Chaos, Solitons and Fractals 82 (2016) 11–21

Contents lists available at ScienceDirect

Chaos, Solitons and Fractals

Nonlinear Science, and Nonequilibrium and Complex Phenomena

journal homepage: www.elsevier.com/locate/chaos

Frontiers

Choosing the optimal model parameters for Granger causality in application to time series with main timescale

Maksim V. Kornilov [a,b,*], Tatiana M. Medvedeva [a], Boris P. Bezruchko [a,b], Ilya V. Sysoev [a,b]

[a] Saratov State University, 83, Astrakhanskaya street, 410012 Saratov, Russia
[b] Saratov Branch of Kotel'nikov Institute of Radioengineering and Electronics of Russian Academy of Sciences, 38, Zelyonaya street, 410019 Saratov, Russia

ARTICLE INFO

Article history:
Received 2 September 2015
Accepted 20 October 2015
Available online 18 November 2015

Keywords:
Coupling identification
Granger causality
Time series

ABSTRACT

The problem of determining the presence and direction of coupling between experimentally observed time series is of immediate interest in many relevant areas of knowledge. One of the approaches to its solution is the method of nonlinear Granger causality. The algorithm is based on the construction of predictive models and its effectiveness depends on the proper selection of model parameters.

The most important of them for signals with a characteristic time scale fluctuations are the time lag used in the reconstruction of the state vector, and the range forecast. In this paper, we propose two criteria for evaluating performance of the method of nonlinear Granger causality, which allows one to select the lag and range forecast and achieves the best sensitivity and specificity. The sensitivity is determined by range of weakness the method can detect and specificity means the ability to avoid false positive results. Because of the proposed criteria on the example of several unidirectionally coupled reference systems were received practical advice on the selection of the following model parameters: lag and range forecast.

1. Introduction

Detection of the presence and direction of interactions between subsystems of complex systems from their experimentally observed time series is an up-to-date problem, which has applications in various fields of knowledge. A variety of methods were developed in order to solve this task, including the cross-correlation function, coherence, phase synchronisation index, information based measures [1,2], a partial directed coherence [3] and approaches based on building predictive models, which include the Granger causality approach [4], and phase dynamics approach [5,6]. The main idea of approaches Granger causality method, transfer entropy and partial directed coherence is similar, and in some cases they can be shown to be completely equal [7]. However in general this is not the case [8].

The author developed Granger causality in relation to economical studies, but now it is successfully applied, for example, in neuroscience to identify the coupling between different brain regions [9–14], in climatology—to predict the behaviour of the monsoon [15]. There are a number of studies, where the Granger causality is used to indicate the evolution of coupling in time [16–18].

Despite the transparency of the idea of the method, its efficiency depends critically on the details of the implementation. For example, the special attention has to be paid to effects of measurement noise [19–21].

* Corresponding author at: Saratov State University, 83, Astrakhanskaya street, 410012 Saratov, Russia. Tel.: +7 9173085778.
E-mail addresses: kornilovmv@gmail.com (M.V. Kornilov), golovatanya@rambler.ru (T.M. Medvedeva), bezruchkobp@gmail.com (B.P. Bezruchko), ivssci@gmail.com (I.V. Sysoev).

http://dx.doi.org/10.1016/j.chaos.2015.10.027
0960-0779/

In this paper, we propose a new approach to choose the method parameters for the case, when the measured time series have a single main time scale which can be detected, for example, by using auto-correlation function. Such systems are quite common in nature, e.g. rhythmic changes in solar activity (sunspot number), oscillatory changes of brain potentials (for example, spike-wave discharges at the absence epilepsy, or θ-rhythm during the sleep), the signals of the heart (the main heart rate at the cardiogram) etc. The idea of the approach is that the parameters of the model, with the meaning of time: the time lag l, used for the reconstruction of the state vector [22–24], and the prediction time τ should be associated with the characteristic scale of the observed oscillations. Some research in this direction has already been held: in [25] it was shown that the lack of sample rate and, therefore, too large values of τ and l lead to systematic, fatal errors in determining the direction of coupling, primarily to the appearance of false positive conclusions. Similarly, in the paper [26] it was shown that mistakes can be caused by too small values of the prediction time, therefore it was proposed to use the value of τ equal to a quarter of the characteristic period of the observed oscillations.

In general, however, the question of the optimal choice of τ and l is still open, even for individual narrow classes of signals, including the class considered in this paper, since the previously used numerical criteria took into account only one parameter of the Granger causality method: either prediction time, as in [26], or sampling interval, as in [25], or a kind of approximation functions, as in [27]. Therefore, in this paper we propose two new numerical criteria characterising the performance of the method according to τ and l and test them in a number of nonlinear coupled reference systems. The application of these criteria helped to make general conclusions about the best and worst values of τ and l for the models used in the method of Granger causality to determine the coupling of signals having a distinct time scale (oscillation period). The degree of efficiency of the method is shown according to the level of nonlinearity of the original data, which is determined by the highest Lyapunov exponent and the effective coefficient of phase diffusion.

In the original study [4] linear approximating functions were used, but now nonlinear functions of different type: polynomials [27,28], radial basis functions [29] and kernel Granger causality [30] can be used instead. The other way to solve the problem of approximation function choice is to use local linear models, as it was proposed in [31]. We examined the model with polynomial nonlinearity of the general form and the local linear models, since they are most frequently used in practice due to simplicity and generality. In this regard, we chose the small values of both the model dimension and polynomial order. So, the models do not require many coefficients and can be reliably estimated from the short time series. This helps to make the results useful from a practical point of view, when the amount of data is very limited due to the features of the experiment, nonstationarity, or a desire for acceptable temporal resolution, when constructing models in a time window, as is done, for example, in [14,17,32].

2. Granger causality

Let us remind the key point of Granger causality. Supposing that we have time series of two systems—a series $\{x_n\}_{n=1}^N$ from the system X and a series $\{y_n\}_{n=1}^N$ from the system Y, where $n = 1, 2, \ldots, N$ is discrete time, N is the length of the series. It requires to determine whether the system Y drives the system X or not by analysing realisations $\{x_n\}_{n=1}^N$ and $\{y_n\}_{n=1}^N$. To solve this problem an individual model (dynamical system) is constructed on the first step:

$$x'_{n+\tau} = f(x_n, x_{n-l}, \ldots, x_{n-(D_s-1)l}, \mathbf{c}^s), \tag{1}$$

where x'_n is a predicted value at the time moment n and it may differ from the measured value x_n, f is an approximating function (if it is nonlinear, method is called a nonlinear Granger causality), l—lag of the model, i.e. the number of discrete time points between the two subsequent values from $\{x_n\}_{n=1}^N$, forming D_s-dimensional state vector of the model $\mathbf{x}_n(x_n, x_{n-1}, \ldots, x_{n-(D_s-1)l}$, τ is the prediction time—the distance in time between the predicted point and the closest point of the state vector, D_s—dimension of the individual model (the number of points of the time series which form the state vector, which is being reconstructed by the time delay method [23,24]), $\mathbf{c}^s$—unknown vector of coefficients which is chosen using least squares fit to minimise the standard error of approximation (1):

$$\varepsilon_s^2 = \frac{1}{N} \sum_{n=\tau+(D_s-1)l+1}^{N} (x'_n - x_n)^2 \tag{2}$$

The next step is to build the joint model, in which D_a members from the series $\{y_n\}_{n=1}^N$ are used besides the data of the series $\{x_n\}_{n=1}^N$:

$$x''_{n+\tau} = g(x_n, x_{n-l}, \ldots, x_{n-(D_s-1)l}, y_n, y_{n-l}, \ldots, y_{n-(D_a-1)l}, \mathbf{c}^j), \tag{3}$$

where x''_n is a model predicted value, $\mathbf{c}^j$—joint model coefficients. The standard prediction error of the joint model similarly to (2) has the form:

$$\varepsilon_j^2 = \frac{1}{N} \sum_{n=\tau+(\max(D_s,D_a)-1)l+1}^{N} (x''_n - x_n)^2. \tag{4}$$

If $\varepsilon_j^2 < \varepsilon_S^2$, the system Y is considered to drive the system X (systems are coupled). *Prediction improvement* index is typically used as a measure of coupling:

$$PI = 1 - \frac{\varepsilon_j^2}{\varepsilon_S^2}. \tag{5}$$

If $PI = 0$ (considering the signal $\{y_n\}_{n=1}^N$ did not help in predicting $\{x_n\}_{n=1}^N$), it is considered that Y has no effect on X. If the $PI \to 1$ (considering the signal $\{y_n\}_{n=1}^N$ has significantly improved the prediction of $\{x_n\}_{n=1}^N$), it should be regarded as Y drives X.

Practice shows that the choice of the parameters of the described procedure (lag l, prediction time τ, dimensions D_s and D_a, type of nonlinear functions f and g) significantly determines the efficiency of the method. For example, the use of too small or too large τ may cause a large number of errors: positive conclusions about the coupling that in fact does not exist [25,26]. Neglecting the nonlinearity in the modelling often leads to a situation, when really existing links are not detected [27,33]. The problem seems to be major since most coupling analysis techniques are very model-dependent [8].