34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Comparing causality measures of fMRI data using PCA, CCA and Vector Autoregressive Modelling

Adnan Shah and Muhammad Usman Khalid and Abd-Krim Seghouane

Abstract— **Extracting the directional interaction between activated brain areas from functional magnetic resonance imaging (fMRI) time series measurements of their activity is a significant step in understanding the process of brain functions. In this paper, the directional interaction between fMRI time series characterizing the activity of two neuronal sites is quantified using two measures; one derived based on univariate autoregressive and autoregressive exogenous (AR/ARX) and other derived based on multivariate vector autoregressive and vector autoregressive exogenous (VAR/VARX) models. The significance and effectiveness of these measures is illustrated on both simulated and real fMRI data sets. It has been revealed that VAR modelling of the regions of interest is robust in inferring true causality compared to principal component analysis (PCA) and canonical correlation analysis (CCA) based causality methods.**

Keywords: **Functional MRI, PCA, CCA, VAR/VARX, causality, effective connectivity**

I. INTRODUCTION

In order to thoroughly characterize brain functions, fMRI should not be used only for functional segregation to accurately delineate activated brain regions but also for functional integration to explore brain networks. Important information on the brain structure can be obtained by measuring to which extent the individual neuronal sites contribute to information production and how they communicate among each other. The individual site contributions can be measured with quantities such as entropy, however the direction of interaction requires special measures. Inferring the interaction structure between two brain regions from the time series measurements of their activity requires three steps from the analysis technique: i) the detection of existence of interaction or coupling, ii) the distinction between direct and indirect interaction and iii) the definition of the direction of interaction or the direction of information flow. Symmetric measures such as temporal correlation and mutual information have been used for the estimation of functional connectivity network [10], [9]. Unfortunately neither of these measures can provide directional information. Model-based directionality measures of coupling such as SEM [6] or DCM [7] have been proposed to address the shortcomings of symmetric measures such as linear correlation. The concept of Granger causality (GC) [8], a data-driven approach, has been adapted to fMRI in order to identify causal relations in the brain [5], [1] and later on extended to multiple time-series based on [11]. Directed information (transinformation [12]) has been introduced to infer functional neuroimage time-series causality. Furthermore, a data driven measure derived using Kullback-Leibler divergence has been used to quantify the information flow in fMRI time-series [13]. However, all these data-driven causality statistics reported in the literature measure the same underlying quantity and are related by some scaling factor [14].

In this work, a comparison in terms of data-driven causality between fitting a univariate autoregressive model and a multivariate vector autoregressive model (VAR) to fMRI time-series data of two brain regions of interest (ROIs) is reported. In a typical fMRI experiment, several ROIs are a priori identified in the brain. Each ROI is represented in the fMRI data set by multiple voxels, where each voxel is a variable comprising a single time series reflecting changes in the underlying metabolic signal. The directional interaction between two ROIs can be assessed by formulating the ROIs interaction as either an interaction among univariate time-series or among multivariate time-series both leading to a drive-response relation between source and target ROI. For univariate approach, dimensionality reduction techniques such as PCA and CCA [17] are used to derive a representative time series for each interacting ROI, which are then used to derive the causality measure using the standard GC framework in the context of AR/ARX models. In multivariate approach, interactions among groups of voxels from interacting ROIs are taken into account by fitting individual vector autoregressive (VAR) model to each ROI. A multivariate causality measure [11] is derived for these interacting ROIs based on the generalized variance of the residual errors in the context of VAR/VARX models.

The next section introduces the VAR framework for computing the causal interaction. In Section 3, PCA and CCA is discussed and univariate causality measure is derived. Section 4 presents performance results on both simulated and real fMRI data. Concluding remarks are given in Section 5.

II. VECTOR AUTOREGRESSIVE APPROACH

To assess the interaction in multivariate framework, ROI time-series are taken together and the ROI as a vector autoregressive model (VAR) [4] take part in determining the causality. With no standard definition for GC when the interacting ROIs are multivariate, GC based on Gewekes

NICTA is funded by the Australian Government as represented by the Department of Broadband, Communications and the Digital Economy and the Australian Research Council through the ICT Centre of Excellence program. This work was supported by grants from the Human Frontier Science Program, grant number: RGY 80/2008.

[1] National ICT Australia, Canberra Research Laboratory, The Australian National University, College of Engineering and Computer Science, Canberra, Australia adnan.shah, muhammad.khalid, abd-krim.seghouane at nicta.com.au

approach [11] with generalized variance (alternatively trace) of the residual errors is used here to find multivariate causality among interacting ROIs.

VAR models simultaneously estimate the interrelationship between more than one endogenous variables by estimating the coefficient matrix of autoregressive coefficient for each time-series involved in model estimation. Assume two multivariate stochastic processes of length N, $X = (x_{v,k} : v = 1, ..., V; k = 1, ..., N)$ and $Y = (y_{v,k} : v = 1, ..., V; k = 1, ..., N)$ corresponding to two ROIs composed of V voxels between which some interaction exists. Then the interacting VAR models based on X and Y in the standard form of a multivariate linear regression are as follows

$$ROI_1 : X = Z\Theta + \eta, \tag{1}$$

$$ROI_1 \leftarrow ROI_2 : X = Z\Theta + K\Phi + \zeta, \tag{2}$$

where $Z = [X_{k-1}, X_{k-2}, ..., X_{k-p}]$ and $K = [Y_{k-1}, Y_{k-2}, ..., Y_{k-q}]$ are the p and q previous multivariate time series samples, X is $((N-p) \times V)$ and Θ and Φ are $(V \times V)$ coefficient matrices of the VAR model to be estimated. η and ζ are i.i.d Gaussian noises with zero mean and $d \times d$ covariance matrices Σ_p and Σ_{pq}, respectively.

In such a VAR model (1) of order p with V channels, there are pV^2 number of unknown coefficients to be estimated, which can easily be satisfied for a least square solution [16] in fMRI data by choosing ROI size in relation to the sample length. The least-squares (LS) solution for the VAR coefficients in (1) is

$$\hat{\Theta} = (Z^\top Z)^{-1} Z^\top X \tag{3}$$

and the corresponding noise covariance estimate is,

$$\hat{\Sigma}_p = \frac{1}{N-l}(X - Z\hat{\Theta})^\top (X - Z\hat{\Theta}) \tag{4}$$

Similarly, for (2) we consider the generic least-squares problem of minimizing

$$T(\Theta, \Phi) = (X - Z\Theta - K\Phi)^\top (X - Z\Theta - K\Phi) \tag{5}$$

First, minimizing (5) with respect to Θ to get $\hat{\Theta} = (Z^\top Z)^{-1} Z^\top (X - K\Phi)$ and then minimizing the updated expression with respect to Φ to get $\hat{\Phi} = (K^\top PK)^{-1} K^\top PX$ where, $P = I - Z(Z^T Z)^{-1} Z^\top$ is the orthogonal projection. Finally, the minimized LS estimate of Θ incorporating $\hat{\Phi}$ into account is

$$\hat{\Theta} = (Z^\top Z)^{-1} Z^\top (X - K\hat{\Phi}) \tag{6}$$

and the corresponding noise covariance estimate is,

$$\hat{\Sigma}_{pq} = \frac{1}{N-l}(X - Z\hat{\Theta} - K\hat{\Phi})^\top (X - Z\hat{\Theta} - K\hat{\Phi}) \tag{7}$$

Based on these estimates, multivariate ROI causal interaction C_1 is defined based on generalized variance as the log of the ratio of the determinants of the residual covariances for the regressions

$$C_1(X \leftarrow Y) = ln(|\hat{\Sigma}_p| / |\hat{\Sigma}_{pq}|) \tag{8}$$

with $C_1(X \rightarrow Y) > C_1(Y \rightarrow X)$ implies that the ROI X is predominantly acting as a source and Y as a target.

III. Univariate Autoregressive Approach

In univariate framework, dimensionality reduction is performed on ROI time-series to define a representative signal. This can be achieved by three methods: i) averaging across all voxels, ii) PCA on all voxels and projection of the first principal component, iii) CCA on all voxels with its delayed version and projection of the most autocorrelated component [17]. We use PCA and CCA here to extract representative signals for interacting ROIs as follows:

PCA: For a V dimensional mean-free ROI time-series X and Y of N observations, for example X having linear transformation

$$P_c = u_1 X_1 + u_2 X_2 + ... + u_V X_V = uX \tag{9}$$

we obtain the first principal component x_k and y_k for each ROI by solving the following maximization problem [18], with weight vector u maximizing the variance of given ROI data

$$\mathrm{L}(\lambda, u_v) = u_v' X_v' X_v u_v - \lambda(u_v' u_v - 1) - \phi(u_v' u_w) \tag{10}$$

Based on the representative signals, the inter-ROI interaction can then be modeled as the following autoregressive models based on $AR(p)$ and $ARX(p, q)$ approach as in [13]

$$\begin{aligned} x_k &= \sum_{j=1}^{p} a_j x_{k-j} + \epsilon_k \\ x_k &= \sum_{j=1}^{p} a_j x_{k-j} + \sum_{l=1}^{q} b_l y_{k-l} + \epsilon_k' \end{aligned} \tag{11}$$

where ϵ_k and ϵ_k' are i.i.d $N(0, \hat{\sigma}_1^2)$ and $N(0, \hat{\sigma}_2^2)$, respectively.

CCA: Consider V dimensional mean-free ROI time-series X of N observations, $k = 1, ..., N$, and its one-sample delayed version X_{k-1}, their canonical variates are as

$$\begin{aligned} F &= r_1 X_1 + r_2 X_2 + ... + r_V X_V = rX \\ G &= s_1 X_{k-1,1} + ... + s_V X_{k-1,V} = sX_{k-1} \end{aligned} \tag{12}$$

(12) gives a representative time-series x_k' for ROI X, where canonical variates r and s which maximizes the correlation for X and X_{k-1} can be obtained by solving the following maximization problem [20]

$$\begin{aligned} \mathrm{L}(\lambda, r, s) = r' X' X_{k-1} s - \frac{\lambda}{2}(r' X' X r - 1) \\ - \frac{\rho}{2}(s' X_{k-1}' X_{k-1} s - 1) \end{aligned} \tag{13}$$

Similarly a representative time-series y'_k can be obtained for ROI Y. Based on the representative signals, the inter-ROI interaction can then be modeled similarly as in (11)

$$x'_k = \sum_{j=1}^{p} a'_j x'_{k-j} + \kappa_k$$
$$x'_k = \sum_{j=1}^{p} a'_j x'_{k-j} + \sum_{l=1}^{q} b'_l y'_{k-l} + \kappa'_k \qquad (14)$$

where κ_k and κ'_k are i.i.d $N(0, \hat{\sigma}_3^2)$ and $N(0, \hat{\sigma}_4^2)$, respectively.

Causality: After extracting the representative time-series for each ROI using PCA and CCA, the causal interaction from $X \rightarrow Y$ at time k is calculated from the discrepancy between the probability densities of $AR(p)$ and $ARX(p,q)$ as in [13]. Based on these estimates, the univariate ROI causal interaction C_2 for PCA and C_3 for CCA extracted time-series are defined as the log of the ratio of the residual variances for the regressions in (11) and (14).

$$C_2(X \leftarrow Y) = ln(\hat{\sigma}_1^2/\hat{\sigma}_2^2) \qquad (15)$$
$$C_3(X \leftarrow Y) = ln(\hat{\sigma}_3^2/\hat{\sigma}_4^2) \qquad (16)$$

with $C_i(X \rightarrow Y) > C_i(Y \rightarrow X)$ for $i \in \{2,3\}$ implies that the ROI X is predominantly acting as a source and Y as a target.

IV. Application Results

A. Simulated fMRI Data

In order to test and validate these methods for identifying the directional interaction in fMRI data, 10 realizations of two ROIs with 20 time series of 600 samples each were simulated. For each time-series the stimulus sequence is a realization of alternating blocks of activity [40 s] and rest [60 s] with short volume repetition time [TR = 1s] that is suitable to recover directed neuronal influences [1]. These ROI time-series were generated with two autoregressive processes where the second being driven by the first according to the following model

$$\begin{cases} x_k = z_k + 0.5x_{k-2} + \epsilon_k \\ y_k = z_k + 0.3y_{k-1} + 0.4f(x_{k-1}) + \epsilon'_k \end{cases} \qquad (17)$$

Where z_k is the voxel response at time instant k obtained by convolving stimulus sequence with the canonical HRF used by SPM software [19] . The function f is used to introduce the type of interaction. For linear interaction, f is set to identity whereas a nonlinear interaction is introduced by setting f to $exp(x)$. The components ϵ_k and ϵ'_k represent i.i.d. Gaussian noise $N(0, \sigma^2 = [0.164 : 0.164 : 3.28])$ for both linear and non-linear interactions. Both time-series were generated based on model (17) using zero initial conditions. (15) and (16) were used to calculate the directional interaction for univariate method with the known models. In practice the order of these models have to be estimated using the AIC_c criterion for example [3], [21] or KIC_c [22]. Similarly, (8) was used to find directional interaction for multivariate method assuming an order $p = q = 1$. Though VAR(1) is proven to give causality measure in fMRI data, it can be further tested by VAR model selection such as [4]. Figure 2 illustrates the causality measure robustness and gain brought by VAR(1) in comparison to univariate approach. With both the linear and nonlinear interactions, there is a clear asymmetry in directional interaction with multivariate VAR(1) approach, in favor of the direction $X \rightarrow Y$ indicating that X causes Y.

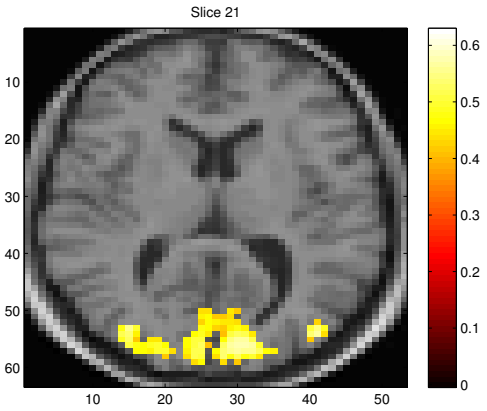

Fig. 1. Activation slice

B. Real fMRI Data

An experiment was performed to investigate regions of the brain that responded to visual motion (using the 'attention to visual motion fMRI' data set) [15]. A 2 Tesla MAGNETOM Vision system (Siemens, Erlangen) whole body MRI system with a head volume coil was used and contiguous multi-slice T_2^*-weighted fMRI images (T_E =40ms; 90 ms/image; 64 x 64 pixels) were obtained over a period of 5 min 22 s producing 100 image volumes in each run. The subjects were requested to look at a fixation point (size 0.3°) in the middle of a transparent screen and images were back projected onto the screen using an LCD video-projector. 250 white dots (size 0.1°) were moved radially from the fixation point in random directions towards the border of the screen, at a constant speed of 4.7°/s, where they vanished. The speed of the moving dots was changed five times during each trial. Subjects were then asked to indicate any change in speed. This condition lasted 32.2 s, giving 10 multi-slice volumes per conditions. Image processing and statistical analysis were carried out using SPM8 and Matlab. As expected, the activation detection task showed activation in V1, V5 and SPC [15][2]. A single activated slice is shown in figure 1. The characterization of the direction of attentional modulation V1 - V5 was of interest. After dimensionality reduction using PCA and CCA over selected 6 voxels from each ROI V1 and V5, two time-series were generated for univariate analysis. Choice of optimal order of the autoregressive fitting models on both time series were obtained using AICc [3]. For

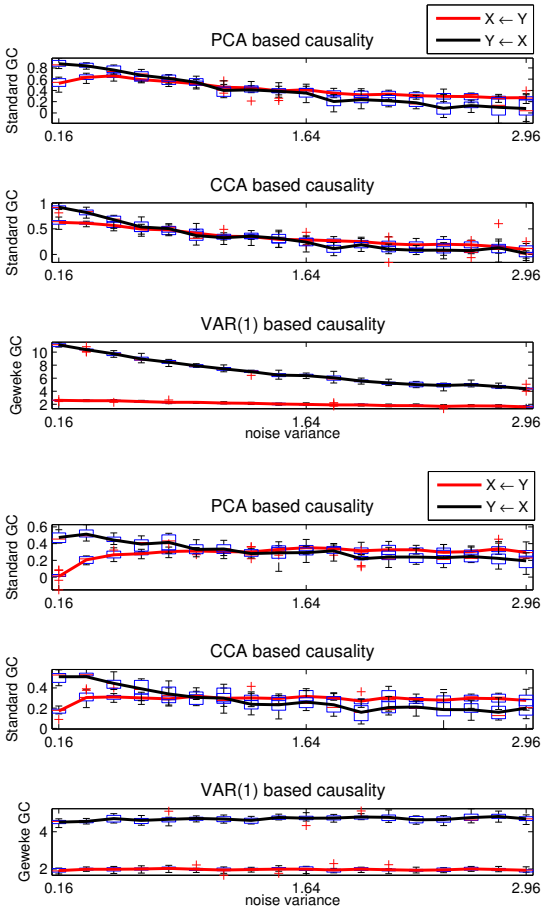

Fig. 2. Information flow between two ROIs - Nonlinear interaction (top 3 subplots) and linear interaction (bottom 3 subplots)

Method	$V1 \rightarrow V5$	$V5 \rightarrow V1$
PCA	0.14	0.05
CCA	0.14	0.07
VAR(1)	0.52	0.17

TABLE I

DIRECTIONAL INFLUENCE

taking the interrelationship between all voxel time-series into account as a multivariate regional interaction, selected group of voxels from ROI V1 and V5 were fitted with VAR/VARX models of order 1. Relation (15) and (16) were used to compute the directional interaction using the residual variances for the fitted univariate regression models. Similarly, relation (8) was used to compute the directional interaction using the residual covariances for the fitted multivariate regression models. The obtained values are reported in Table-1. Though both methods confirm the forward attentional modulation $V1 \rightarrow V5$, multivariate approach based on VAR(1) is not only better in quantifying the directional influence but also waiving off the need for VAR model selection criterion.

V. CONCLUSION

In this paper, the problem of extracting true causal interaction between activated brain regions from fMRI time series measurements has been approached in univariate and multivariate frameworks. Inter-ROI causality analysis has been carried out using dimensionality reduction techniques PCA and CCA; and as VAR modelling of the ROIs. Though both univariate and multivariate methods have the potential of extracting meaningful causality in the real fMRI data, multivariate VAR(1) method is robust in identifying true causal interaction in low power fMRI signals. The performance and effectiveness of these methods as tested on simulated and real fMRI data reveal that VAR based causality is an efficient measure for computing effective connectivity.

REFERENCES

[1] A. Roebroeck, E. Formisane, and R. Goebel, Mapping directe influence over the brain using Granger causality and fMRI, NeuroImage, vol. 25, pp. 230-242, 2005

[2] A. K. Seghouane and A. Shah, HRF Estimation in fMRI Data With an Unknown Drift Matrix by Iterative Minimization of the Kullback-Leibler Divergence, IEEE Trans. on Medical Imaging, vol. 2, pp. 192-206, 2012

[3] A. K. Seghouane and S. I. Amari, The AIC criterion and symmetrizing the Kullback-Leibler divergence, IEEE Trans. on Neural Networks, vol. 8, pp. 97-106, 2007

[4] A. K. Seghouane, Vector autoregressive model-order selection from finite samples using Kullback's symmetric divergence, IEEE Trans. on Circuits and Systems, vol. 53, pp. 2327-2335, 2006

[5] L. Harrison, W. D. Penny, and K. Friston, Multivariate autoregressive modeling of FMRI time series, NeuroImage, vol. 19, pp. 1477-1491, 2003

[6] A. R. McIntosh and F. Gonzalez-Lima, Structural equation modeling and its application to network analysis in functional brain imaging, Human Brain Mapping, vol. 2, pp. 2-22, 1994

[7] K. Friston, L. Harrison, and W. D. Penny, Dynamic Causal Modelling, NeuroImage, vol. 19, pp. 1273-1302, 2003

[8] C. W. J. Granger, Investigating Causal relations by econometric models and cross-spectral models, Econometrica, vol. 37, pp. 424-438, 1969

[9] P. Jezzard, p. Matthews, and S. Smith, Functional Magnetic Resonance, New York, Oxford University Press, 2001

[10] N. A. Lazar, The Statistical Analysis of Functional MRI Data, Springer, 2008

[11] J. Geweke, Measurement of linear dependence and feedback between multiple time-series, J. Amer. Stat. Assoc., vol. 77, pp. 304-313, 1982

[12] Y. Saito and H. Harashima, Recent Advances in EEG and EMG Data Processing, Amsterdam, The Netherlands: Elsevier, pp. 133-146, 1981

[13] A. K. Seghouane, Quantifying information flow in fMRI using the Kullback-Leibler Divergence, IEEE International Symposium on Biomedical Imaging from Nano to Macro, pp. 1569-1572, 2011

[14] A. K. Seghouane, S. I. Amari, Identification of Directed Influence, Granger Causality, Kullback-Leibler and Complexity, Neural Computation, vol. 24, pp. 1722-1739, 2012

[15] C. Buchel and K. Friston, Modulation of connectivity in visual pathways by attention: Cortical interactions evaluated with structural equation modelling and fMRI, Cerebral Cortex, vol. 7, pp. 768-778, 1997

[16] P. A. Tyraskis and O. G. Jensen, Multichannel Linear Prediction and Maximum-Entropy Spectral Analysis Using Least Squares Modeling, IEEE Transactions on GeoSci. Remote Sen., vol. GE-23, pp. 101-109, 1985

[17] O. Friman, M. Borga, P. Lundberg, H. Knutsson, Exploratory fMRI analysis by autocorrelation maximization, Neuroimage, vol. 16(2), pp. 454-464, 2002

[18] I. T. Jolliffe, Principal Component Analysis, Springer, 2nd ed., 2002

[19] K. J. Friston, A. P. Holmes, and J. Ashburner and J. B. Poline, SPM8, http://www.fil.ion.ucl.ac.uk/spm/

[20] DR. Hardoon, S. Szedmak, J. Shawe-Taylor, Canconical correlation analysis: an overview with application to learning methods, Neural Comput., vol. 16(12), pp. 2639-2664, 2004

[21] A. K. Seghouane, Asymptotic bootstrap corrections of AIC for linear regression models, Signal Processing, vol. 90, pp. 217-224, 2010

[22] A. K. Seghouane and M. Bekara, A small sample model selection criterion based on Kullbacks symmetric divergence, IEEE Transactions on Signal Processing, vol. 52, pp. 3314-3324, 2004