NeuroImage 181 (2018) 382–394

Contents lists available at ScienceDirect

NeuroImage

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

Nonlinear effective connectivity measure based on adaptive Neuro Fuzzy Inference System and Granger Causality

Mona Farokhzadi [a], Gholam-Ali Hossein-Zadeh [a,b], Hamid Soltanian-Zadeh [a,b,c,*]

[a] *Control and Intelligent Processing Center of Excellence (CIPCE), School of Electrical and Computer Engineering, University College of Engineering, University of Tehran, Tehran, Iran*

[b] *School of Cognitive Sciences, Institute for Research in Fundamental Sciences (IPM), Tehran, Iran*

[c] *Radiology Image Analysis Laboratory, Henry Ford Health System, Detroit, MI, 48202, USA*

A R T I C L E I N F O

Keywords:
Granger causality
Nonlinear effective connectivity
ANFIS predictor
EEG/MEG

A B S T R A C T

Exploring brain networks is an essential step towards understanding functional organization of the brain, which needs characterization of linear and nonlinear connections based on measurements like EEG or MEG. Conventional measures of connectivity are mostly linear and bivariate. This paper proposes an effective connectivity measure called Adaptive Neuro-Fuzzy Inference System Granger Causality (ANFISGC). The proposed measure is based on the symplectic geometry embedding dimension, Adaptive Neuro-Fuzzy Inference System (ANFIS) predictor, and Granger Causality (GC). It is a powerful predictor that detects both linear and nonlinear causal information flow. It is not bivariate and thus can distinguish between direct and indirect connections. The performance of the proposed method is evaluated and compared with those of the Linear Granger Causality (LGC), Kernel Granger Causality (KGC), combination of Pairwise Granger Causality and Conditional Granger Causality (PwGC + CGC), Transfer Entropy (TE), and Phase Transfer Entropy (PTE) methods using simulated and experimental MEG data. Simulation results show that ANFISGC outperforms the other methods in detecting both linear and nonlinear connections and, by increasing the coupling strength between nodes, the value of ANFISGC increases. In the analysis of the time series of the brain sources of epilepsy patients obtained from the MEG inverse problem, the regions found by ANFISGC were more similar to the clinical findings than those found by the other methods.

1. Introduction

The human brain is a complex biological system. Exploring the brain function and analyzing the interactions among its different regions are challenging tasks. They can be accomplished using MEG and EEG signals that provide high temporal resolution (in the order of millisecond) but poor spatial resolution (He et al., 2011). These modalities are suitable for studying the brain effective connectivity as they provide enough data to explore the brain dynamics. The effective connectivity deals with the causal influence of the regions on each other (Greenblatt et al., 2012; Sakkalis, 2011).

The brain networks may be identified through an investigation of the brain connectivity. They have clinical applications for neurological disorders (e.g., autism, epilepsy, and Alzheimer's) (Menassa et al., 2018; Han et al., 2017; Boutros et al., 2015; Amini et al., 2011; Wilke et al., 2010; Chen et al., 2014). The brain networks of patients and healthy individuals can be compared to identify abnormalities of the patients' brains.

A variety of linear effective connectivity criteria have been developed based on the autoregressive (AR) and multivariate autoregressive (MVAR) models in the time and frequency domains. They include Linear Granger Causality (LGC) (Ding et al., 2007), combination of Pairwise Granger Causality and Conditional Granger Causality (PwGC + CGC) (Stramaglia et al., 2014), Partial Directed Coherence (PDC), Generalized Partial Directed Coherence (GPDC), Directed Transfer Function (DTF), full frequency Directed Transfer Function (ffDTF), and spectrum weighted Directed Transfer Function (swDTF) (Gurisko, 2014; Wu et al., 2011; Siggiridou et al., 2017).

The most widely used measure of linear effective connectivity is LGC. Based on LGC, time series Y is considered as the cause of time series X if incorporating the past information of Y improves the prediction of X from its past (Ding et al., 2007; Khadem and Hossein-Zadeh, 2014). In details,

* Corresponding author. School of Electrical and Computer Engineering, College of Engineering, University of Tehran, North Kargar Avenue, Tehran, Iran.
E-mail addresses: hsoltan1@hfhs.org, hszadeh@ut.ac.ir (H. Soltanian-Zadeh).

https://doi.org/10.1016/j.neuroimage.2018.07.024
Received 17 February 2018; Received in revised form 9 July 2018; Accepted 11 July 2018

to investigate a causal relation from Y to X, a model must be created to predict time series X using the lagged samples of X and Y time series. Similarly, a second model must be constructed to predict time series X using only lagged samples of X. If the error of predicting X is significantly higher in the second model relative to the first model, the causal link from Y to X (Y → X) will be accepted. A limitation of this method is the assumption of linear interactions among the brain regions (or M/EEG channels) because the prediction is done based on the AR and MVAR models.

Pairwise Granger Causality (PwGC) as a bivariate measure does not distinguish direct and indirect interactions. The Conditional Granger Causality (CGC), specially its full version, is one of the solutions but it is not applicable when the number of data samples is small. In (Stramaglia et al., 2014), a combination of PwGC and CGC (PwGC + CGC) is recommended. Variables with high PwGC to a target node are used in CGC as the conditional variables.

Since the interactions between the brain regions can be nonlinear (Marinazzo, 2011; Ioannides and Mitsis, 2010), linear measures suffer from low sensitivity. To address this shortcoming, researchers have developed measures of nonlinear connectivity based on Granger and entropy concepts.

Transfer Entropy (TE) (Schreiber, 2000), Partial Transfer Entropy (Gomez-Herrero, 2010), and Phase Transfer Entropy (PTE) (Lobier et al., 2014; Wang et al., 2017) are the nonlinear nonparametric methods developed based on the information theory. While PTE outperforms TE in the presence of noise, it is a bivariate measure and thus may not distinguish the direct and indirect connections.

Kernel Granger Causality (KGC) (Marinazzo et al., 2008a) is a nonlinear parametric measure where LGC is done in the feature space of a kernel function (Gaussian or inhomogeneous polynomial). The performance of KGC certainly depends on the kernel used. Also, a limitation of KGC is that it is a bivariate measure and thus may not distinguish the direct and indirect connections. To overcome this limitation, a multivariate version of KGC is proposed (Marinazzo et al., 2008b; Stramaglia et al., 2014). In (Chen et al., 2004), a new approach is presented where a nonlinear relation is approximated by locally linear AR models. However, this approximation may not be sufficient in highly nonlinear relationships.

Recently, an integrated method using β minimal Redundancy Maximal Relevance (βmRMR) regressor selection, Multi-Layer Perceptron (MLP), and Granger Causality (GC), named βmRMR–MLP-GC was developed as a nonlinear connectivity measure (Khadem and Hossein-Zadeh, 2014). In this measure, the MLP neural network was employed as a predictor instead of the MVAR model. The original implementation of βmRMR–MLP-GC does not include the GC concept as it was based on creating only one model. Here, to investigate the existence of the causal connection Y → X, they first form a model to predict time series X using the lagged samples of X and Y time series and then, to quantify the effect of Y on X, they set the lagged samples of Y to zero in the first model.

This paper describes our new effective connectivity measure called ANFISGC. We utilize ANFIS as an appropriate prediction tool to discover both linear and nonlinear connections. According to (Samanta, 2011), ANFIS is superior to other tools in predicting time series, especially for nonlinear and chaotic systems. In Section 2, we describe the structure and capabilities of ANFIS. We also describe an approach superior to βmRMR–MLP-GC in which, instead of constructing one model, two models based on the Granger concept are used to identify casual links.

In (Farokhzadi et al., 2016), we proposed a bivariate version of the measure. Here, we present a complete conditional version of the measure, which it is not bivariate and may distinguish between direct and indirect causal connections. It is not model based and thus does not have the limitation of the methods like KGC.

We apply ANFISGC on both simulated and experimental MEG dataset. The simulation design and real MEG data are described towards the end of Section 2. The results of applying ANFISGC and five alternative methods (LGC, KGC, PwGC + CGC, TE, and PTE) on the simulated and real data are reported in Section 3. Discussions, conclusions, and future works are presented in Sections 4–6, respectively.

2. Materials and methods

In this section, the theories behind ANFIS, embedding dimension based on symplectic geometry, and the Granger Causality are reviewed. Afterwards, these three tools are combined in a unified structure.

2.1. Adaptive Neuro Fuzzy Inference System

ANFIS is an adaptive neural network that works like a Fuzzy Inference System (FIS) and is applicable in different fields such as in modeling and time series prediction. Neural Networks (NN) have the ability of learning and generalization in solving nonlinear problems. FIS has the capability of approximate reasoning and employing fuzzy information theory. In ANFIS, the advantages of NN and FIS are combined. The structure of ANFIS is illustrated in Fig. 1.

ANFIS is a model with five layers and m inputs and includes adjustable and fixed nodes. The adjustable nodes are shown by squares and the fixed nodes are shown by circles in Fig. 1. The first layer generates membership functions using several parameters. The remaining layers implement multiplication (logical AND), normalization, linear regression, and summation as shown in Fig. 1.

ANFIS maps the inputs to the outputs using the membership functions, a set of rules, and some parameters. It learns an approximating function (f) to estimate the output (y) from the inputs ($x_1, x_2, \ldots, x_m$):

$$y = f(x_1, x_2, \ldots, x_m) \tag{1}$$

The gradient descend and the least square methods are used to adjust the parameters of the membership and regression functions in the first and fourth layers, respectively (Samanta, 2011).

2.2. Embedding dimension based on symplectic geometry

The model order of a system (time series) can be determined via a variety of approaches. Commonly, Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are used for the AR and MVAR models. For nonlinear systems, these criteria oversimplify the problem but the embedding dimension (Liu and Aviyente, 2012) may be applied using a variety of methods like Singular Value Decomposition (SVD), false nearest neighbors, and Cao's algorithm (Lie et al., 2002). However, these methods have the following weaknesses:

1. Data intensive, computationally complex, and subjective
2. Dependent on the number of data samples
3. Affected by noise
4. Inappropriate for nonlinear dynamics due to their linearity

In recent years, Symplectic Geometry (SG) is proposed as a more appropriate tool that does not have the above weaknesses. SG is similar to SVD but can reflect nonlinearity. The implementation of this method for estimating the embedding dimension of a single time series is explained in (Lie et al., 2002). For the multivariate time series, it can be applied by generalizing the trajectory matrix introduced in (Ataei et al., 2003).

2.3. Granger Causality index

Wiener Granger Causality provides a linear bivariate effective connectivity measure called the Granger Causality (GC). Based on the GC approach, a causal link will be assigned from Y to X, if incorporating the past samples of Y(t) reduces the prediction error of X(t), i.e., causes the improvement of X(t) prediction. The formulation of this method is as follows.