Demonstrating causal links between time series associated with fMRI data using Granger analysis

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

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

***Abstract*— An econometric autoregressive modelling-based technique, Granger analysis, has increasingly been used for measuring causal connections between brain regions in fundamental magnetic resonance imaging (fMRI). This paper summarises the possibilities and limitations of applying Granger analysis to fMRI, describes a replication of a previous theoretical study, and an application currently under way. Previous researchers have described methods for detecting time-lagged correlation between activations in region of interest (ROI)– primarily variants on 'Granger causality' (GC). With appropriate caveats, GC can draw inferences from temporal precedence 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 method of combining an hrf. We also examined whether GC time-lag analysis can estimate the direction of causation in brain regions known to act together in spatial working memory tasks. Independent Component Analysis (ICA) is used to identify independent spatial components, whose interactions are then examined using the Granger causality method. The method successfully identified causal relationships on replicated, artificially simulated data and in between extracted components in the data. Work is now underway to determine the implications for these relationships.**

Keywords-Granger analysis; independent component analysis; effective connectivity; fMRI

I. INTRODUCTION

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

II. GRANGER THEORY

A. MAR Modelling/basic principle

Granger analysis indicates the directionality of influence, predictability, 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.

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_{BAR} of a bivariate autoregressive model, as in

$$F(\mathrm{X},\mathrm{Y}) = \log\left(\frac{var(r_{\mathrm{BAR}})}{var(r_{\mathrm{AR}})}\right). \quad (1)$$

B. Multivariate Models

While the model presented above generates causality scores for two time series, other models can infer causality in the context of more than two time series [3]. One such method is Conditional Granger Causality [4], which calculates the causal relationship between two time series while controlling for a third;

C. Statistical significance

Statistical significance can be calculated through the bootstrap technique [5], widely employed in fundamental magnetic resonance imaging (fMRI) studies [6] using Granger analysis.

D. 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 [2, 7] justifies and defines use of the word "causality", arguing that if all other plausible causes can reasonably be excluded, the one event remaining must be causal. Other plausible causes could be reasonably excluded either by using a multivariate Granger analysis (e.g., [8]), which tests for causal relationship between two time series X and Y given time series Z, or by some a priori knowledge that there exists functional connection between two regions and only two of interest.

A single bivariate Granger test, such as that described above, does not itself rule out other causes potential causes; it may only show that time-lagged correlated time series Y predicts X rather than necessarily causes it. However, where it is known there exists a functional connection primarily between two, and only two, regions, then such a bivariate Granger test could shed light on the effective connectivity of that region including the directionality of that connection.

III. Advantages over Other Methods

Granger analysis offers at least one advantage over both structural equation modelling (SEM) and dynamic causal modelling (DCM), two widely-used statistical methods for measuring effective connectivity in fMRI data. SEM and DCM require a priori hypotheses about the network nodes and interactions between them [6], while Granger analysis is a data-driven approach. Although the bivariate technique described above cannot be relied upon to describe *causality* without assuming some previous theory regarding interactions, a multivariate Granger technique can do; furthermore, the bivariate approach can still describe predictivity, and can indicate direction of effective connectivity where a connection is known to exist.

Granger analysis is designed to take advantage of temporal precedence, which is a feature of data not captured by SEM, which assumes instantaneous connections, although DCM uses a coupling parameter which describes speed of change of nodes within a model. Furthermore, Granger analysis can work with a large number of ROIs, [6]– another reason that it is a good exploratory method for effective connectivity – although the method does lose power with increasing numbers of ROIs.

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

IV. Experiment 1: Simulation

Londei et al. [9] described a Granger analysis 'proof of concept' using simulated fMRI data, which has been replicated with minor changes in the current project. Here, five ground truth 'independent components' – activation maps - have been created. Activation of each over time is described by a set of equations (Figure 1), and random noise is added. The activation of each independent spatial component can be isolated using independent component analysis (Figure 3), specifically the fastICA 2.5 algorithm [11], which seeks to maximize nongaussianity using negentropy. In contrast to Londei et al. [9], 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, altering the number of extracted eigenvectors, random or pre-determined initial guesses for the mixing matrix, and different model orders were tested. Results reported here use 20 initial eigenvectors extracted, random initial guesses, and an autoregressive model order of 2. In contrast to the formula shown in Figure 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 (Figure 4), and those most closely matching can be chosen for Granger analysis.

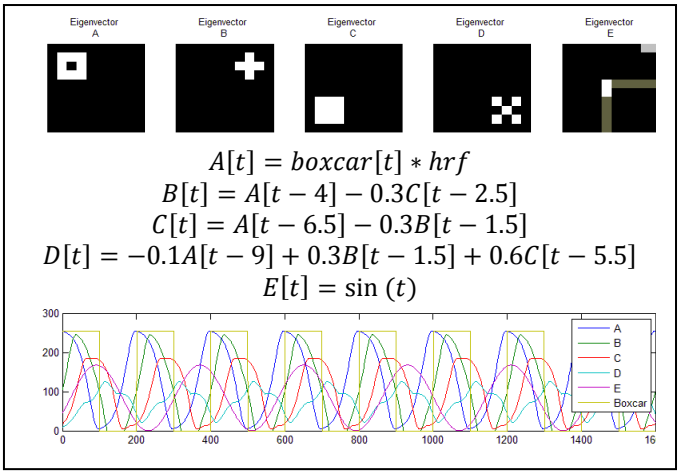

Figure 1. Londei et al. (2006) simulation design, showing the images mixed and time series used to mix each of them.

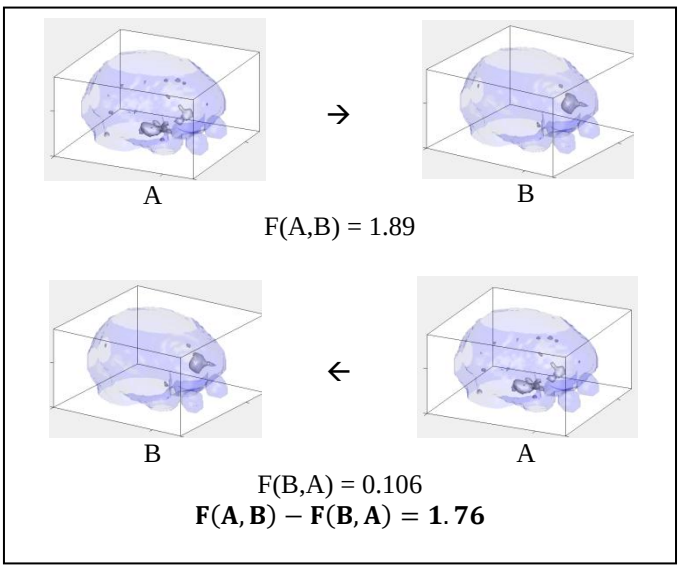

Figure 2. Calculating predictivity from a Granger score (see Equation (2)), using a comparison of Granger-causal scores of two independent components (mask and 99.9% threshold values shown).

We apply the Granger analysis described previously to examine the relationships between each of the time series, using a bivariate Granger analysis comparison (Figure 2). Granger causal relationships represent relationships between different components; where these components can be identified with particular regions, 'Granger-causal' or predictive relationships are found between those regions (Figure 5).

V. Theoretical Considerations

The context in which Granger analysis can usefully find predictive and causal connections in fMRI-based brain research is subject to ongoing debate. Although a number of recent studies [3, 6, 10-18] have used Granger analysis in an experimental context, there remain serious questions about its applicability to reliably demonstrate effective connectivity. There exist few studies comparing Granger analysis to other

effective connectivity measures [19]. Witt & Meyerand [20] compared DCM, Granger causality, and SEM, finding mixed results for Granger causality, but as they acknowledged, results may have been biased because the modelled system was created using DCM.

Haemodynamic lag variation due to biological artefacts across brain regions can obscure haemodynamic lag variations due to difference in neural activation time; comparing difference in lag across experimental conditions has been proposed as a solution to this [23], although this doesn't address other problems introduced by the haemodynamic lag, such as smoothing of neuronal activity [21]. The poor temporal resolution of fMRI data at typical repetition times (RT) impedes data collection. Where neural lag durations between two time series are less than the RT, the probability that the lag will be detected decreases proportional to neural lag decreases, and increases proportionally to a decrease in RT.

Attempts have been made to outline circumstances where lag-based methods can be useful when applied to fMRI data. Smith et al. [21] claims that neural lags greater than 100 ms could possibly be detected with a very low repetition time (RT) and signal-to-noise ratio (SNR), where the haemodynamic lag can be accurately estimated. It has been argued that a 250 ms RT allows for detection of neural lags as small as 50 ms (with SNR=6) [22]. Data in [23] supported that, but suggests that even the absence of hrf confounds, where RT = 2 s and SNR=1, neural lags may have to be as large as 500 ms to be detected significantly. Where the hrf opposed the neural lags, lags would have to be larger in order to be detected.

Variation in the extent to which time series in question are similar strongly affects statistical power [23]. 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 rate of 2 s, such as that which was used in the data which is described below, as a 500 ms delay is largely unrealistic for many or most tasks. Even the bleak figures quoted above assume a lack of hrf confound, which can be very difficult to eliminate. However, significant results in many previous studies using Granger analysis leave open the possibility that real-world analyses of data using Granger causality may be of some use. Many studies utilizing Granger causality have used an RT greater than 1.5 s. If significant results could be found with the existing data, and hrf confounds can be ruled out, the results could be interpreted as further evidence for the utility of time-lag analysis of fMRI.

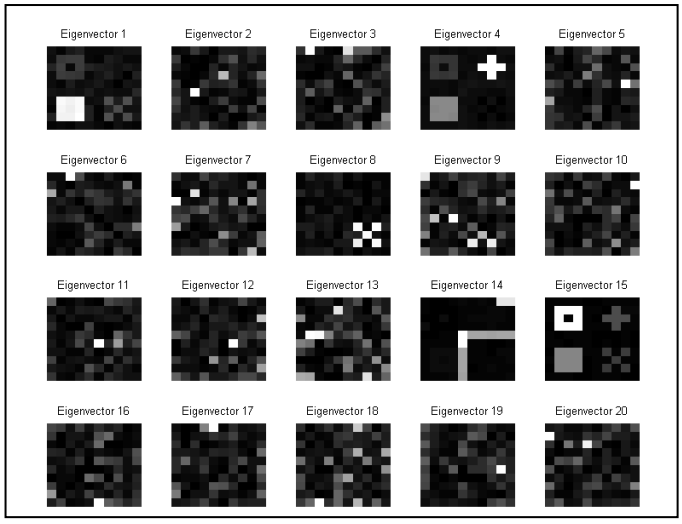

Figure 3. Eigenvectors extracted using ICA.

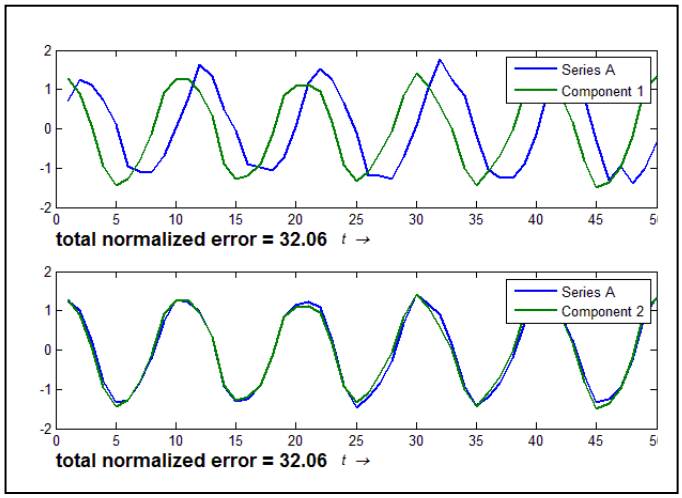

Figure 4. Each component eigenvector is compared to each original series; the closest match eigenvector for each of the original series is chosen for subsequent Granger analysis.

VI. Experiment 2: Spatial Perception and Working Memory

A. Introduction

Distinct regions of activity during spatial working memory have been identified in [24], and we have replicated this study here. Previous unpublished research by Iwabuchi using the dataset can be used to predict relationships that will be observed 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.

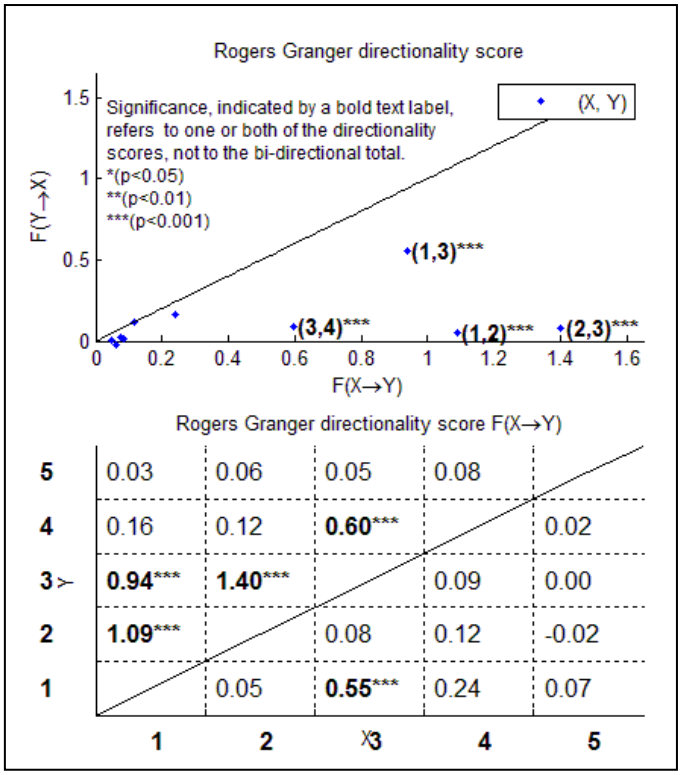

Figure 5. Simulated data Granger results. Eigenvectors chosen from extracted eigenvectors (Figure 3) according to their match (Figure 4) with original timeseries (Figure **1**).

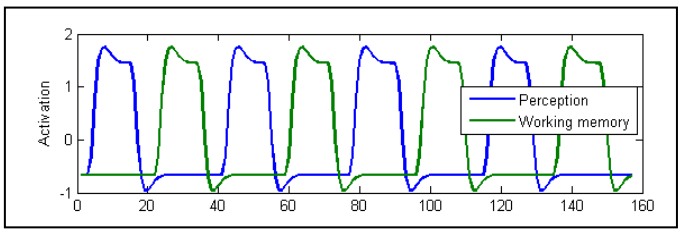

Figure 6. Activation conditions, (A) spatial perception and (B) spatial working memory, describe the time period at which brain regions related to each are expected to show peak BOLD signal.

The experiment was set up on boxcar design on 4 sets of spatial perception (control) and 4 sets of spatial perception working memory (experimental) blocks, all of 15 trials of length 1990 ms each, with a subject-determined break of approximately 15 s in between them. In each trial, the stimulus was presented for 500 ms and was followed by an inter-stimulus interval of 1500 ms.

Independent component analysis (FastICA 2.5) applied separately to the pre-processed data from each subject to extract 80 independent components, using a symmetric approach and a cubic non-linear convergence function ("pow3"). The associated time series of those spatial components were compared to the convolution of each of the time series of activation (Figure 6) of each of the two stimulus condition series (Figure 4); the closest 5 matches in each case were selected. The Granger-causal relationships between each of the selected time series were compared. If the matching to stimulus condition series can accurately identify regions associated with activity in the mental tasks which they represent – respectively, spatial perception and spatial working memory - then causal relationships across those conditions indicate Granger-causal relationships between networks associated with the working memory task and those associated with the spatial perception task.

B. Results

Several significant Granger-causal relationships between components were found for each subject, including relationships between conditions. Exemplars from one subject are shown in Figure **8**Figure 7.

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 (Figure 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 a form of group ICA; see [25]) 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.

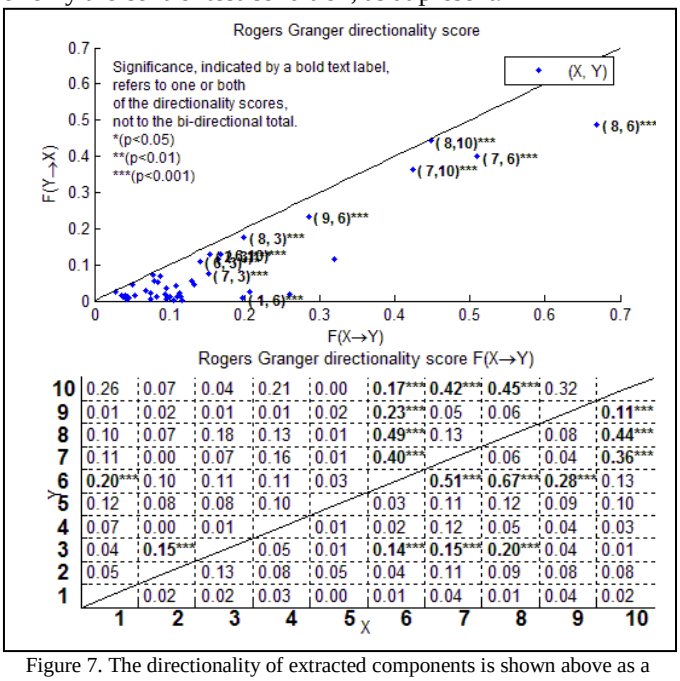

Figure 7. The directionality of extracted components is shown above as a comparison between two axes (a high x-axis values indicates X is predictive of Y; a high y-axis value indicates Y is predictive of X). Components 1-5 are associated with the working memory condition; components 6-10 are associated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

VII. 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 [9]. 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 Granger-causal relationships between spatial perception and spatial working memory can be detected using Granger causality, and the extent to which Granger causality is able to detect delays between regions.

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

[6] 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.

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

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] S. Graham, E. Phua, C. S. Soon, T. Oh, C. Au, B. Shuter, S.-C. Wang, and I. B. Yeh, "Role of medial cortical, hippocampal and striatal interactions during cognitive set-shifting," *NeuroImage*, vol. 45, pp. 1359-1367, May 2009.

[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] J. R. Sato, A. Fujita, E. F. Cardoso, C. E. Thomaz, M. J. Brammer, and E. Amaro, Jr., "Analyzing the connectivity between regions of interest: An approach based on cluster Granger causality for fMRI data analysis," *NeuroImage*, vol. 52, pp. 1444-1455, Oct 2010.

[15] 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.

[16] 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.

[17] 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.

[18] 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.

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

[20] 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.

[21] 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.

[22] 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.

[23] 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.

[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] 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 [9, 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 (Figure **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 timesteps. A bivariate autoregressive model,

$$Y[t]_{BAR} = a_1X[t-1] + a_2X[t-2] + \cdots + a_mX[t-m] + b_1Y[t-1] + b_2Y[t-2] + \cdots + b_mX[t-m] + r_{BAR} \quad (3)$$

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 X[t-1], X[t-2], and so on. Error can also be calculated for the bivariate autoregressive model in the same way as for the autoregressive model (Figure **9**). Then, errors between the autoregressive model (based on T), 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 directionality of causality between time series *X* and *Y* (Figure 2).

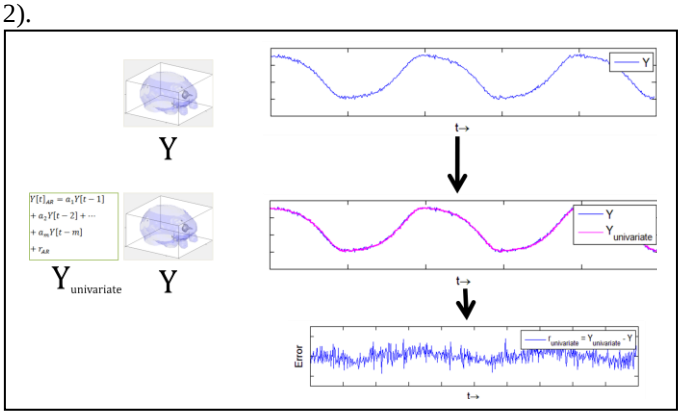

Figure 8. Calculating the error in the univariate autoregressive model.

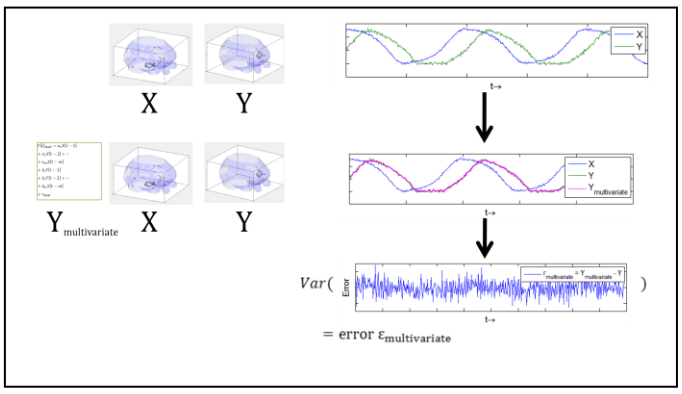

Figure 9. Calculating the error in the multivariate autoregressive model.