Demonstrating causal links between fMRI time series using time-lagged correlation

Ben J. Smith, Patrice Delmas
Intelligent Vision Systems NZ
The University of Auckland Department of Computer Science
Auckland, New Zealand
bsmi080@aucklanduni.ac.nz

Sarina Iwabuchi, Ian J. Kirk
Research Centre for Cognitive Neuroscience
The University of Auckland Department of Psychology
Auckland, New Zealand

***Abstract*—An autoregressive modelling-based technique, Granger analysis, has increasingly been used for measuring causal connections between brain regions in fundamental magnetic resonance imaging brain imaging (fMRI). This paper summarises the possibilities and limitations of applying Granger analysis to fMRI. It also describes a replication of a previous theoretical study, and an application to spatial working memory currently under way. Previous researchers have described methods for detecting time-lagged correlation between neural activations in brain regions of interest (ROI)–often variants of 'Granger-causality analysis' (GCA). With appropriate caveats, GCA can draw inferences from time-lagged correlation about effective connectivity between ROIs in a way other popular methods do not. We replicated an existing theoretical model using a different non-linear function, and different method of combining a haemodynamic response function. We then examined whether GCA can estimate the direction of causation in brain regions shown to act together in spatial working memory tasks. Independent Component Analysis (ICA) was used to identify independent spatial components. Then, GCA tested the interactions between those components. The method identified causal relationships on replicated, artificially simulated data and in between extracted components in the data. Work is now underway to determine the implications of these relationships.**

Keywords—Granger causality analysis; independent component analysis; effective connectivity; fMRI; time-lagged correlation

I. INTRODUCTION

In neuroscience, "effective connectivity" is defined as the influence one neural system exerts over another" [1, 2]. *Granger-causality analysis* (GCA), which developed out of econometrics, is one method proposed to measure effective connectivity. Granger analysis measures the extent to which activity in a brain region can predict future activity in another brain region.

Section II discusses multivariate autoregressive theory, Granger theory, and the extent to which Granger-causality is truly causal. Section III compares Granger-causality analysis with other effective connectivity methods, while Section IV discusses Granger limitations from a theoretical perspective.

II. GRANGER THEORY

A. MAR Modelling/basic principle

GCA indicates the directionality of influence, predictivity, or *Granger-causality* between two (bivariate) or more (multivariate) regions. The usefulness of a time series Y in predicting the behavior of another time series X is the 'Granger-causality' of Y on X; if Y significantly predicts X, then Y is 'Granger-causal' of X. *Predictivity* can be measured by measuring time-lagged correlation between two time series.

In Granger analysis, the predictivity of one time series on another is measured by comparing the error r_{AR} of a univariate autoregressive model with the error r_{MAR} of a multivariate autoregressive model, as in

$$F(\mathrm{X} \rightarrow \mathrm{Y}) = \log(\frac{var(r_{\mathrm{MAR}})}{var(r_{\mathrm{AR}})}). \quad (1)$$

Autoregressive models are further discussed in the Appendix.

B. Multivariate Models

While the model presented in (1) generates a causality score for two time series, other models can infer causality in the context of more than two time series [3]. Extension of the model can be as simple as extending both models compared to include an extra time series or set of time series. The resulting Granger-causal score describes causality from one time series added in the second model, to the one time series being modelled One multivariate method is Conditional Granger Causality, which calculates the causal relationship between from series X to series Y, given series Z, described as $F(\mathrm{X} \rightarrow \mathrm{Y}|\mathrm{Z})$ [4].

C. Inferring causality

'Granger-causality' of the kind described above refers to the temporally predictive power that a second time series has in predicting the future value of a time series of interest. Granger justifies using the word "causality" by arguing that if all other plausible causes can reasonably be excluded, the one event remaining must be causal [2, 5]. Other causes could be excluded by using a multivariate GCA, which tests for causal relationship between two time series X and Y given multiple time series Z [6]. Alternatively, *a priori* knowledge could

indicate that there exists functional connection between exactly two regions of interest (ROIs).

A single bivariate Granger test does not exclude other causes potential causes. Hence, it can't independently describe causation – only predictivity. Where it is known there exists a functional connection between exactly two regions, then a bivariate Granger test could shed light on the effective connectivity of that relationship, including directionality and hence causation.

III. Advantages over Other Methods

GCA offers at least one advantage over both structural equation modelling (SEM) and dynamic causal modelling (DCM), two widely-used model-driven statistical measures for effective connectivity in fMRI data [7, 8]. These measures generally require *a priori* hypotheses about network model nodes and interactions between them [9]. In contrast, Granger analysis is a data-driven approach [7].

GCA's measurement of time-lag correlation captures temporal precedence relationships, unlike SEM, which assumes instantaneous connections, although DCM uses a coupling parameter which describes speed of change of nodes within a model. GCA can work with a large number of brain regions of interest (ROIs) [10]– facilitating its use as an exploratory method for effective connectivity–but statistical power required increases with the number of ROIs tested.

Although Granger analysis is not the only way an autoregressive model has been used to measure predictivity in fMRI, it may be currently the most widely used.

IV. Theoretical Considerations

There is ongoing controversy over the context in which GCA can usefully find predictive and causal connections in fMRI-based brain research. Although a number of recent studies [3, 7, 10-16] have used Granger analysis in an experimental context, there remain doubts about its applicability to reliably demonstrate effective connectivity. Few studies exist comparing Granger analysis to other effective connectivity measures [17]. Witt & Meyerand [8] compared DCM, Granger causality, and SEM, finding mixed results for Granger causality, but as they acknowledged, results may have been skewed because the modelled system was created using DCM.

Haemodynamic lag is the time delay between neural activity and the haemodynamic (blood flow) response. Variation including haemodynamic lag variation due to biological artefacts across brain regions can obscure neural lag variations. Correctly detecting neural lags is essential for good results using time-lagged correlation. Comparing difference in lag across experimental conditions could distinguish neural lag from haemodynamic lag [23]. Smoothing of neuronal activity still remains a problem [9]. Additionally, the poor temporal resolution of fMRI data at typical repetition times (RT; the time between one full-volume scan and the next one) impedes data collection. Where neural lag duration between two time series is less than the RT, the probability of successful lag detection is proportional to neural lag duration and inversely proportional to RT.

Attempts have been made to outline circumstances where lag-based methods can be useful when applied to fMRI data. Smith et al. [9] claims that if haemodynamic lag can be accurately estimated, neural lags greater than 100 ms could possibly be detected with a very low Repetition Time (RT) and SNR. It has been argued [18] that a 250 ms RT allows for detection of neural lags as small as 50 ms (with SNR = 6). supported that, but suggests that even the absence of haemodynamic confounds, where RT = 2 s and SNR = 1, neural lags may have to be as large as 500 ms to be detected. Where the haemodynamic lag difference opposed neural lag difference, neural lags would have to be larger in order to be detected.

Variation in the extent to which potentially causally-related time series are similar strongly affects statistical power [19]. Where time series mimicked each other exactly except for a lag, even very small neural lags could be detected; but where time series were mixtures of more than one signal, only one of which they shared, larger neural lags were required for detection.

It may be difficult to find Granger directionality using a RT of 2 s, as is used in the current article, if even a 500 ms delay is undetectable for many or most tasks. Even the bleak figures quoted above assume a lack of haemodynamic-neural lag confound, which can be very difficult to eliminate.

In spite of theoretical obstacles, significant results in many previous studies using Granger analysis suggest that real-world analyses of data using Granger causality is some use [3, 7, 10-16]. Many studies utilizing Granger causality have used an RT greater than 1.5 s. If haemodynamic-neural confounds can be ruled out (e.g., by controlling for lag with different conditions as suggested by [20]), significant results could provide further evidence for the utility of time-lag analysis of fMRI.

V. Method

Fundamental magnetic resonance imaging (fMRI) was used to map the presence of deoxygenated haemoglobin (BOLD data) in the brain, indicating changing neural activity over time. The resultant time series of volumes is separated into independent component spatial maps using independent component analysis (ICA). Components corresponding to time series which are most active in each experimental condition are assumed to correspond with cognitive processes triggered in the experiment. Time-lag correlation of those time series are measured using GCA.

Statistical significance is calculated through bootstrapping [21] with a null distribution of time series created for each time series using random sampling with replacement. Similar methods are widely employed in fundamental magnetic resonance imaging studies using GCA [10].

VI. Experiment 1: Simulation

Londei et al. [22] tested the GCA method using simulated fMRI data. Their test has been replicated with minor changes in the current project. Five ground truth 'independent components' – activation maps - have been created (Fig. 1). These represent activations of specific areas in an fMRI slice.

Activation of each over time is described by a set of equations (Fig. 1), and random noise is added. The activation of each independent spatial component can be isolated using independent component analysis (Fig. 2), specifically the fastICA 2.5 algorithm [11], which seeks to maximize nongaussianity using negentropy. In contrast to Londei et al. [22], a cubic rather than hyperbolic tangent non-linear function was used to converge on a maximization solution. A variety of conditions, including mean centring along each dimension, number of extracted eigenvectors, random or pre-determined initial guesses for the mixing matrix, and different model orders were tested.

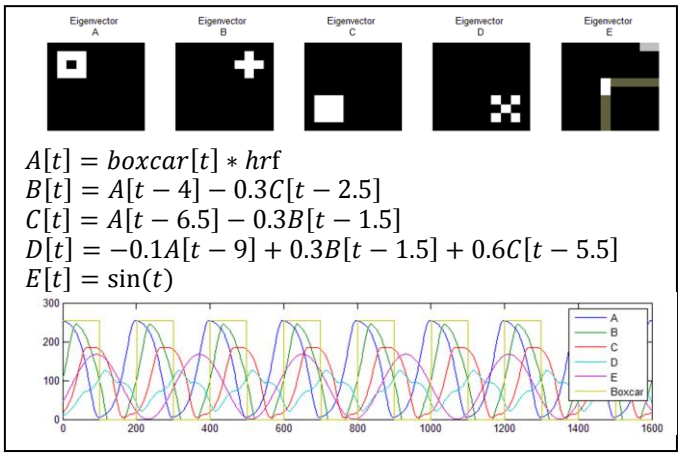

Fig. 1. Londei et al. (2006) simulation design, showing the images mixed and time series used to mix each of them.

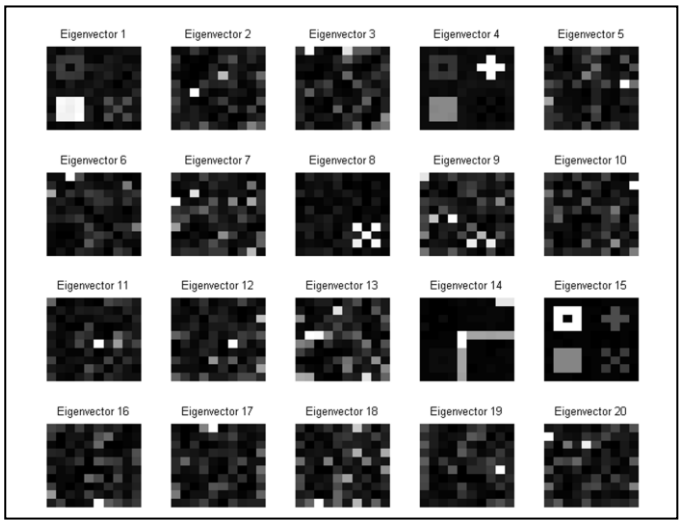

Fig. 2. Eigenvectors extracted using ICA from image time series generated from ground truth images in Fig. 1.

Results reported here use 20 initial eigenvectors extracted and random initial guesses, consistent with [22]. An autoregressive model order of 2 was used, as higher model orders would not increase the chance of detecting the short neural lags expected. In contrast to the formula shown in Fig. 1 for $A[t]$, the hrf convolution was here applied to each pixel in the images after all had been generated, not to the first time series. Each extracted spatial component corresponds to a time series describing the extent to which that component represents the data at any given time. Each time series can be compared to the reference time series (Fig. 4), and those most closely matching can be chosen for Granger analysis.

GCA plots the relationships between each of the time series, using a bivariate Granger analysis comparison (Fig. 3). Granger-causality represents relationships between different components; where these components can be identified with particular regions, 'Granger-causal' or predictive relationships are found between those regions (Fig. 5).

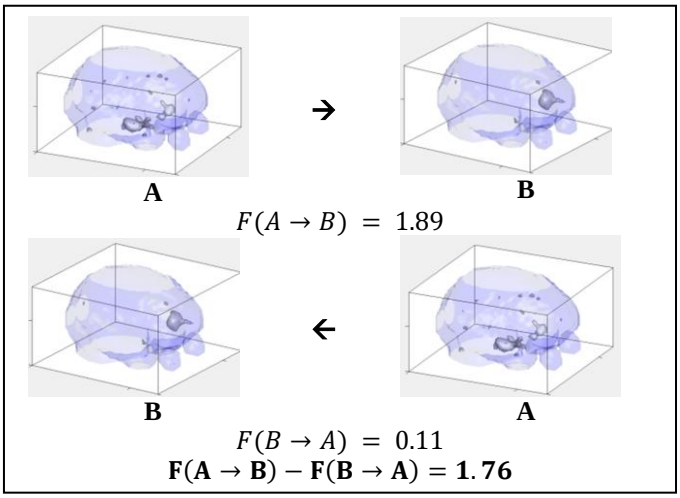

Fig. 3. Calculating predictivity from a Granger score (see (2)), using a comparison of Granger-causal scores of two independent components from the fMRI data set (mask and 99.9% threshold values shown). Here, the greater predictivity value of F(B → A) compared to F(A → B) suggests that B is predictive and possibly causal of A.

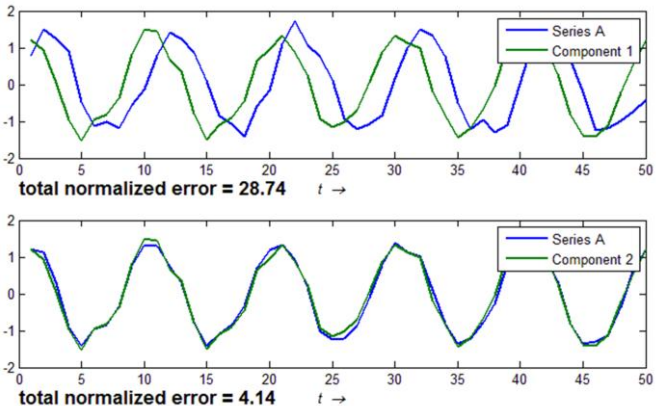

Fig. 4. Each component eigenvector is compared to each original series; the closest match eigenvector for each of the original series is chosen for subsequent Granger analysis.

VII. Experiment 2: Spatial Perception and Spatial Working Memory

A. Introduction

Working memory is a theorized cognitive system for temporary information storage and retrieval [23]. Distinct regions of activity during working memory processing spatial information (*spatial working memory*) have been identified in [24]. That study has been replicated here. Previous

unpublished research by Iwabuchi using the same dataset can be used to predict relationships that might be observed using the method described here. Specifically, connections between activity in the prefrontal cortex (particularly the right middle frontal gyrus) and right parietal cortex during spatial working memory should show consistent directionality or Granger-causality.

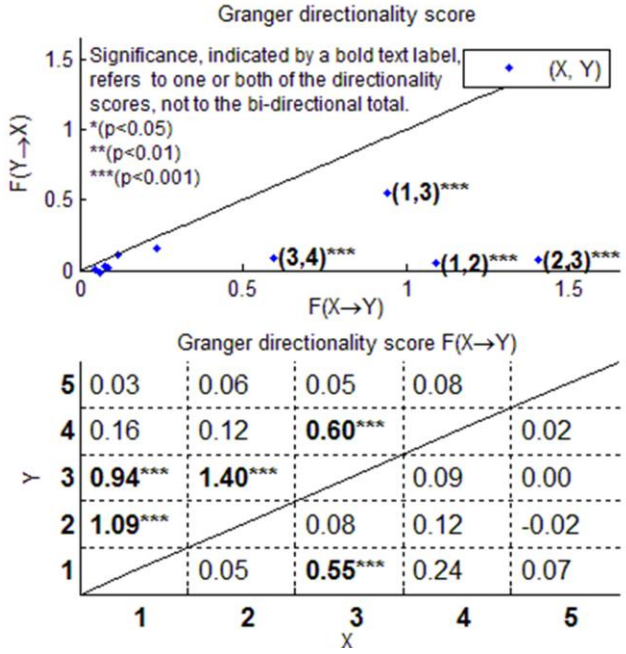

Fig. 5. Simulated data Granger results. Eigenvectors chosen from extracted eigenvectors (Fig. 2) according to their match (Fig. 4) with original timeseries (Fig. 1).

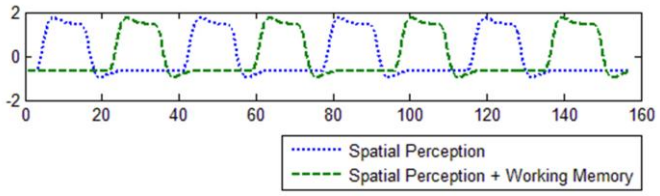

Fig. 6. Activation conditions, (A) spatial perception and (B) spatial working memory, describe the time period at which brain regions related to each are expected to show peak BOLD signal.

In *boxcar design*, a subject is exposed to repeated blocks of trials, where blocks alternate between experimental conditions. The boxcar design experiment exposed the subject on four sets of spatial perception only (control) and four sets of spatial perception and working memory (experimental) blocks. All blocks contained 15 trials of length 1990 ms each, break of approximately 15 s in between each block of trials. In each trial, the stimulus (a word on different locations on a screen) was presented for 500 ms and was followed by an inter-stimulus interval of 1500 ms. Subjects had to remember the word that was presented two trials previously, utilizing spatial working memory cognitive process.

Independent component analysis (FastICA 2.5, [25]) was applied separately to the pre-processed data from each of 20 subjects to extract 80 independent components, using a symmetric approach and a cubic non-linear convergence function. Condition durations were convolved with a haemodynamic response function approximating the haemodynamic response function of neural activity (Fig. 6). The associated time series of those spatial components were compared to the convolved activation time series of each of the two stimulus condition series (Fig. 4); the closest five matches to each condition time series were selected.

The Granger-causal relationships between each of the selected time series were compared (Fig. 7). If components chosen match regions associated with spatial perception and spatial working memory, then Granger-causal relationships between those components indicate Granger-causal relationships between working memory and spatial perception.

B. Results

Several significant Granger-causal relationships between components were found for each subject, including relationships between conditions. Exemplars from Subject 1 are shown in Fig. 7. For Subject 1, more causal relationships were found between time series associated with the perception and working memory condition than between the series in the perception-only condition. However, this trend was observed in just 17 of the 20 subjects (Table I.).

No consistent trend was seen indicating that either component category exerted a greater number of causal relations on any other.

C. Discussion and further work

Comparison of independent components to regions of interest has not yet been completed, so inferences can't yet be made about interactions between spatial perception and spatial working memory.

A number of significant causal relationships were observed between the independent components extracted (Fig. 7). Identifying the independent components with particular brain networks would enable inferences about the casual connectivity of those brain networks, assuming that regional variations in haemodynamic lag due to factors other neural activity could be accounted for.

Work is currently under way to combine data from different subjects; this could be done either prior to the component analysis (using ICA to combine subject data; see [26]) or subsequent to it, with the extracted time series. Different mean-centring and convergence functions will be examined. The "control" stimulus time series which is used to extract data could be associated with *both* the test conditions instead of only the control test condition, as at present.

VIII. SUMMARY

Distinct time series can extracted from fMRI-like data using ICA to separate the data into different components. These time series can be compared using Granger analysis to detect time-lagged predictive relationships [22]. These predictive relationships could even be described as causal relationships, where influences to external areas can be ruled out with prior knowledge of a system, or the use of Conditional Granger Causality.

Finding causal relationships – effective connectivity - in fMRI data can help determine how brain systems are structured and work together. Currently, more work is needed to determine whether GCA can detect predictive or causal relationships between spatial perception and spatial working memory, and the extent to which GCA can detect delays between regions.

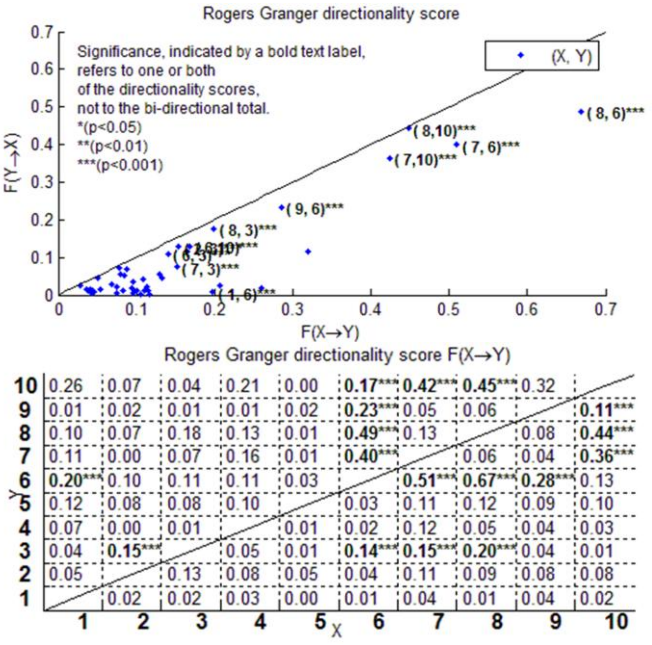

Fig. 7. The directionality of extracted components is shown above as a comparison between two axes (a high x-axis values indicates X is predictive of Y; a high y-axis value indicates Y is predictive of X). Components 1-5 are associated with the working memory condition; components 6-10 are associated with the spatial perception and working memory condition.

TABLE I. SIGNIFICANT CAUSAL RELATIONS BETWEEN EACH COMPONENT CATEGORY

Causal relation[a]	**F(P → P)**	**F(P → WM)**	**F(WM → P)**	**F(WM → WM)**
Subject #				
1	0	0	0	7
2	2	7	0	8
3	1	1	0	0
4	2	2	0	4
5	0	0	2	4
6	1	0	4	2
7	1	0	1	10
8	0	0	1	6
9	1	2	5	1
10	0	3	1	1
11	7	1	7	0
12	0	2	2	8
13	2	2	12	8
14	1	0	3	7
15	4	6	7	0
16	0	0	2	1
17	2	0	2	6
18	1	1	1	3
19	0	0	0	6
20	1	3	5	4

a. "P" indicates Spatial Perception condition; "WM" indicates Spatial Perception + Spatial Working Memory Condition

REFERENCES

[1] K. J. Friston, C. D. Frith, and R. S. J. Frackowiak, "Time-dependent changes in effective connectivity measured with PET," *Human Brain Mapping*, vol. 1, pp. 69-79, 1993.

[2] C. W. J. Granger, "Investigating Causal Relations by Econometric Models and Cross-spectral Methods," *Econometrica*, vol. 37, pp. 424-438, 1969.

[3] W. Liao, D. Mantini, Z. Zhang, Z. Pan, J. Ding, Q. Gong, Y. Yang, and H. Chen, "Evaluating the effective connectivity of resting state networks using conditional Granger causality," *Biological Cybernetics*, vol. 102, pp. 57-69, Jan 2010.

[4] M. Ding, Y. Chen, and S. L. Bressler, *Granger causality: basic theory and application to neuroscience*: Wiley-VCH Verlag, 2006.

[5] C. W. J. Granger, "Testing for causality," *Journal of Economic Dynamics and Control*, vol. 2, pp. 329-352, 1980.

[6] Z. Zhou, Y. Chen, M. Ding, P. Wright, Z. Lu, and Y. Liu, "Analyzing brain networks with PCA and conditional granger causality," *Human Brain Mapping*, vol. 30, pp. 2197-2206, Jul 2009.

[7] M. Wilke, K. Lidzba, and I. Krageloh-Mann, "Combined functional and causal connectivity analyses of language networks in children: A feasibility study," *Brain and Language*, vol. 108, pp. 22-29, Jan 2009.

[8] S. T. Witt and M. E. Meyerand, "The effects of computational method, data modeling, and TR on effective connectivity results," *Brain Imaging and Behavior*, vol. 3, pp. 220-231, 2009.

[9] S. M. Smith, K. L. Miller, G. Salimi-Khorshidi, M. Webster, C. F. Beckmann, T. E. Nichols, J. D. Ramsey, and M. W. Woolrich, "Network modelling methods for FMRI," *NeuroImage*, vol. 54, pp. 875-971, 2011.

[10] H. Chen, Q. Yang, W. Liao, Q. Gong, and S. Shen, "Evaluation of the effective connectivity of supplementary motor areas during motor imagery using Granger causality mapping," *NeuroImage*, vol. 47, pp. 1844-1853, Oct 2009.

[11] B. B. Biswal, D. A. Eldreth, M. A. Motes, and B. Rypma, "Task-dependent individual differences in prefrontal connectivity," *Cerebral Cortex*, vol. 20, pp. 2188-2197, Sep 2010.

[12] Q. Gao, X. Duan, and H. Chen, "Evaluation of effective connectivity of motor areas during motor imagery and execution using conditional Granger causality," *NeuroImage*, vol. 54, pp. 1280-1288, Jan 2011.

[13] A. Londei, A. D'Ausilio, D. Basso, C. Sestieri, C. Del Gratta, G. L. Romani, and M. O. Belardinelli, "Brain network for passive word listening as evaluated with ICA and Granger causality," *Brain Research Bulletin*, vol. 72, pp. 284-292, May 2007.

[14] C. A. Seger, E. J. Peterson, C. M. Cincotta, D. Lopez-Paniagua, and C. W. Anderson, "Dissociating the contributions of independent corticostriatal systems to visual categorization learning through the use of reinforcement learning modeling and Granger causality modeling," *NeuroImage*, vol. 50, pp. 644-656, Apr 2010.

[15] M. C. Stevens, G. D. Pearlson, and V. D. Calhoun, "Changes in the interaction of resting-state neural networks from adolescence to adulthood," *Human Brain Mapping*, vol. 30, pp. 2356-2366, Aug 2009.

[16] L. Q. Uddin, A. M. C. Kelly, B. B. Biswal, F. X. Castellanos, and M. P. Milham, "Functional connectivity of default mode network components: Correlation, anticorrelation,and causality," *Human Brain Mapping*, vol. 30, pp. 625-637, Feb 2009.

[17] J. B. Rowe, "Connectivity analysis is essential to understand neurological disorders," *Frontiers in Systems Neuroscience*, vol. 4, pp. 1-13, 2010.

[18] B. P. Rogers, S. B. Katwal, V. L. Morgan, C. L. Asplund, and J. C. Gore, "Functional MRI and multivariate autoregressive models," *Magnetic Resonance Imaging*, vol. 28, pp. 1058-1065, 2010.

[19] G. Deshpande, K. Sathian, and X. Hu, "Effect of hemodynamic variability on granger causality analysis of fMRI," *NeuroImage*, vol. 52, pp. 884-896, Sep 2010.

[20] A. Roebroeck, E. Formisano, and R. Goebel, "Mapping directed influence over the brain using Granger causality and fMRI," *NeuroImage*, vol. 25, pp. 230-242, 2005.

[21] B. Efron and R. Ribshirani, *An introduction to the bootstrap*. New York: Chapman & Hall, 1993.

[22] A. Londei, A. D'Ausilio, D. Basso, and M. O. Belardinelli, "A new method for detecting causality in fMRI data of cognitive processing," *Cognitive Processing*, vol. 7, pp. 42-52, Mar 2006.

[23] A. D. Baddeley, "Working memory," *Science*, vol. 255, pp. 556-559, 1992.

[24] C. Lycke, K. Specht, L. Ersland, and K. Hugdahl, "An fMRI study of phonological and spatial working memory using identical stimuli," *Scandinavian Journal of Psychology*, vol. 49, pp. 393-401, 2008.
[25] A. Hyvärinen and E. Oja, "A fast fixed-point algorithm for independent component analysis," *Neural Computation*, vol. 9, pp. 1483-1492, 1997.
[26] V. Schöpf, C. Windischberger, C. H. Kasess, R. Lanzenberger, and E. Moser, "Group ICA of resting-state data: a comparison," *Magnetic Resonance Materials in Physics*, vol. 23, pp. 317-325, 2010.

APPENDIX

Autoregressive models [18, 22] can predict the behaviour of a time series X at time t based on the values of the time series at previous times $X[t-1]$, $X[t-2]$, and so on, and error can be calculated as a measure of how accurate the model is (Fig. 8). A typical autoregressive model is

$$Y[t]_{AR} = a_1 Y[t-1] + a_2 Y[t-2] + \cdots + a_m Y[t-m] + r_{AR} \quad (2)$$

Error can be calculated as the variance of the residuals for an autoregressive model, over all time steps. A bivariate autoregressive model,

$$Y[t]_{BAR} = a_1 X[t-1] + a_2 X[t-2] + \cdots + a_m X[t-m] + b_1 Y[t-1] + b_2 Y[t-2] + \cdots + b_m X[t-m] + r_{BAR} \quad (3)$$

can predict the behaviour of a time series Y at time t, based on the values of the same time series Y, as well as another time series X, at $t-1$, $t-2$, and so on. Error can also be calculated for the bivariate autoregressive model in the same way as for the autoregressive model (Fig. 9). Then, errors between the autoregressive model (based on X), and the bivariate autoregressive model (based on X and Y) can be compared (1), and the result describes causal or predictive power of X on Y. The causal or predictive power of the reverse relationship, Y on X can then be calculated and compared to the causal power of X on Y to find the direction of causality between time series X and Y (Fig. 3).

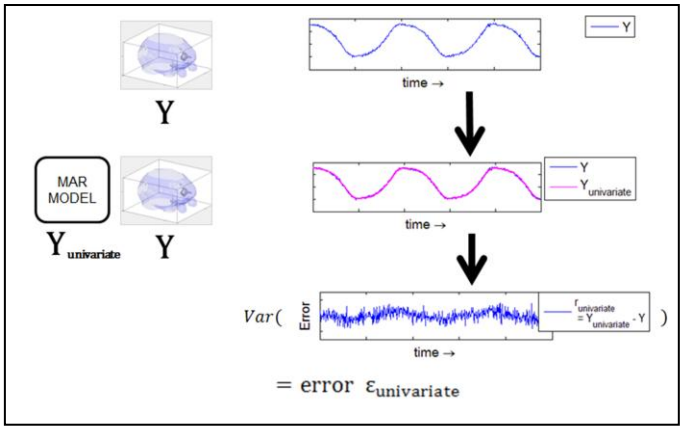

Fig. 8. Calculating the error in the univariate autoregressive model.

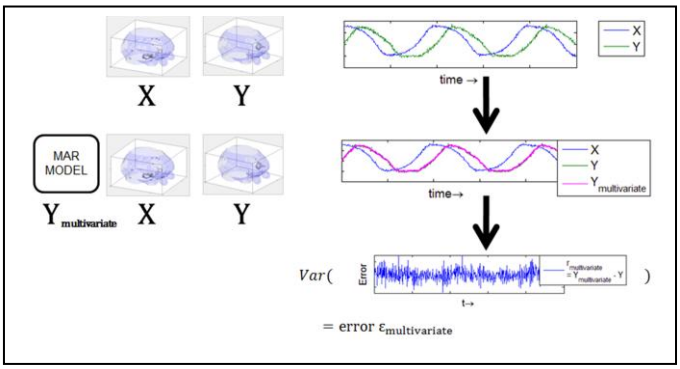

Fig. 9. Calculating the error in the multivariate autoregressive model.