Computers in Biology and Medicine 41 (2011) 1132–1141

Contents lists available at ScienceDirect

Computers in Biology and Medicine

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

A comparison of multivariate causality based measures of effective connectivity

Meng-Hung Wu [a,*], Richard E. Frye [b,1], George Zouridakis [c,d,e,2]

[a] Departments of Computer Science, South Texas College, 3201 W. Pecan Blvd., McAllen, TX 78501, USA
[b] Arkansas Children's Hospital Research Institute, University of Arkansas Medical Sciences, Slot 512-41B, Room R4025, 13 Children's Way, Little Rock, AR 72202, USA
[c] Departments of Computer Science, University of Houston, 218 Philip G. Hoffman Hall, Houston, TX 77204, USA
[d] Departments of Electrical and Computer Engineering, University of Houston, 218 Philip G. Hoffman Hall, Houston, TX 77204, USA
[e] Departments of Engineering Technology, University of Houston, 218 Philip G. Hoffman Hall, Houston, TX 77204, USA

ARTICLE INFO

Keywords:
Connectivity measures
Autoregressive models
Multivariate causality
Directed transfer function
Partial directed coherence
Granger causality
Dynamic autoregressive neuromagnetic causal imaging (DANCI)
MEG

ABSTRACT

During the past several years a variety of methods have been developed to estimate the effective connectivity of neural networks from measurements of brain activity in an attempt to study causal interactions among distinct brain areas. Understanding the relative strengths and weaknesses of these methods, the assumptions they rely on, the accuracy they provide, and the computation time they require is of paramount importance in selecting the optimal method for a particular experimental task and for interpreting the results obtained. In this paper, the accuracy of the six most commonly used techniques for calculating effective connectivity are compared, namely directed transfer function, partial directed coherence, squared partial directed coherence, full frequency directed transfer function, direct directed transfer function, and Granger causality. These measures are derived from the coefficients and error terms of autoregressive models calculated using the dynamic autoregressive neuromagnetic causal imaging (DANCI) algorithm. These techniques were evaluated using magnetoencephalography recordings as well as several synthetic datasets that simulate neurophysiological signals, which varied on several parameters, including network size, signal-to-noise ratio, and network complexity, etc. The results show that Granger causality is the most accurate method across all experimental conditions explored and suggest that large multisensor data sets can be accurately analyzed using Granger causality with the DANCI algorithm.

© 2011 Published by Elsevier Ltd.

1. Introduction

A number of recent neuroimaging studies have focused on the importance of brain connectivity across small- and large-scale neural networks [1–6]. Several methods are used to study brain connectivity. Anatomic connectivity is most commonly studied using diffusion tensor imaging (DTI) while functional connectivity is most commonly studied by analyzing functional brain activity derived from functional neuroimaging techniques such as functional magnetic resonance imaging (fMRI) or magnetoencephalograph (MEG) [7–22]. These studies can provide insight into the relationships between brain structure and function. More recently, *effective connectivity* methods have been developed that can quantify *causal* relationships in neurophysiological neural networks [7–11].

In any complex (neural) network, nodes (sources of activity) can be connected through direct or indirect connections, and such networks can be analyzed by considering the interaction between nodes. Some techniques only consider the interactions between two nodes at a time while other techniques can consider the interactions between more than two nodes simultaneously and even between all of the nodes simultaneously. Fig. 1 shows a three nodes network. Node 1 has a direct causal influence on node 2 and node 2 has a direct causal influence on node 3 (direct connections shown by solid arrows). However, at the same time, node 1 has an indirect influence on node 3 via node 2 (indirect connections shown by a dashed arrow). Indirect connections become a problem in pair-wise analysis (i.e., when only two nodes are considered simultaneously) [11], especially with larger networks. The analysis of larger networks with many indirect connections requires multivariate methods that account for the influence of all sources simultaneously.

* Corresponding author. Tel.: +1 956 872 8364; fax: +1 956 872 2008.
E-mail addresses: mwu1@southtexascollege.edu (M. Wu), Richard.E.Frye@uth.tmc.edu (R.E. Frye), zouridakis@uh.edu (G. Zouridakis).
[1] Tel.: +1 713 500 3245; fax: +1 713 500 5711.
[2] Tel.: +1 713 743 8656; fax: +1 713 743 1250.

0010-4825/$ - see front matter © 2011 Published by Elsevier Ltd.
doi:10.1016/j.compbiomed.2011.06.007

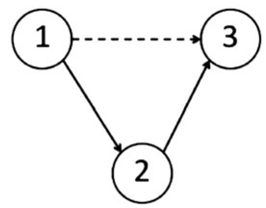

Fig. 1. Three-node network. Channel 1, 2, and 3 are interacting with each other in a network. Channel 1 has a causal direct influence on 2 and channel 2 has a causal direct influence on 3, shown by solid arrows. Channel 1 has an indirect influence on 3 via 2, as shown by a dashed arrow.

In general, there are three main classes of multivariate measures to estimate causal connectivity: the direct transfer function (DTF) and its derivatives, partial directed coherence (PDC) and its derivatives, and Granger causality. Autoregressive modeling is the basic framework for calculating multivariate causality. Linear autoregressive modeling has considerable flexibility and can model interactions within and between multiple sources simultaneously. The DTF, PDC, and their derivatives are based on the coefficients of linear autoregressive (AR) models of the recorded signals, while Granger causality is based on the error of the AR model. Which of these measures is the most advantageous and accurate is still an open question.

The DTF was defined by Kaminski and Blinowska [12,13] and was designed to measure directional information flow. DTF is calculated as the ratio between the inflow into a particular node and all nodes connected to the particular node. DTF is capable of finding the connections in a multivariate system of interacting signals by considering all signals simultaneously. As such, this method was the first multivariate approach used to assess causality in neural signals. However, this method is not able to distinguish between direct and indirect connections. DTF has several derivatives, including the full frequency DTF (ffDTF), which was defined by Korzeniewska [14] and is derived by integrating over a wide frequency spectrum to account for all connections over all frequency bands.

PDC was defined by Baccala and Sameshima [15,16] and represents the ratio between the outflow from one particular node and all nodes to which the particular node is connected. Since PDC does not use a transfer function, this method is able to disentangle direct and indirect connections [17,18]. Square partial directed coherence (sPDC) was introduced by Astolfi [19] and is derived from PDC by squaring the frequency dependent coefficients. Thus, sPDC emphasizes the strongest outflow connection and better differentiates between nodes with lower and higher outflow connectivity, thus increasing the sensitivity of the measure.

The direct DTF (dDTF) proposed by Korzeniewska [14] is derived from the PDC and DTF by multiplying ffDTF by PDC and combines the strengths of both measures. dDTF can determine whether a connection between two nodes is mediated by a third one (i.e., indirect connection) and the directionality of the connection.

Finally, Granger causality (GC) is based on the relative change in the AR model error when a new signal is added to improve the prediction of a modeled signal [20]. The GC method is designed to consider interactions among multiple sources simultaneously, thereby providing a true measure of causal connectivity between two sources, while allowing the indirect influence of other sources on the relationship to be taken into account.

Understanding the relative strengths and weaknesses of these methods, the assumptions they rely on, the accuracy they provide, and the computation time they require is of paramount importance in selecting the optimal method for a particular experimental task and for interpreting the results obtained. In this paper, the six most widely used techniques found in the literature are compared, namely DTF, ffDTF, PDC, sPDC, dDTF and GC. The coefficients and model error are derived using the Dynamic Autoregressive Neuromagnetic Causal Imaging (DANCI) algorithm [20,21,22]. The simulated data are used to compare the accuracy of the measures across a wide range of network architectures, AR model orders, network complexity, and signal-to-noise ratios. In these simulations, extra indirect connections between channels were added to increase the complexity of the networks. Those indirect connections between nodes cause partial influence between nodes and make it more difficult to estimate the true effective connectivity. We then compared the six methods described above using clinical MEG data obtained during a reading experimental task.

2. Methods

2.1. Data preprocessing

Stationary data are required for accurate AR modeling. If the data are not stationary, the multivariate AR models will be invalid and may contain 'spurious regression' results [23]. To prevent non-stationarity the data are normalized both within and across data epochs [20]. Stationarity is verified by examining the unit roots using the Dickey–Fuller unit root test (see Fig. 2A).

2.2. Linear autoregressive modeling

An AR process can be defined for any time series $X\{x(1), \ldots, x(t)\}$, so that $x(t)$ can be predicted by the values of the signal at some past times as follows:

$$x(t) = \sum_{j=1}^{p} a_{xj} x(t-j) + e_{(x|x)}(t) \quad (1)$$

where a_{xj}, $j=1, \ldots, p$ are the AR coefficients, p is the model order that determines the number of previous time points to be considered, and $e_{x|x}(t)$ is the error term, i.e., the portion of the signal $x(t)$ at time t, which is not accounted for by the previous values of $x(t-1), \ldots, x(t-p)$.

The time series $x(t)$ can also be predicted by combining its own past activity with past values of another signal, as follows:

$$x(t) = \sum_{j=1}^{p} a_{xj} x(t-j) + \sum_{j=1}^{p} a_{yj} y(t-j) + e_{(x|xy)}(t) \quad (2)$$

In Eq. (2) signal $x(t)$ at time t is predicted by the previous activity of both $x(t-1), \ldots, x(t-p)$ and $y(t-1), \ldots, y(t-p)$, and $e_{(x|xy)}(t)$ represents the error in predicting $x(t)$ given the previous values of $x(t-1), \ldots, x(t-p)$ and $y(t-1), \ldots, y(t-p)$.

2.3. Least-squares linear regression

The causal connectivity measures described below are derived from the coefficients and error terms of AR models calculated using the DANCI algorithm. The DANCI algorithm uses least-squares linear regression to calculate the AR coefficients for multiple sources, as described in detailed elsewhere [24] (Fig. 2B). In a companion paper within this issue, we demonstrate that DANCI provides a fast, accurate, unbiased, and robust estimation of Granger causality [24].

Directed Transfer Function (DTF) [12] relies on AR modeling, but the signal is transformed to the frequency domain to estimate the spectral properties of the process. The multivariate AR model of

Download English Version:

https://daneshyari.com/en/article/505558

Download Persian Version:

https://daneshyari.com/article/505558

Daneshyari.com