex has many edges, because vertices with a high degree are affected by many other vertices and the created predictor is likely to be bad. Therefore, we calculate the average of the time series prediction accuracy of the vertices of individual degrees when we change the number of neighboring points. We call it the average prediction accuracy, which is defined by Eq. (11).

\[ \bar{\rho}_\delta = \frac{1}{N} \sum_{j=1}^{N} \sum_{i \in \Omega} A_{ij} \rho_{ij}. \]  

Eq. (11)

In Eq. (11), \( \bar{\rho}_\delta \) is the average prediction accuracy of the degree \( \delta \), \( \Omega \) is the set of indices of vertices defined by \( \Omega \equiv \{ i \mid k_i = \delta, i = 1, 2, \cdots, N \} \), \( k_i \) is the degree of vertex \( i \), \( A_{ij} \) is the \((i, j)\)th element of the adjacency matrix, and \( \rho_{ij} \) is the prediction accuracy of vertex \( j \) predicted by the time series observed from vertex \( i \).

To surround the predicted points in an \( m \)-dimensional state space, we need at least \( m + 1 \) points. Namely, we need at least \( K = m + 1 \) neighboring points when the dimension of the reconstructed state space is \( m \) [2]. Figure 7 shows the relation between \( \bar{\rho}_\delta \) and degree \( \delta \) when the number of neighboring points used for prediction \( K \) is \( m + 1 \) and \( 2m \). In Fig. 7, we use ten random networks (\( p = 1 \)) with 20 vertices and an average degree of four. In addition, we change the dimension of the reconstructed state space for prediction in the range of \( 2 \leq m \leq 5 \).

![Fig. 7. Relation between degree \( \delta \) and average prediction accuracy \( \bar{\rho}_\delta \) on ten random networks \( p = 1 \).](image)

(a) The number of neighbors \( K = m + 1 \) \((2 \leq m \leq 5)\).  
(b) The number of neighbors \( K = 2m \) \((2 \leq m \leq 5)\).

As shown in Fig. 7, the time series prediction accuracy decreases as the degree increases. In addition, from Figs. 7(a) and (b), there is no significant change in the prediction accuracy even when the number of neighboring points used for the prediction and the reconstruction dimension are changed. Therefore, it is sufficient that the number of neighboring points is \( m + 1 \) in the numerical experiments using the coupled logistic map. Then, we set the number of neighboring points \( K \) to \( m + 1 \) in the numerical experiments conducted for the coupled logistic map. However, in both Figs. 7(a) and (b), the prediction accuracy is slightly higher when the dimension of reconstruction is three. To make this result easier to understand, we replace the horizontal axis \( \delta \) and legend \( m \) in Fig. 7. This is shown in Fig. 8. In other words, Fig. 8 shows the relation between the average prediction accuracy \( \bar{\rho}_\delta \) and the reconstruction dimension \( m \) for each degree \( \delta \).

Figure 8 shows that although the average prediction accuracies \( \bar{\rho}_\delta \) take almost the same value even when the dimension of the reconstructed state space is changed, the average prediction accuracies \( \bar{\rho}_\delta \) are slightly improved in the case of \( m = 3 \). Although we investigate this reason in various ways, we cannot clarify the reason. Thus, we will investigate the reason for a future work. Figures 7 and 8 shows
that the mutual prediction accuracy for vertices with a higher degree is low. Vertices with higher
degrees receive information from many other vertices, which makes it difficult to predict. In other
words, the accuracy of causality detection using CCM depends on how the time series are affected.

4.3 Connectivity of community structure
In this section, using CCM, we compare the difference in the connectivity estimation accuracy between
a network with a community structure and a random network with the same number of vertices and
edges. We use the network shown in Fig. 1. Figure 9 shows the results of comparing the prediction
accuracy. The connectivity estimation accuracy of a network with a community structure is higher
than 0.8, whereas that of a random network is less than 0.6. Since the degrees of all vertices in the
two networks are the same, the degree has no effect on the connectivity estimation accuracy. In other
words, it is relatively easier to estimate a network having a community structure than a random
network when we perform connectivity estimation by CCM.

5. Conclusion
We performed the connectivity estimation of networks generated from complex network models using
CCM [2], which is a method for estimating causality. Here, causality means connectivity, and thus,
estimating causality can be considered as estimating connectivity. In particular, we investigated
whether there is a difference in the accuracy of connectivity estimation by CCM when the edge
rewiring probability of the WS model is changed. Thus, we used two nonlinear dynamical systems, a logistic map and a chaotic neuron model, as the network vertices.

As a result, we found that the connectivity estimation accuracies in the ring-lattice regular network were slightly higher than those in random networks. In addition, the variance in the accuracy of connectivity estimation in random networks was larger than that in regular ring-lattice networks. We also found that it is difficult to estimate the connectivity for high-degree vertices by CCM. This result supports the fact that the variance in the accuracy of connectivity estimation in a random network is large.

Furthermore, we compared the accuracies of connectivity estimation in a network with communities and its randomized networks with the same conditions (the numbers of vertices, edges, and degrees). As a result, we found that it is easier to perform connectivity estimation for a network with a community structure than that for random networks. In other words, these results imply that networks with a community structure have less disturbed information of time series than random networks.

In addition, to estimate the coupling using CCM, we need to manually decide the threshold that determines whether a coupling exists. However, the appropriate value of the threshold varies depending on the conditions. Therefore, in this paper, we proposed a method to estimate the network structure by automatically determining the threshold of the prediction accuracy and detecting the causality by combining the Otsu method and CCM. The results of the numerical experiments show that the coupling estimation accuracy obtained by the proposed method works well, which implies that it is a useful method.

In this paper, the WS model, which can easily change the network randomness, was used as a complex network model. However, it is also important to use other network models, such as the BA model [13]. Then, it is an important future work to investigate the performance of CCM using several network models including the BA model.

In addition, we would like to extend the method so that it is possible to estimate the network structure even if there are unobserved time series data, because it is not often the case that all the time series data generated from a network are observed in the real world. Therefore, it is also an important issue to analyze the case where we cannot observe all variables from target dynamical systems.

Acknowledgments
This research is partially supported by the JSPS Grant-in-Aids for Scientific Research (Nos. 15KT0112, 17K00348, 18K18125).

Appendix
In this appendix, we describe the Otsu method. The Otsu method of binarization is a discriminant analysis method often used in the field of image processing [5]. In this paper, the estimated adjacency matrix is generated by applying Otsu binarization to the prediction accuracy matrix obtained by CCM. In case that the time series is obtained from connected vertices, the prediction accuracy of time series signals obtained by CCM converges to a high value if the length of the time series signals used for prediction is sufficient. On the other hand, for unconnected vertices, the prediction accuracy of time series signals obtained by CCM converges to 0 when the length of the time series signals used for the prediction is sufficient. Therefore, if this length is sufficiently long, the prediction accuracy of the time series signals obtained by CCM is expected to be a bimodal distribution. Therefore, the Otsu binarization method is effective in estimating the causality from the prediction accuracy of time series signals obtained by CCM. The algorithm of the Otsu binarization method is as follows.

\begin{enumerate}
\item The average value \( \mu_0 \), maximum value \( I_{\text{max}} \), and minimum value \( I_{\text{min}} \) of all obtained prediction accuracies are calculated.
\item The threshold \( \theta \) is defined in the range of \( I_{\text{min}} \leq \theta \leq I_{\text{max}} \), and the data are classified into two classes, class 1 and class 2, by the threshold \( \theta \).
\end{enumerate}
The number of data in class 1, \( n_1 \), the average value of data in class 1 \( \mu_1 \), the variance of data in class 1 \( \sigma_1^2 \), the number of data in class 2, \( n_2 \), the average value of data in class 2 \( \mu_2 \), and the variance of data in class 2 \( \sigma_2^2 \) are calculated.

The within-class variance \( \sigma_w^2 \), between-class variance \( \sigma_b^2 \), and separation rate \( S \) are calculated.

\[
\sigma_w^2 = \frac{n_1 \sigma_1^2 + n_2 \sigma_2^2}{n_1 + n_2}, \tag{A-1}
\]
\[
\sigma_b^2 = \frac{n_1 (\mu_1 - \mu_0)^2 + n_2 (\mu_2 - \mu_0)^2}{n_1 + n_2}, \tag{A-2}
\]
\[
S = \frac{\sigma_b^2}{\sigma_w^2}. \tag{A-3}
\]

Steps (2) to (4) are repeated by changing the threshold value \( \theta \) in the range of \( I_{\text{min}} \leq \theta \leq I_{\text{max}} \). Then, the threshold value, when the separation rate \( S \) takes the maximum value \( S_{\text{max}} \), is used for discrimination of two classes: coupled and uncoupled.

References

[1] C.W.J. Granger, “Investigating causal relations by econometric models and cross-spectral methods,” *Econometrica: Journal of the Econometric Society*, vol. 37, no. 3, pp. 424–438, 1969.

[2] G. Sugihara, R. May, H. Ye, C. Hsieh, E. Deyle, M. Fogarty, and S. Munch, “Detecting causality in complex ecosystems,” *Science*, vol. 338, no. 6106, pp. 496–500, 2012.

[3] L. da F. Costa, F.A. Rodrigues, G. Travieso, and P.R. Villas Boas, “Characterization of complex networks: A survey of measurements,” *Advances in Physics*, vol. 56, no. 1, pp. 167–242, 2007.

[4] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, “Complex networks: Structure and dynamics,” *Physics Reports*, vol. 424, nos. 4–5, pp. 175–308, 2006.

[5] N. Otsu, “A threshold selection method from gray-level histograms,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[6] N. Packard, J. Crutchfield, D. Farmer, and R. Shaw, “Geometry from a time series,” *Physical Review Letters*, vol. 45, no. 9, pp. 712–716, 1980.

[7] F. Takens, “Detecting strange attractors in turbulence,” *Lecture Notes in Mathematics*, pp. 366–381, 1981.

[8] D.J. Watts and S.H. Strogatz, “Collective dynamics of ‘small-world’ networks,” *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.

[9] S. Maslov and K. Sneppen, “Specificity and stability in topology of protein networks,” *Science*, vol. 296, no. 5569, pp. 910–913, 2002.

[10] A.L. Lloyd, “The coupled logistic map: a simple model for the effects of spatial heterogeneity on population dynamics,” *Journal of Theoretical Biology*, vol. 173, no. 3, pp. 217–230, 1995.

[11] S. Jalan and A. Singh, “Cluster synchronization in multiplex networks,” *EPL (Europhysics Letters)*, vol. 113, no. 3, p. 30002, 2016.

[12] K. Aihara, T. Takabe, and M. Toyoda, “Chaotic neural networks,” *Physics Letters A*, vol. 144, nos. 6–7, pp. 333–340, 1990.

[13] A.L. Barabási and R. Albert, “Emergence of scaling in random networks,” *Science*, vol. 286, no. 5439, pp. 509–512, 1999.