PHYSICAL REVIEW E **70**, 050902(R) (2004)

Granger causality and information flow in multivariate processes

Katarzyna J. Blinowska, Rafał Kuś, and Maciej Kamiński
Laboratory of Medical Physics, Institute of Experimental Physics, Warsaw University, Hoża 69, 00-681 Warszawa, Poland
(Received 8 June 2004; published 22 November 2004)

The multivariate versus bivariate measures of Granger causality were considered. Granger causality in the application to multivariate physiological time series has the meaning of the information flow between channels. It was shown by means of simulations and by the example of experimental electroencephalogram signals that bivariate estimates of directionality in case of mutually interdependent channels give erroneous results, therefore multivariate measures such as directed transfer function should be used for determination of the information flow.

DOI: 10.1103/PhysRevE.70.050902 PACS number(s): 87.10.+e, 05.45.Tp, 02.50.Sk, 02.50.Ey

In 2003, the Nobel prize in economics was awarded to C. W. J. Granger. One of his main contributions concerned the dependencies between time series: in 1969 he formulated the causality principle called, since, the "Granger causality measure" or "Granger causality" [1].

The causality dependence between time series is a topic of interest not only in econometrics, but also in biology and other natural sciences. This is especially the case in brain studies where information about the mutual influence between signals such as electroencephalogram (EEG), local-field potentials (LFP), and spike trains is important for theoretical studies as well as for clinical practice. Attempts to find measures of influence between two EEG signals, in the sense of propagation, started in the early 1980's [2,3] and concerned the propagation between two channels. In more recent papers concerning the estimation of directionality, usually bivariate measures are used: e.g., [4,5]. However, brain activity measured on different sites is highly correlated and there exists a multitude of relations between different recorded channels. In such a situation it is difficult to judge if two given channels interact with each other, or if they are driven by a third channel or channels. In his later work [6] Granger addressed the problem of missing information and he stated that a test for causality is impossible unless the set of interacting channels is complete.

An estimator of directionality of information flow for an **arbitrary number** of channels—the directed transfer function (DTF)—was introduced by Kamiński and Blinowska in 1993 [7]. It was based on a multivariate autoregressive (AR) model formulated by Franaszczuk *et al.* [8]. DTF in its non-normalized version is an extension of Granger causality to an arbitrary number of channels, as was shown in [9]. The DTF method has been applied to localize epileptic foci [10], to determine LFP propagation between brain structures of animals in different behavioral states [11], to investigate EEG activity propagation in different sleep stages [12], and to study the epileptogenesis [13].

The aim of this work will be a comparison between bivariate and multivariate estimators of directionality, by means of simulations and evaluation of experimental EEG signals. We shall point out some pitfalls in applying bivariate measures in case of a mutually dependent set of channels. The problem is important, since the commonly used bivariate measures often lead to erroneous results.

The concept of Granger causality is based on the predictability of time series. Namely, if a series $X_2(t)$ contains information in past terms that helps in the prediction of $X_1(t)$, and this information is contained in no other series used in the predictor, then $X_2(t)$ is said to cause $X_1(t)$. Formally, it can be written as

$$X_1(t) = \sum_{j=1}^{p} A_{11}(j)X_1(t-j) + \sum_{j=1}^{p} A_{12}(j)X_2(t-j) + E_1(t),$$
$$X_2(t) = \sum_{j=1}^{p} A_{21}(j)X_1(t-j) + \sum_{j=1}^{p} A_{22}(j)X_2(t-j) + E_2(t). \tag{1}$$

The above equations are identical with the definition of bivariate AR model. The generalized AR model defined for an arbitrary number of channels (MVAR) is represented as a vector $\mathbf{X}$ of k signals recorded in time: $\mathbf{X}(t) = [X_1(t), X_2(t), \ldots, X_k(t)]^{\mathrm{T}}$. Then the MVAR model can be expressed as

$$\mathbf{X}(t) = \sum_{i=1}^{p} \hat{\mathbf{A}}(i)\mathbf{X}(t-i) + \mathbf{E}(t), \tag{2}$$

where $\mathbf{X}(t)$ is the data vector at time t, $\mathbf{E}(t)$ is the vector of white-noise values, $\hat{\mathbf{A}}(i)$ are the model coefficients, and p is the model order. The model order can be found by means of criteria derived from information theory; in this work the Akaike information criterion (AIC) [14] was used.

Without loss of generality we can rewrite (2) as

$$\mathbf{E}(t) = \sum_{i=0}^{p} \mathbf{A}(i)\mathbf{X}(t-i),$$
$$\mathbf{A}(0) = \mathbf{I}, \quad \mathbf{A}(i) = -\hat{\mathbf{A}}(i) \quad \text{for } i = 1, \ldots, p. \tag{3}$$

Then, transforming the model equation to the frequency domain we get

$$\mathbf{E}(f) = \mathbf{A}(f)\mathbf{X}(f),$$
$$\mathbf{X}(f) = \mathbf{A}^{-1}(f)\mathbf{E}(f) = \mathbf{H}(f)\mathbf{E}(f). \tag{4}$$

The $\mathbf{H}(f)$ matrix is called the transfer matrix of the system, f denotes frequency, $\mathbf{E}(f)$ is the transformation of $\mathbf{E}(t)$. From the transfer matrix, we can calculate power spectra $\mathbf{S}(f)$ and coherences. If we denote by $\mathbf{V}$ the variance matrix

 ©2004 The American Physical Society

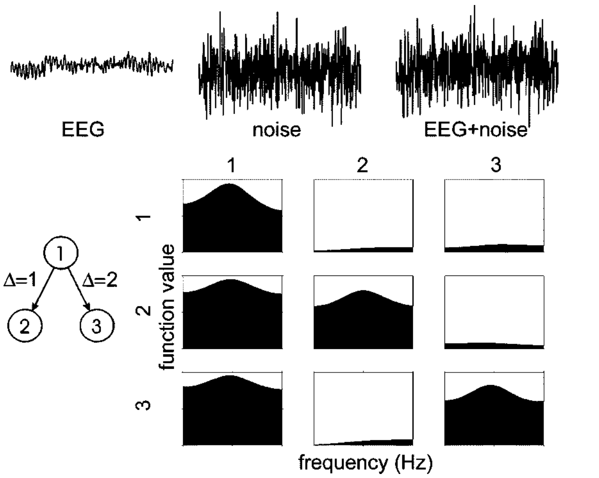

FIG. 1. Top: Simulated signals. Bottom left: Simulation scheme. Bottom right: In each box DTF as functions of frequency (0–25 Hz); the number above the columns indicates output channels, the numbers on the left of the rows indicate destination channels.

of the noise $\mathbf{E}(f)$, the power spectrum is defined by (asterisk means transposition and complex conjugate)

$$\mathbf{S}(f) = \mathbf{H}(f)\mathbf{V}\mathbf{H}^{*}(f). \tag{5}$$

The DTF was defined [7] in terms of the elements of the transfer matrix H_{ij},

$$\gamma_{ij}^{2}(f) = \frac{|H_{ij}(f)|^{2}}{\sum_{m=1}^{k} |H_{im}(f)|^{2}}. \tag{6}$$

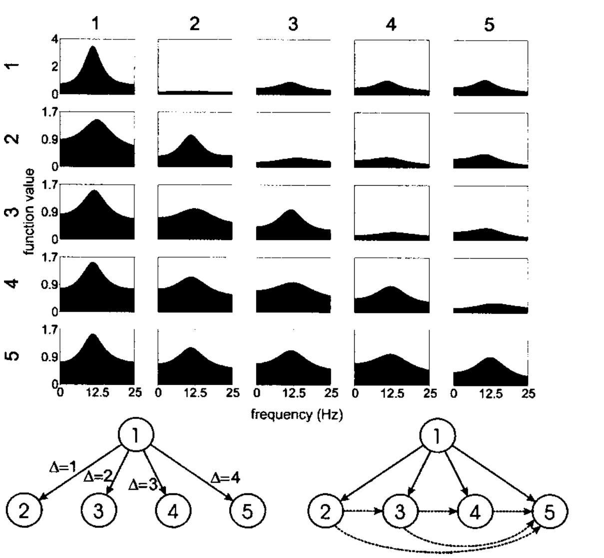

FIG. 2. Top: Granger causality calculated pairwise; each graph represents the function describing transmission from the channel marked above the row to the channel marked on the left of the row. Granger causality in arbitrary units on vertical axes; graphs on the diagonal contain power spectra; frequency on horizontal axes (0–25 Hz range). Bottom left: Simulation scheme. Bottom right: Resulting flow scheme. The black arrows represent true (simulated) flows; the dotted arrows represent false flows found by the applied method.

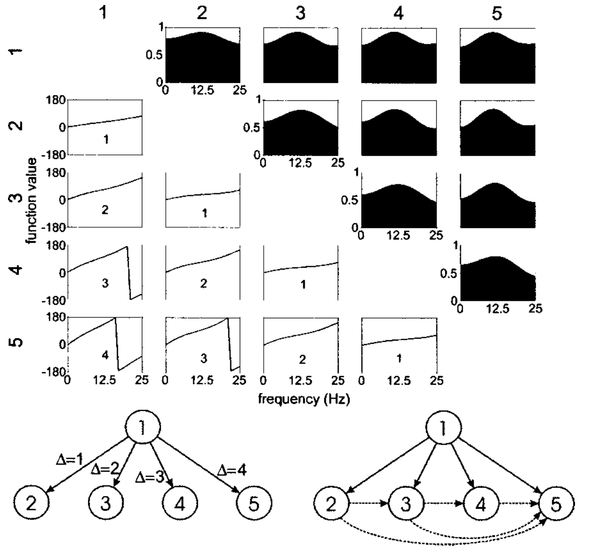

FIG. 3. Pairwise coherences and resulting flows. Top: Coherence amplitude (black graphs above diagonal) and coherence phase (graphs below diagonal); each graph represents the function for the pair of channels marked on the left of the row and above the column; on the horizontal axes frequency (0–25 Hz); on vertical axes coherence amplitudes (0–1 range) or phases ($-\pi$–π range); delay values (in samples) estimated from phases, marked by the numbers shown over the phase graphs. Bottom left: Simulation scheme. Bottom right: Resulting flow scheme. The same convention in drawing arrows as in Fig. 2.

The normalization of DTF is performed in such a way that γ_{ij} describes the ratio between the inflow from channel j to channel i and all the inflows of activity to the destination channel i. Such a ratio takes values in the range [0, 1]. A value close to 1 indicates that most of the signal in channel i consists of the signal from channel j; values of DTF close to 0 indicate that there was almost no flow from channel j to channel i at this frequency.

The Granger causality measure is equivalent to the non-normalized version of DTF. In this case only the element $H_{ij}(f)$ of the transfer matrix is used to describe the transmission. In the application to multivariate biological time series, Granger causality has the interpretation of information flow between channels.

In our calculations we have used the Yule-Walker algorithm for MVAR model fitting and the Akaike [14] criterion for estimation of the model order. In the simulations we have constructed the signal in channel 1 from an experimental EEG (2560 samples long, sampled at 128 Hz, highpass filtered with cutoff frequency 3 Hz) plus a random noise. The signals in destination channels were constructed by introducing delays and adding to each delayed channel an extra noise (drawn from different random noise generators). The performance of DTF is shown by a simple example illustrated in Fig. 1, where the variance of noise in channels 2 and 3 is nine times as big as in the input channel. Comparing values of DTF functions in Fig. 1, we can easily observe that the flows come only from channel 1; even in the presence of a noise several orders of magnitude bigger than the signal, DTF is able to detect the right direction of propagation.

The next simulations concern a situation common in EEG

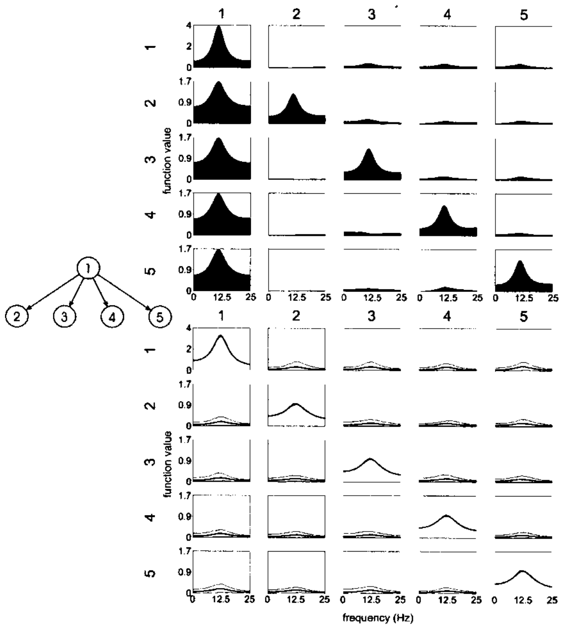

FIG. 4. Top: Non-normalized multichannel DTFs for the simulation (Fig. 2). The picture organization is similar to Fig. 3 (on the diagonal power spectra). Bottom: DTFs obtained from surrogate data. Thick line: The average obtained from 100 surrogates. Ninety five percent of surrogate realizations are contained between the thin lines. The plots in both panels are in the same scale in arbitrary units. Frequency on horizontal axes, 0–25 Hz range. At left is the resulting flow pattern.

studies. The activity is propagating from a certain location in the brain with different delays to the sites where it is recorded. The simulation scheme is shown in Fig. 2 (construction of signals as in the previous simulation, only the variance of noise in destination channels was two times as big as the variance of signal 1). In order to test bivariate measures of propagation, the AR model was fitted to two channels at a time and bivariate Granger causality was calculated. The results (Fig. 2) show that a number of false flows was found. Namely, a flow was found always when a phase difference existed between a pair of channels.

Another way of determining the direction of flow is through estimation of phases of coherence functions. Ordinary coherence between channels i and j is defined using elements of spectral matrix $\mathbf{S}(f)$ as

$$K_{ij}(f) = \frac{S_{ij}(f)}{\sqrt{S_{ii}(f)S_{jj}(f)}}. \tag{7}$$

The application of this method to the simulation scheme from Fig. 2 gave the same erroneous result (Fig. 3) as the one obtained by bivariate Granger causality.

The DTF functions, obtained by fitting the MVAR model simultaneously to all channels of the simulated process, and the resulting flow scheme are shown in Fig. 4. In this case the pattern of propagations is reproduced correctly. DTF utilizes the full multivariate power of MVAR, which allows for identification of multiple causality relations between signals. The accuracy of the method can be tested by means of surrogate data. They were constructed by randomizing phases in our simulation scheme, according to the procedure proposed in [15]. The procedure of calculating surrogates was repeated 100 times. The maximal level of flows obtained from surrogate data was similar to the "leak" flows obtained by DTF.

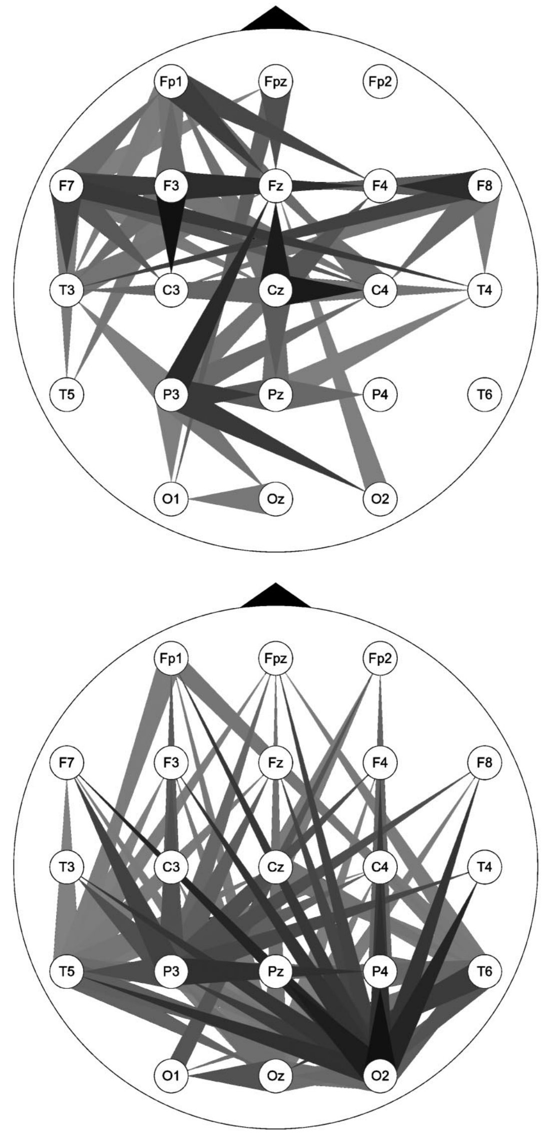

FIG. 5. The directions of flows found for experimental EEG signals recorded from 21 channels (10/20 system), for awake state, obtained by the bivariate Granger causality measure (top) and non-normalized DTF function calculated by means of fitting MVAR to all 21 channels simultaneously (bottom). The shade of gray of the arrow represents the strength of the connection (black=the strongest). In order to make the pattern of flows clear, the cutoff of relative intensity was applied on the level 0.53 with respect to the strongest flow, which was assigned the value of 1.

In this paper, we have concentrated our attention on the simulations illustrating at best the difference between multivariate and bivariate measures. Simulations illustrating DTF performance for a more complicated configuration of sources are described, e.g., in [16], where also different normalizations of the estimator are discussed.

In order to illustrate the performance of bivariate versus multivariate estimates we have tested them on physiological time series, namely EEG signals recorded in awake state, eyes closed (referenced to linked ears, sampling frequency 128 Hz, 2560 time points).[1] Multivariate Granger causality (non-normalized DTF functions) were calculated from the MVAR model fitted to all 21 channels simultaneously and bivariate measures of Granger causality were found from two-channel AR models. The flows in the alpha activity range were estimated by integrating the functions in the 8–14 Hz range. In case of DTF a consistent pattern of flows is observed. The sources of propagation can be seen in the posterior regions of the head (Fig. 5). It is known that in awake state with eyes closed the most prominent alpha activity comes from the visual cortex located in the posterior regions of head, which agrees well with our results. In contrast to these findings the bivariate estimates gave a rather inconsistent pattern of flows with multiple sources in different regions of the head and "sinks" to which the activity flows from all possible directions (e.g., Fz, T3 electrodes). This last effect can be easily explained with our simulations.

Recently in the literature one can find many methods aimed at finding the direction of propagation in time series. Practically all of them are based on bivariate estimates. For example, for detecting direction of coupling between multichannel EEG or magnetoencefalogram (MEG) data, a bivariate measure based on the assumption of two coupled oscillators and an estimation of their phases was proposed [16]. The method can be helpful in case of two signals independent from other influences, e.g., a physical experiment under control; however, for EEG and MEG studies it is hardly applicable, because: (1) as we have demonstrated above, bivariate measures in case of a mutually dependent set of channels are likely to give erroneous results, and (2) the method is based on the assumption of two interacting oscillators. This assumption limits the application for brain activity study, since in this case we deal with oscillators acting with different frequencies, corresponding to EEG rhythms. DTF is a spectral estimate and it can detect the case when activities of different frequencies propagate in different directions, as was demonstrated in experimental studies [12].

A great deal of attention of the scientific society is devoted to measures of causality between nonlinear time series, e.g., [17–19]. In [19] the concept of Granger causality was extended for the nonlinear case. The authors admit that pairwise analysis is inappropriate for more than two channels and introduce conditional extended Granger causality. However, in case of nonlinear measures a number of pitfalls are encountered and the characterization of nonlinear dependence can be quite problematic [18]. Fortunately, it was demonstrated by numerous studies based on linear versus nonlinear forecasting [20] or surrogate data techniques [21,22], that EEG and LFP can be considered a colored noise time series and that one can trace chaotic deterministic behavior only in certain phases of epileptic seizure [23]. However, even in this case linear techniques perform well: e.g., in [10] the power of the multivariate DTF function in epileptic focus localization was demonstrated.

The problem of the directedness of information flow is not limited to EEG or LFP signals; the same problem is encountered, e.g., in geophysics or other natural sciences when numerous mutually interacting processes take place.

In this study we would like to emphasize the importance of a multivariate approach, which merits more attention, since the pitfalls in evaluating the direction of causal relations in EEG or LFP connected with the use of bivariate instead of multivariate techniques are much more serious than limitations connected with the assumption of the linearity of time series.

This work was partly supported by the Polish Committee for Scientific Research grant to the Institute of Experimental Physics.

[1]The signals are obtained by courtesy of Professor W. Szelenberger from Warsaw Medical University.

[1] C. W. J. Granger, Econometrica **37**, 424 (1969).
[2] Y. Saito and H. Harashima, in *Recent Advances in EEG and EMG Data Processing*, edited by N. Yamaguchi and K. Fujisawa (Elsevier, New York, 1981).
[3] T. Inouye, A. Yagasaki, H. Takahashi, and K. Shinosaki, Electroencephalogr. Clin. Neurophysiol., **51**, 265 (1981).
[4] C. Bernasconi and P. König, Biol. Cybern. **81**, 199 (1999).
[5] W. A. Freiwald, P. Valdes, J. Bosch, R. Biscay, J. C. Jimenez, L. M. Rodriguez, V. Rodriguez, A. K. Kreiter, and W. Singer, J. Neurosci. Methods **94**, 105 (1999).
[6] C. W. J. Granger, J. Econ. Dyn. Control **2**, 329 (1980).
[7] M. Kamiński and K. J. Blinowska, Biol. Cybern. **65**, 203 (1991).
[8] P. J. Franaszczuk, K. J. Blinowska, and M. Kowalczyk, Biol. Cybern. **51**, 239 (1985).
[9] M. Kamiński, M. Ding, W. Truccolo, and S. Bressler, Biol. Cybern. **85**, 145 (2001).
[10] P. J. Franaszczuk and G. K. Bergey, Brain Topogr **11**, 13 (1998).
[11] A. Korzeniewska, S. Kasicki, M. Kamiński, and K. J. Blinowska, J. Neurosci. Methods **73**, 49 (1997).
[12] M. Kamiński, K. J. Blinowska, and W. Szelenberger, Electroencephalogr. Clin. Neurophysiol. **102**, 216 (1997).
[13] A. Medvedev and J. O. Willoughby, Int. J. Neurosci. **97**, 149 (1999).
[14] H. Akaike, IEEE Trans. Autom. Control **19**, 716 (1974).
[15] J. J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and D. Farmer, Physica D **58**, 77 (1992).
[16] A. Korzeniewska *et al.*, J. Neurosci. Methods **125**, 195 (2003).
[17] M. G. Rosenblum and A. S. Pikovsky, Phys. Rev. E **64**, 045202(R) (2001).
[18] R. G. Andrzejak *et al.*, Phys. Rev. E **68**, 066202 (2003).
[19] Y. Chen *et al.*, Phys. Lett. A **324**, 26 (2004).
[20] K. J. Blinowska and M. Malinowski, Biol. Cybern. **66**, 159 (1991).
[21] P. Ackerman *et al.*, Electroencephalogr. Clin. Neurophysiol. **90**, 384 (1994).
[22] C. J. Stam *et al.*, Clin. Neurophysiol. **110**, 1801 (1999).
[23] J. P. M. Pijn, Brain Topogr **9**, 249 (1997).