ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 25 (2005) 478–490

Evaluating frequency-wise directed connectivity of BOLD signals applying relative power contribution with the linear multivariate time-series models

Okito Yamashita,[a,b,*] Norihiro Sadato,[c,d] Tomohisa Okada,[e] and Tohru Ozaki[a]

[a]*Institute of Statistical Mathematics, Minami-Azabu 4-6-7, Mintato-ku, Tokyo 106-8569, Japan*
[b]*ATR Computational Neuroscience Laboratories, Kyoto, Japan*
[c]*National Institute for Physiological Sciences, Okazaki, Japan*
[d]*JST (Japan Science and Technology Corporation)/RISTEX (Research Institute of Science and Technology for Society), Kawaguchi, Japan*
[e]*Institute of Biomedical Research and Innovation, Kobe, Japan*

Received 10 May 2004; revised 29 November 2004; accepted 30 November 2004
Available online 29 January 2005

In this article, we propose a statistical method to evaluate directed interactions of functional magnetic-resonance imaging (fMRI) data. The multivariate autoregressive (MAR) model was combined with the relative power contribution (RPC) in this analysis. The MAR model was fitted to the data to specify the direction of connections, and the RPC quantifies the strength of connections. As the RPC is computed in the frequency domain, we can evaluate the connectivity for each frequency component. From this, we can establish whether the specified connections represent low- or high-frequency connectivity, which cannot be examined solely using the estimated MAR coefficients. We applied this analysis method to fMRI data obtained during visual motion tasks, confirming previous reports of bottom-up connectivity around the frequency corresponding to the block experimental design. Furthermore, we used the MAR model with exogenous variables (MARX) to extend our understanding of these data, and to show how the input to V1 transfers to higher cortical areas.

Keywords: Causality; Frequency-wise directed connectivity; Multivariate autoregressive model (with exogenous variables); Relative power contribution

Introduction

'Functional specialization' and 'functional integration' are two fundamental principles of brain organization. Functional specialization suggests that a cortical area is specialized for certain aspects of perceptual or motor processing. The cortical infrastructure supporting a single function might then involve many specialized areas, the cooperation of which is mediated by the functional integration of these areas.

In functional magnetic-resonance imaging (fMRI), statistical parametric mapping (SPM) is a well-established procedure used to localize the cortical areas related to a certain cognitive process (Friston et al., 1995; Worsley et al., 2002). Much of the experimental evidence regarding cortical specialization has been analyzed using SPM. Nowadays, increased attention is paid to the integration of specialized areas, and the development of statistical methods that can be used to analyze this integration is highly desired (Lee et al., 2003). Integration within a distributed system can be understood in terms of 'effective connectivity'. Effective connectivity originally was defined as "the influence that one neural system exerts over another" (Friston, 1994). Related to this definition of effective connectivity, in this article, we use the term directed connectivity to define any directed relationship of the BOLD signals between specific cortical areas and this definition could be operational or statistically descriptive.

One of the most common methods for the evaluation of effective connectivity is the structural equation model (SEM) proposed by McIntosh and Gonzalez-Lima (1994). Although the SEM works well to quantify the *strength* of the influence one area exerts on another, the *direction* of the connections between the regions of interest (ROIs) must be determined a priori. This limitation exists because SEM exploits only the instantaneous covariance structure and therefore discards detailed temporal information. However, it is possible to overcome this limitation by using causal analysis in the field of time-series analysis. The comprehensive definition of causality was introduced first by Granger (1969). His definition was based on the common belief that "the cause should precede the result", and thus he defined 'X_t causes Y_t', such that the past of a time series X_t can predict the future of another time series Y_t. The primary point of this

* Corresponding author. Fax: +81 3 5421 8751.
E-mail address: yamashi@ism.ac.jp (O. Yamashita).
Available online on ScienceDirect (www.sciencedirect.com).

1053-8119/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2004.11.042

definition is that a directed relationship between X_t and Y_t is based on their temporal order.

In fMRI, few studies have attempted to establish causality based on time series. Granger causality in relation to fMRI effective connectivity has been explicitly introduced for the first time by Goebel et al. (2003). They have proposed Granger causality map applying the multivariate autoregressive (MAR) model and Geweke's measure (Geweke, 1982) in order to investigate the directed interactions between a reference region and the remaining brain regions. Harrison et al. (2003) have also done some pioneering work using the Granger causality concept of effective connectivity. In this literature, the MAR model (including the bilinear term) is fitted to a set of time series obtained from the ROIs, and the direction of connections is specified by testing whether the estimated coefficients are significantly larger than zero. In one paper (Lahaye et al., 2003), the authors reported that the inclusion of temporal information improved the model's fit significantly; this group defined their connectivity measure based on the likelihood-ratio test for a bivariate time series. Finally, the dynamical causal modeling (DCM) proposed by Friston et al. (2003) is a novel approach that emphasizes the construction of a reasonably realistic neuronal model of interacting cortical regions.

In this article, we focus on a statistical method to evaluate directed connectivity of BOLD signals. Initially, the MAR model was fitted to a set of time series in the ROIs. Then, the relative power contribution (RPC) as proposed by Akaike (1968) was calculated using the estimates of the parameters. The RPC quantifies the contribution of a specific time series as distinct from the others. Hence, we obtained a directed measure for every pair of connections between the ROIs. This work can be considered an extension of Harrison's method, in the sense that it is possible to evaluate both the direction and the strength of connectivity at the same time. In addition, this RPC is extended to the MAR model with exogenous variables (MARX), so, for example, the box-car function in a block design can be established as the input to a cortical area separate from white noise (intrinsic noise).

Methods

We used a three-step procedure to evaluate the directed connectivity in our fMRI data: firstly, specification of the ROIs and extraction of time series from the ROIs; secondly, identification of the model; and thirdly, post-processing using the estimated parameters to quantify connectivity.

To pick up time series from ROIs, the following processes in SPM software (http://www.fil.ion.ucl.ac.uk/spm/) were used: slice-timing correction for the adjustment of the data-acquisition timing; realignment to correct for subject movement; and normalization to better define cerebral areas. Statistical analysis was conducted with a general linear model and a representative time series from each region, which was defined as the first eigen vector of an ensemble of time series from voxels within a 6-mm-radius sphere, where the center voxel of each region is the local maximum of t-MAP. Some reports have explained the possible effects of the various ways of specifying ROIs (Gavrilescu et al., 2004; Goncalves and Hall, 2003); however, this is beyond the scope of the current paper.

In previous papers, the second step corresponds to the use of the MAR model with a bilinear term (Harrison et al., 2003), a higher-order polynomial (Lahaye et al., 2003), or the directed graph of the SEM assumed by an analyst (McIntosh and Gonzalez-Lima, 1994). Here, we employed the MAR model,

$$\mathbf{Z}_t = \sum_{k=1}^{p} A(k)\mathbf{Z}_{t-k} + \varepsilon_t \tag{1}$$

or the MARX model,

$$\mathbf{Z}_t = \sum_{k=1}^{p} A(k)\mathbf{Z}_{t-k} + \mathbf{w}S_t + \varepsilon_t \tag{2}$$

where ε_t is a white-noise process (innovation) with mean zero and variance C_ε. The $d \times 1$ vector $\mathbf{Z}_t = (z_{1,t}, \cdots, z_{d,t})'$ comprises the ensemble of d time series (adjusted to have a mean of zero) in the ROIs. S_t is an external input of a scalar process and $\mathbf{w}$ is a vector of the size $d \times 1$ that determines the magnitude of S_t. The AR order is denoted by p.

After estimating the parameters from the data (see Appendix A) to quantify connectivity, it is necessary to manipulate these estimates. This is because not a single coefficient, but rather p coefficients $A_{ij}(1), \cdots, A_{ij}(p)$, are associated with the connectivity from $z_{j,t}$ to $z_{i,t}$. In the field of time-series analysis, various measures and tests of causality have been proposed: for example, the measures in frequency/time domain have been proposed for (block) bivariate stationary time series purely based on the Granger's definition of causality (i.e., predictability) by Geweke (1982) and Hosoya (1991), and their measures has been extended to the multivariate stationary time series, respectively (Geweke, 1984; Hosoya, 2001). Akaike (1968) has proposed the RPC in frequency domain, which measures the causal relationship of every pair of the variables in a MAR model of any dimension by estimating the parameters only once. In the field of biological engineering, some frequency-domain measures, such as the directed coherence (DC) (Saito and Harashima, 1981), the partial directed coherence (PDC) (Baccala and Sameshima, 2001), and the directed transfer function (DTF) (Kaminski et al., 2001), have been proposed. As the RPC seems to have intuitive interpretation of the innovation contributions (see Discussion) and makes use of the whole covariance structure of ROIs, we applied this measure (see Fig. 1 for schematics of the RPC).

Relative power contribution

The parametric spectrum of the MAR model can be calculated from the AR coefficient $A(i)$, $(i = 1, \cdots, p)$ and the innovation covariance matrix C_ε as follows (see any standard textbook in time-series analysis for more details; for example, Chapter 3 of Shumway, 2000),

$$P_{\mathbf{Z}}(f) = H(f)C_\varepsilon \overline{H}(f)', \tag{3}$$

where $\overline{H}$ and H' are the complex conjugate and transpose of a matrix H, respectively. f denotes the frequency ranging from 0 to 0.5 (the highest frequency 0.5 corresponds to $1/(2TR)$ Hz). The power-spectrum matrix is denoted by $P_{\mathbf{Z}}(f)$, of which the diagonal

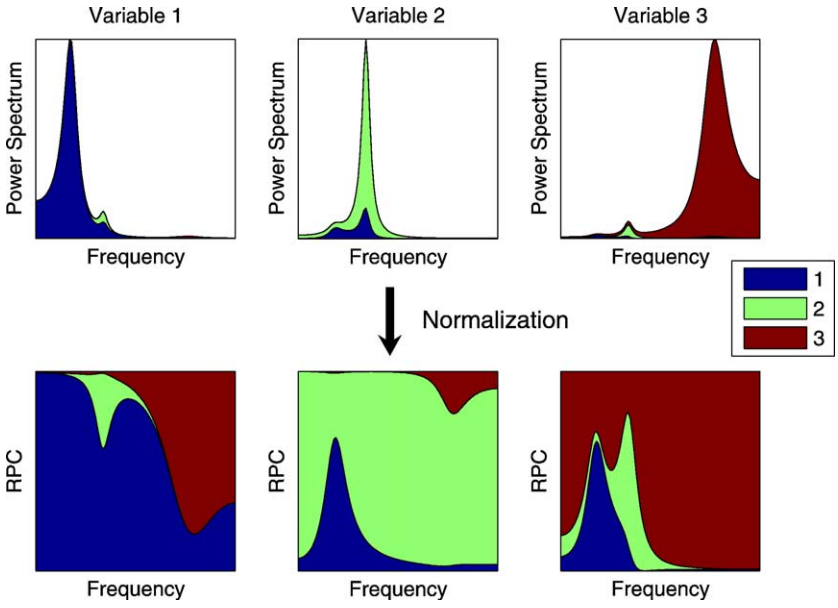

Fig. 1. Schematic figure of the RPC measure in the case of three variables. The power spectrum of each variable can be decomposed to the power contribution from every variable which is represented by the shaded areas in each plot. The RPC measure can be obtained by normalizing these contributions.

entries $P_{ii}(f)$ and nondiagonal entries $P_{ij}(f)$ represent the power spectrum of the ith time series $z_{i,t}$ and the cross spectrum between $z_{i,t}$ and $z_{j,t}$, respectively. The matrix $H(f)$, which describes the transfer functions (frequency response) from a set of the innovation process ε_t to a set of the time series $\mathbf{Z}_t$, is given by

$$H(f) = \left(I - \sum_{j=1}^{p} A(j) e^{\{-i2\pi fj\}} \right)^{-1} \tag{4}$$

where $i = \sqrt{-1}$.

Here, the innovation processes are assumed to be mutually uncorrelated, that is, the matrix C_ε is diagonal,

$$C_\varepsilon = \mathrm{diag}\left(\sigma_1^2, \cdots, \sigma_d^2\right). \tag{5}$$

Then substituting Eq. (5) into Eq. (3) leads to

$$P_{ii}(f) = \sum_{j=1}^{d} \mid H_{ij}(f)^2 \mid \sigma_j^2, \quad (i = 1, \cdots, d). \tag{6}$$

This equation indicates that the power spectrum of z_{it} at frequency f can be decomposed to d terms $|H_{ij}(f)|^2\ \sigma_j^2$, $(i = 1; \cdots, d)$, each of which can be interpreted as the power contribution of the jth innovation $\varepsilon_{j,t}$ transferring to $z_{i,t}$ via the transfer function H_{ij} (f). Thus, $|H_{ij}(f)|^2\ \sigma_j^2$, $(i = 1; \cdots, d)$ can be regarded as the power contribution of the innovation $\varepsilon_{j,t}$ on the power spectrum of $z_{i,t}$. Finally, the RPC is defined as a ratio of each contribution to the power spectrum $P_{ii}(f)$:

$$R_{j \to i}(f) = \frac{\mid H_{ij}(f)^2 \mid \sigma_j^2}{P_{ii}(f)}. \tag{7}$$

The RPC is computed for every pair of i and j, and for every frequency f, hence the RPC gives a quantitative measurement of the strength of every connection for each frequency component. Note that for any i, j, or f, the ratio $R_{j \to i}(f)$ always ranges from 0 to 1.

Extended relative power contribution

In real applications, we sometimes observe that the specified innovations are correlated. Since the RPC assumes that the innovations are mutually uncorrelated, this measure cannot be applied directly. One ad hoc way might be to neglect nondiagonal elements of the innovation correlation matrix and apply the original RPC. The other way is to employ the extended RPC (ERPC) proposed by Tanokura and Kitagawa (2004), which is the extension of the RPC to the case when the innovations are modestly correlated. The ERPC supposes that the correlated innovation consists of inseparable cross-correlated component and its own uncorrelated component. Hence, the power spectrum of each time series is decomposed to the contributions of uncorrelated portion of the innovation as well as those of inseparable cross-correlated portion.

For the derivation of the ERPC measure, we firstly consider the following decomposition of the innovation correlation matrix R_ε,

$$R_\varepsilon = \sum_{i=2}^{d} \sum_{j=1}^{i-1} \mid \rho_{ij} \mid \mathbf{J}_{ij} \mathbf{J}_{ij}^{'} + \mathrm{diag}(\tau_1, \cdots, \tau_d) \tag{8}$$

where ρ_{ij} is a ij component of R_ε, $\mathbf{J}_{ij}$ is a vector of size $d \times 1$ defined as,

$$\mathbf{J}_{ij} = \left[0, \cdots 0, \overset{i}{1}, 0 \cdots 0, \overset{j}{1}, 0 \cdots 0 \right]' \quad \text{if } \rho_{ij} \geq 0$$

$$\mathbf{J}_{ij} = \left[0, \cdots 0, \overset{i}{1}, 0 \cdots 0, \overset{j}{-1}, 0 \cdots 0 \right]' \quad \text{if } \rho_{ij} < 0$$

and $\tau_i = 2 - \sum_{j=1}^{d} |\rho_{ij}| \, (i = 1, \cdots, d)$. Note that this decomposition is uniquely determined as easily shown from the following example,

$$\begin{bmatrix} 1 & 0.5 & 0.3 \\ 0.5 & 1 & -0.2 \\ 0.3 & -0.2 & 1 \end{bmatrix} = 0.5 \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} + 0.3 \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} + 0.2 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix} + \mathrm{diag}(0.2, 0.3, 0.5).$$

Since the covariance matrix and the correlation matrix are related by the equation $C_\varepsilon = \mathrm{diag}(\sigma_1, \cdots, \sigma_d) \cdot R_\varepsilon \cdot \mathrm{diag}(\sigma_1, \cdots, \sigma_d)$, we obtain the decomposition of the covariance matrix (from Eq. (8)) as follows,

$$C_\varepsilon = \sum_{i=2}^{d} \sum_{j=1}^{i-1} |\rho_{ij}| \, \hat{\mathbf{J}}_{ij} \hat{\mathbf{J}}_{ij}' + \mathrm{diag}\left(\tau_1 \sigma_1^2, \cdots, \tau_d \sigma_d^2\right) \tag{9}$$

where $\hat{\mathbf{J}}_{ij}$ is a vector of size $d \times 1$ defined as

$$\hat{\mathbf{J}}_{ij} = \left[0, \cdots 0, \overset{i}{\sigma_i}, 0 \cdots 0, \overset{j}{\sigma_j}, 0 \cdots 0\right]' \text{ if } \rho_{ij} \geq 0$$

$$\hat{\mathbf{J}}_{ij} = \left[0, \cdots 0, \overset{i}{\sigma_i}, 0 \cdots 0, \overset{j}{-\sigma_j}, 0 \cdots 0\right]' \text{ if } \rho_{ij} < 0.$$

From Eqs. (3) and (9), the power spectrum of the ith time series in the case of the correlated innovations can be represented as:

$$P_{ii}(f) = \sum_{j=2}^{d} \sum_{k=1}^{j-1} |\sigma_j H_{ij}(f) + \mathrm{sign}(\rho_{jk}) \sigma_k H_{ik}(f)|^2 |\rho_{jk}| + \sum_{j=1}^{d} |H_{ij}(f)|^2 \tau_j \sigma_j^2 \tag{10}$$

By $\mathrm{sign}(\rho_{jk})$, the sign of the argument ρ_{jk} is denoted. Thus, the power spectrum of the ith time series can be decomposed the contributions from the correlated part and those from the diagonal part. Finally, the ERPC can be defined as

$$R_{jk \to i}(f) = \left\{ \begin{array}{c} \frac{|H_{ij}|^2 \sigma_j^2 \tau_j}{P_{ii}(f)} \, (j = k) \\ \frac{|\sigma_j H_{ij}(f) + \mathrm{sign}(\rho_{jk}) \sigma_k H_{ik}(f)|^2 |\rho_{jk}|}{P_{ii}(f)} \, (j \neq k) \end{array} \right\}. \tag{11}$$

Thus, the ERPC quantifies the relative contributions from the correlated parts (second line of Eq. (11)) as well as the diagonal parts while the original RPC only takes the diagonal contributions into consideration. It can be seen that when the innovations are mutually uncorrelated (i.e., $\rho_{jk} = 0$), the ERPC is consistent with the original RPC because there is no contribution from the correlated parts (i.e., $R_{jk \to i}(f) = 0$ for $j \neq k$). Note that for each terms in Eq. (11) to be positive value (to be power), $\tau_i = 2 - \sum_{j=1}^{d} |\rho_{ij}| \, (i = 1, \cdots, d)$ must be positive. Therefore, this measure cannot be applied to the highly correlated process.

Extension of relative power contribution to the MARX model

The RPC can also be extended to the MARX model (Eq. (2)). Similar to the derivation of the parametric spectrum of the MAR model, the parametric spectrum of the MARX model can be written as the sum of the spectra resulting from the innovation and an exogenous variable:

$$P_{\mathbf{Z}}(f) = H(f) C_\varepsilon \overline{H}(f)' + H(f) \mathbf{w} P_S(f) \mathbf{w} \overline{H}(f)', \tag{12}$$

where $P_S(f)$ is the power spectrum of the input process S_t.

In the application below, we assume that S_t is the box-car function of the block design and only V1 receives this input. In this case, the vector $\mathbf{w}$ contains an unknown in the corresponding element and zeros in the other elements [that is, $\mathbf{w} = (w_1, 0, 0)'$ if the first element corresponds to V1]. The RPC can be computed as in Eq. (7) by replacing σ_1^2 in Eqs. (6) and (7) with $\sigma_1^2 + w_1^2 P_S(f)$. $P_S(f)$ can be obtained by regarding the periodic input S_t as a quasi-stationary process, which shows a line spectrum (see section 2.3 of Ljung, 1999 for theoretical details). For ease of computation, we apply the sample spectrum defined as:

$$P_S(f) = \frac{1}{T} \left| \sum_{t=1}^{T} S_t e^{\{-i2\pi f t\}} \right|^2. \tag{13}$$

FMRI data acquisition

Five normal subjects (three men and two women) who had no history of neurological or psychiatric illness or developmental disorders took part in this study. The ethics committee of the National Institute for Physiological Sciences, Japan, approved the protocol, and all subjects provided written informed consent. The experiment consisted of viewing alternating a white fixation cross (control) and moving dots (MD). Subjects were instructed to look at a white fixation cross (size 0.3°) at the center of the black screen throughout scan sessions. Similar to Buchel and Friston (1997), two hundred white dots (size 0.1°) were presented on the screen (size 40° horizontally and 30° vertically). These dots were moving radially from the fixation cross toward the border of the screen at a constant speed of 4.0°/s during the MD condition whereas only the fixation point was shown in the control condition. In the experiment, subjects viewed alternating the control condition or MD (30 s each), starting with initial control condition, followed by alternates of MD and the control condition for four times. During the experiment, echo-planer imaging (EPI) images were obtained (TR 1 s, TE 30 ms, FA 62°, FOV 19.2 cm, and 64 × 64 matrices) in 10 oblique slices of 6-mm thickness without a gap, which covered the visual to parietal areas using a 3T scanner (Allegra; Siemens, Erlangen, Germany). In all of the experiments, 10 extra scans were collected at the beginning to establish the steady-state longitudinal magnetization. A three-dimensional structural MRI also was acquired on each subject using a T1-weighted MPRAGE sequence (TR/TE/TI/NEX 1970 ms/4.3 ms/990 ms/1, FA 8°, FOV 210 × 210 mm, and 256 × 256 matrices), yielding 160 sagittal slices with a slice thickness of 1.2 mm and an in-plane resolution of 0.82 mm.

Results

In order to evaluate the strength and direction of the connectivity between three cortical areas associated with the

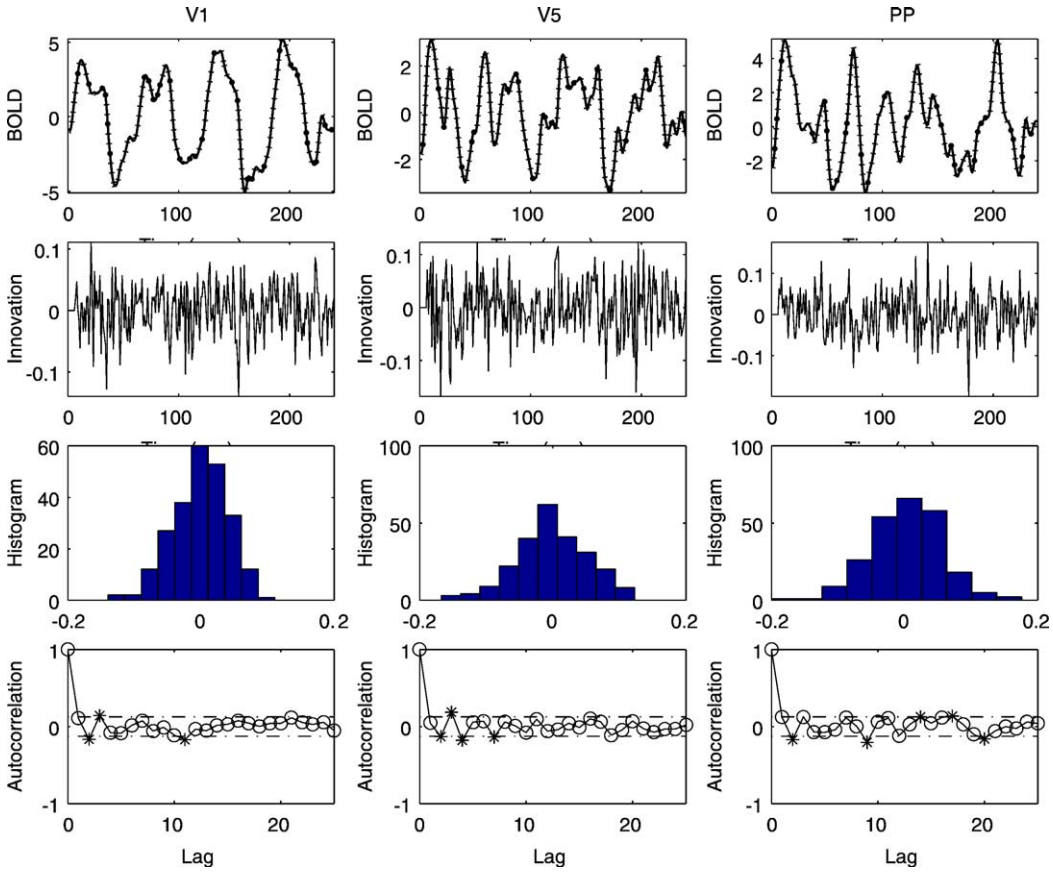

Fig. 2. Time series of the BOLD signal in three ROIs and three kinds of plots of the innovations after MAR-model fitting. In the top row, time series extracted from three ROIs (V1, V5, and PP) are shown. From the second to the bottom row, time series, histograms, and the autocorrelation functions of the innovations (one-step prediction error) after the MAR-model fitting are shown, respectively. In the plots of the autocorrelation function, the approximate 95% confidence limits of the autocorrelation function of a white noise process are also displayed. The sample autocorrelation lying outside the 95% confidence intervals is marked by an asterisk.

perception of visual motion: the primary visual cortex (V1), visual motion-detection area (V5), and the posterior parietal cortex (PP), the abovementioned measures are applied. All of these areas have well-established connections involved in the perception of visual motion (McIntosh and Gonzalez-Lima, 1994).

In the first analysis, the MAR model is fitted to the time series in three ROIs in the right hemisphere, and then the RPC and ERPC are calculated using the estimated parameters. In the second analysis, the MARX model with the box-car input into V1 area is fitted and the ERPC is calculated. The detailed result of these analysis will be shown only for one of the five subjects in Figs. 2–9. The outlined result of all the subjects will be also demonstrated in the case of the MARX model.

MAR model

The MAR model of order $p = 6$, determined by the minimization of AIC, was fitted to the time series shown in the three top panels in Fig. 2. On the remaining panels in Fig. 2, the time series, histogram, and autocorrelation functions of the resulting innovations are shown for the diagnosis of the model fit. From these plots, the distribution of the innovations follows the Gaussian distribution and their lagged correlations are considerably small. Therefore, the model fits the data reasonably well. The estimated correlation of the innovations is 0.3 at the highest. The diagonal assumptions of the innovation covariance matrix was rejected by the likelihood ratio test with P value < 0.001 (Hamilton, 1994, pp. 309–314).

Firstly, the original RPC measure is computed, although the likelihood test shows the diagonal assumption (Eq. (5)) does not hold. In Fig. 3, the power spectrum and the RPC are illustrated in the three top panels and the remaining panels, respectively. Plots of the RPC corresponding to self-contribution (i.e., V1 → V1, etc.) are omitted. The power spectrum of V1 has a clear peak at the frequency $f_0 = 0.018$ Hz around the 60-s period of the block design, whereas the power spectra of V5 and PP have a frequency peak at around $f_1 = 0.033$ Hz around the half period of the block design. These power spectra indicate that the major periodic components captured by the MAR model consist of the frequencies f_0 and f_1. Therefore, the RPC at these frequencies is especially important. The two vertical lines (dotted and real) in the plots of the RPC represent the frequencies f_0, f_1. At f_0, the strength of the bottom-up connections of V1 → V5 and V1 → PP yields values as high as 0.6, whereas at the frequency f_1, the connections of V1 → V5 and V1 → PP vanish, although weak connections between PP and V5 (PP ↔ V5) are observed. We also observed that periodic components of a frequency higher than 0.1 Hz have values close to zero, indicating that there were no fast-component connections. This implies that from a BOLD signal sampled with a low temporal resolution, it is difficult to identify the fast components of

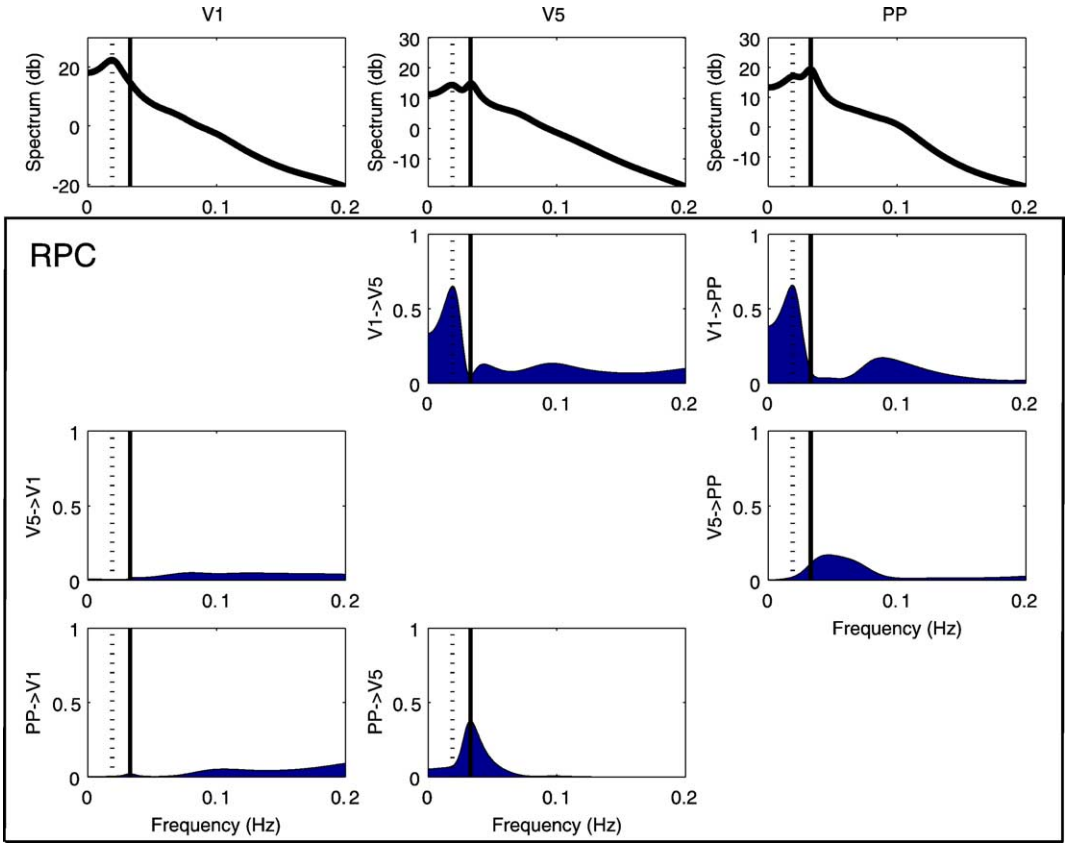

Fig. 3. Power spectra of three ROIs (V1, V5, and PP) and the RPC among those regions. From left to right in the top row, the parametric power spectra of V1, V5, and PP are plotted, respectively. The six panels in the second, third, and fourth rows show the RPC for all of the two-region pairs. The horizontal axis in all panels represents the frequency in hertz.

connectivity (for example, within several seconds), which might be of interest to physiologists.

Secondly, the ERPC is computed in order to know how the innovation correlations affect the results. By the ERPC, each power spectrum is represented by the sum of six terms; the contributions from V1, V5, PP, and the cross-correlated contributions from V1/V5, V1/PP, V5/PP. On the last three rows in Fig. 4, the power contributions of three cross-correlated terms are shown, respectively. It should be noted that in these plots, the contributions from V1/V5 to PP, V1/PP to V5, and V5/PP to V1 are of particular interest for discussing how the ROIs are connecting because the remaining contributions can be considered as self-contributions as in plots omitted in the top three rows. The first thing we notice is that the shape of the RPC curves corresponding to the uncorrelated portion are almost same as that in Fig. 3 (i.e., the original RPC curves when the correlation is neglected), although the RPC values decrease. This result is natural because even if we apply a MAR model with the diagonal constraint onto the innovation covariance matrix, the estimate of MAR coefficients does not change. Next, we observe that the cross-correlated portions of the innovations V1/V5 and V1/PP have some contributions to PP and V5 around the frequency f_0, respectively, whereas the power spectrum of any of these three time series is not contributed by the cross-correlated portions between the remaining two time series (for example, V5/PP has little contribution on V1 and so on). However, it is not easy to discuss a physiological interpretation of these cross-correlated contributions unless we identify the cause of the correlations. In Fig. 5, the results are schematically demonstrated.

MARX model

Physiologically, it is more natural to take into consideration a MAR model with the input to V1 that is associated with the experimental design. This can be done by a MARX model (Eq. (2)), where S_t is the box-car function with a value of 0.5 in the task and -0.5 in the control conditions, and $\mathbf{w}$ has an unknown in the first element ($\mathbf{w} = (w_1,0,0)^t$) (see Fig. 6). Based on this model, we can expect to see how the box-car function influences the three ROIs directly or indirectly through the estimated parameters.

The MARX model of the order $p = 6$, determined by the minimization of the AIC, was fitted to the time series shown in the three top panels in Fig. 7 (the same data as in Fig. 2). In the remaining panels in Fig. 7, the time series, histogram, and autocorrelation functions of the resulting innovations are shown to diagnose the fit of the model. As was the case with the MAR model, the fit of the model to the data is reasonable and the estimated correlation of the innovations is 0.2 at its highest (the likelihood ratio test still rejects the null hypothesis of uncorrelated innovaitons). Furthermore, the fit via the MARX model improves 20 in the AIC value compared with the fit via the MAR model, and therefore this result is statistically more likely. Note, for example, the innovations for V1 are slightly smaller for MARX model than for the MAR model.

In Fig. 8, the parametric spectra and the ERPC computed from the estimated parameters are illustrated in the three top panels and the remaining panels, respectively. Firstly, we can see that the parametric spectra are not as smooth as those obtained by the MAR

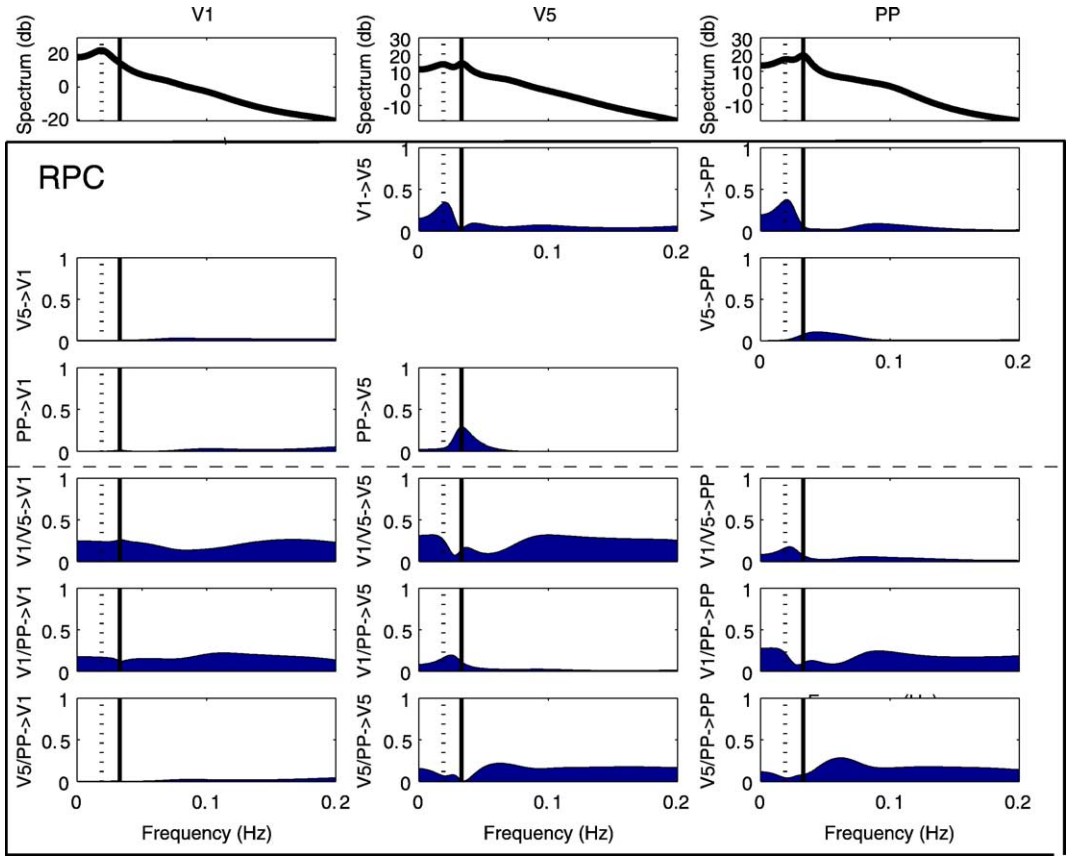

Fig. 4. Power spectra of three ROIs (V1, V5, and PP) and the ERPC among those regions. From the left to right in the top row, the parametric power spectra of V1, V5, and PP are plotted, respectively. The six panels in the second, third, and fourth rows show the relative contributions of the uncorrelated portions and the remaining panels show the relative contributions of the cross-correlated portions to V1,V5, and PP, respectively, from the first to the third column. The horizontal axis in all panels represents the frequency in hertz.

model. This is because these spectra comprise a compound of the smooth spectrum resulting from the innovations and the line spectrum resulting from the deterministic periodic box-car function, respectively. The power spectrum of V1 shows a strong effect of the box-car function compared with the spectra of the other two regions. It should be noted that, although the stimulus is periodic, it will have power at multiple frequencies as it is box-car rather than sinusoidal in shape. It is also important to note that the spectra in Figs. 4 and 8 are similar if the line spectrum does not exist.

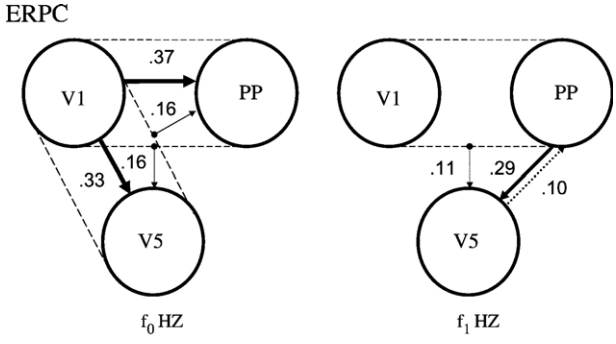

Fig. 5. Result of the ERPC analysis with the MAR model. The specified connections at the frequencies f_0, f_1 are shown by arrows. Only connections more than 0.1 are shown and the self-contributions are neglected.

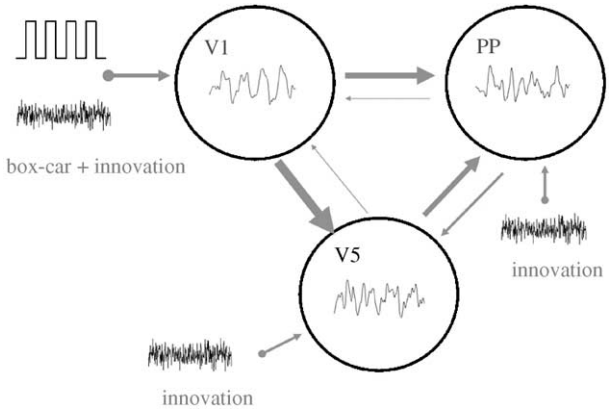

Fig. 6. Schematic figure of the MARX model and the RPC analysis. By the MARX model, the box-car of the experimental design can be modeled as the input into V1 as well as intrinsic white noise processes as in the MAR model. The RPC quantifies the strength of every interregional connection by computing how much the input in each region (that is, the innovations and the box-car function) drives the time series of other regions. The resulting strength of the connections is schematically represented by the thickness of the arrows. These arrows can be obtained for all frequency components. V1: the primary visual cortex for initial visual information processing. V5: the area for visual motion detection. PP: the posterior parietal cortex for the integration of visual spatial information.

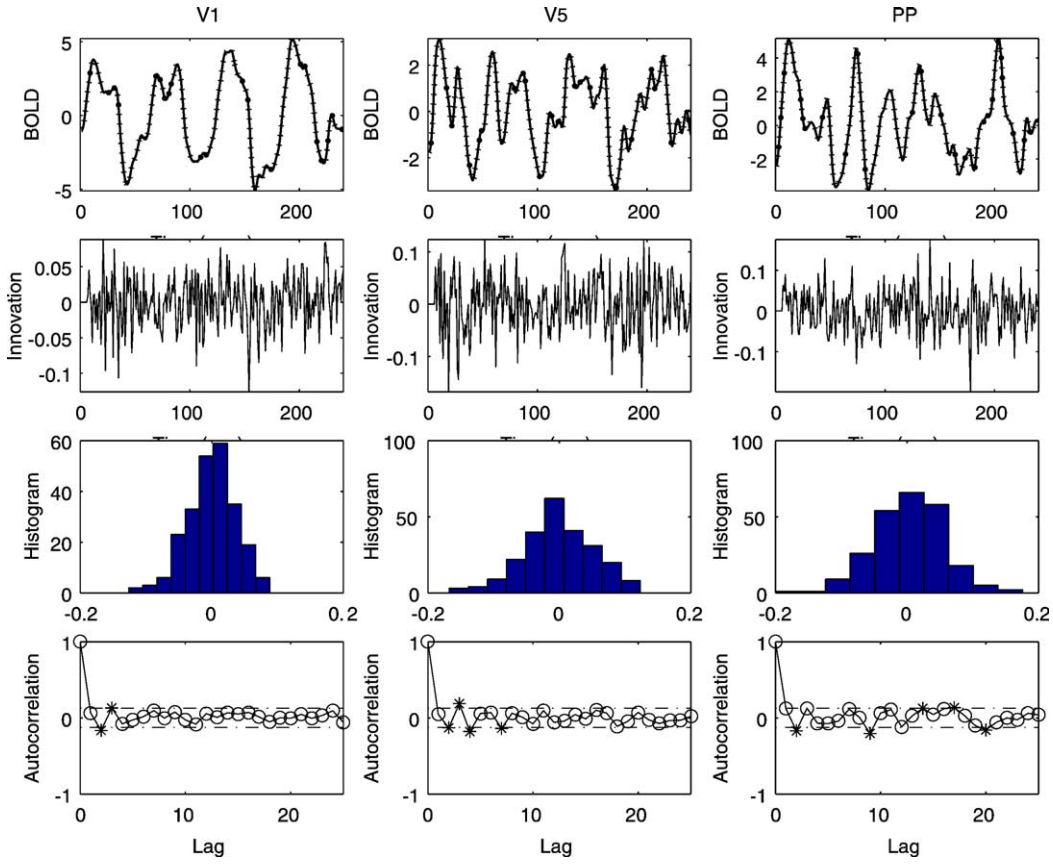

Fig. 7. Time series of the BOLD signal in three ROIs and three kinds of plots of the innovations after the MARX-model fitting. In the top row, the time series extracted from three ROIs (V1, V5, and PP) are shown. From the second to the bottom row, time series, histograms, and the autocorrelation functions of the innovations (one-step prediction error) after the MARX-model fitting are shown, respectively. The dotted line and asterisks in the plots of the autocorrelation function are used as in Fig. 3.

The power contributions of the major periodic component f_0 = 0.017 Hz shows the strong bottom-up connections of V1 → V5 and V1 → PP. The weak connection PP → V5 is also observed around the frequency f_1 = 0.033 Hz (see Fig. 9). The contributions of the cross-correlated portions look similar to those in Fig. 4, although the effect of the correlations relatively decreases around some frequencies due to the introduction of the exogenous variable into the model. Interestingly, we observed that the line spectrum had a strengthening effect on V1 → V5 and V1 → PP connectivity, whereas its effect on the remaining connections was to weaken connectivity (see notches in lower RPC plots).

Figs. 10 and 11 summarize the results of the ERPC at $f_0$and f_1 from the five subjects. The RPC values (denoted by *) and their 95% confidence intervals obtained computationally by the bootstrap method (see Appendix A) are shown as to the contributions of the uncorrelated portions of the innovation. In Fig. 10, which shows the 60 s/cycle frequency component, only the RPC of V1 → V5 and V1 → PP attain significantly high values, and this tendency is consistent across the subjects. In Fig. 11, which shows the 30 s/cycle frequency component, the RPCs of V1 → V5 and V5 → PP are not as strong as in the 60 s/cycle component, but the weak top-down connections PP → V5 can be found for four of the five subjects and the modestly strong connections PP → V1 is found for the remaining subject. The 30 s/cycle component may have some relation with the timing of the change from the task to the control condition or vice versa because two experimental conditions alter in every 30 s.

Discussion

In this article, we used the RPC measure developed by Akaike and its extended version in order to quantify the strength of directed interregional cortical connections. This method can be considered as an extension of Harrison's method in the sense that the post-processing after the identification with the MAR model is more complex, but it provides much more detailed information about the directed and frequency-wise connectivity contained in the model. Using this approach, we can determine whether the specified connections represent low-frequency or high-frequency connectivity, which cannot be examined by the estimated MAR coefficients themselves. Furthermore, by extending the RPC to the MARX model, we can construct models that are more physiologically feasible and see the influence of exogenous input variables. It should be noted that for this model, unlike the structural equation model, no assumption about the direction of connections is necessary, because the directed interregional relationship can be determined from the temporal order of a set of time series (that is, Granger causality). We shall discuss the methodological issues, the results of the analysis, and the problems in the future.

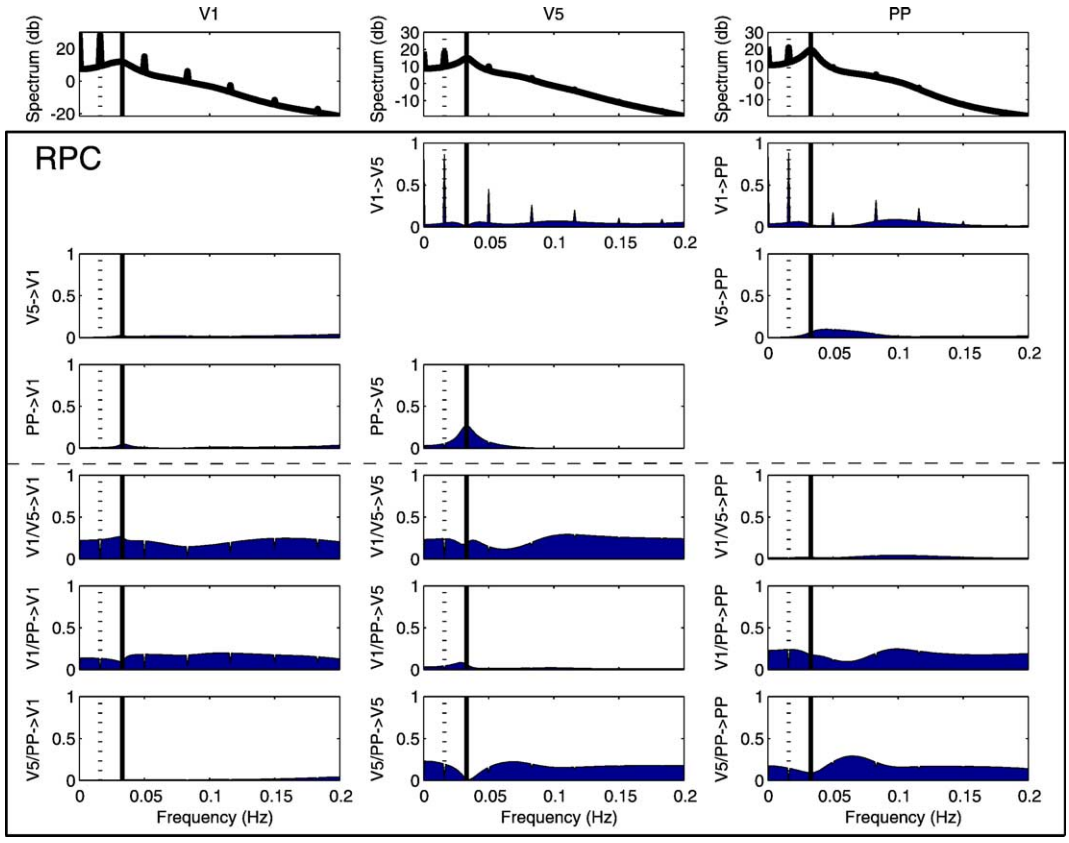

Fig. 8. Power spectra of three ROIs (V1, V5, and PP) and the RPC between those regions. From left to right in the top row, the parametric power spectra of V1, V5, and PP are plotted, respectively. The six panels in the second, third, and the fourth row show the RPC for all of the two-region pairs. The horizontal axis of all the panels represents the frequency in hertz.

Methodological issues

The MAR model is broadly employed as a model of the linear feedback system in real applications. In the MAR model, the white noise processes (innovations) driving the observed time series are essential because if without innovations, the MAR model only can represent simple behaviors. Thus, the innovations can be considered as the source of the information producing the time series. It is a main idea of the RPC that the power spectrum of each observed time series is consisting of the power contributions of each innovation and that the directed interactions are measured by the ratios of these innovation contributions (note that every time series has its own innovation). Following this idea, the RPC is naturally extended to the case of the MARX model by considering the exogenous variables as another inputs as well as the case of the correlated innovations by decomposing the correlated innovations into their own uncorrelated portions and the inseparable correlated portions (i.e., extended RPC). Furthermore, the time-domain RPC can be defined, which is equivalent to the integration of the frequency-domain RPC over all the frequencies (Yamashita, 2004).

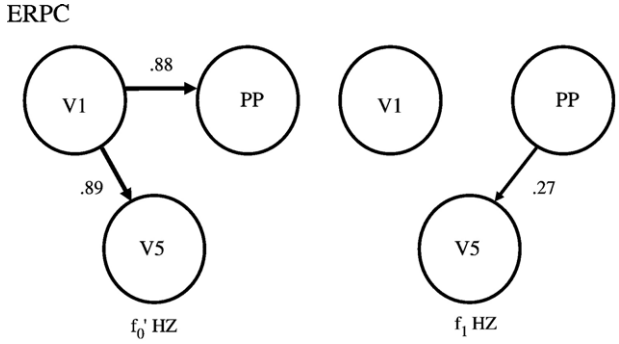

Fig. 9. Result of the ERPC analysis with the MARX model. The specified connections at the frequencies f_0, f_1 are shown by arrows. Only connections more than 0.1 are shown and the self-contributions are neglected.

There are some directed measures based on the MAR model closely related to the RPC. The (generalized) directed coherence (DC) is consistent with square root of the RPC in order to develop the directed version of the coherence (see Baccala and Sameshima, 2001 for the review). The directed transfer function (DTF) (Kaminski et al., 2001) can be obtained as a special case of the RPC when the variance of all innovations is normalized to 1. The measure proposed by Geweke has been also widely employed in the literature on the connectivity analysis to biological data such as local field potentials, EEG, MEG, and fMRI (Bernasconi and Konig, 1999; Bernasconi et al., 2000; Brovelli et al., 2004; etc.). His measure in the frequency domain has the form similar to the RPC. Actually, in the case of the bivariate MAR process with uncorrelated innovations, these two measures are linked by the exponential transformation. However, the RPC and Geweke's measure in frequency domain is generally different because the latter overcomes

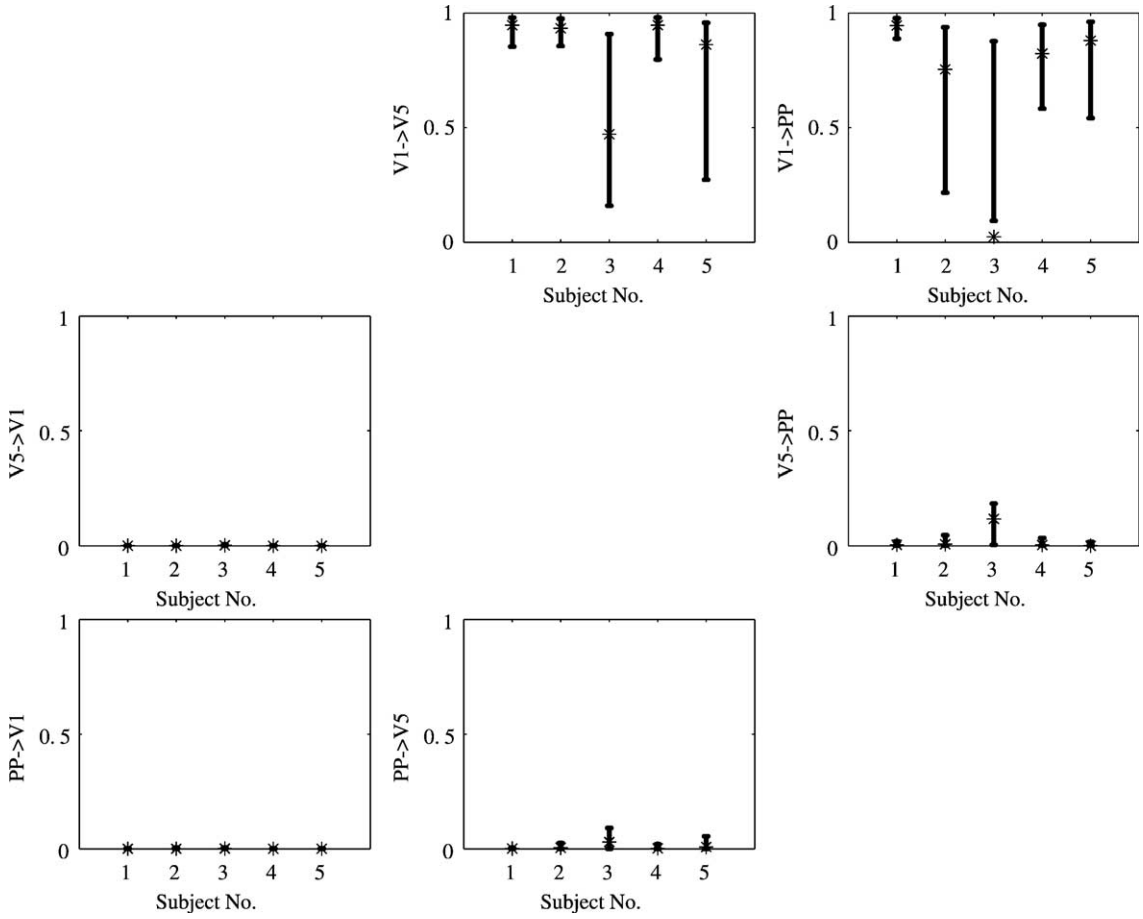

Fig. 10. Summarized results of the RPC analysis for the 60 s/period component for all the five subjects. On each panel, the RPC values and their 95% confidence intervals for the five subjects are represented by asterisks and error bars. The error bars are computed using 250 bootstrap samples.

difficulty in the correlated innovations by applying the orthogonal transformation to the innovation covariance matrix which changes the coordinate of the innovations.

One of advantages of the RPC (and DC, DTF) is low computational cost. The RPC only requires the parameter estimation of the MAR model consisting of all the variables to quantify the strength of the interactions among every pair of the time series with the parameters whereas the number of the parameter estimation needed to compute Geweke's measure is rapidly increasing as the number of time series increases. In addition, the RPC can be extended to the MARX model. This extension could be attractive because it is natural in some experiments to consider that lower brain regions have some direct inputs corresponding to the experimental stimulus. On the other hand, the RPC has some limitations: even the ERPC cannot be applied when the innovations are highly correlated; the RPC measures the total effects of all the paths between two regions, in other words, the influence of the indirect interactions is included (this issue is discussed in the Appendix of Kaminski et al., 2001 and in Baccala and Sameshima, 2001). The partial directed coherence (PDC) or Geweke's measure in time domain could be a good candidate to evaluate the direct effect between two regions. Although due to the second limitations the RPC may not be appropriate to evaluate the amount of direct interactions between two regions, the RPC still can provide some clues on connectivity among some distributed brain regions as shown in this analysis.

It would be helpful to clarify the role of a time-series model in connectivity analysis. In a parametric time-series model, generally, the information related to temporal order is captured by the coefficients of lagged variables ('dynamical part'), and the information related to the instantaneous effect, which might be an intrinsic correlation or caused by the lack of temporal resolution, is captured by the innovation–correlation structure. The former is essential to specify the directed relationship based on the idea of Granger causality, whereas the latter provides no clue as to the direction of the relationship. Thus, the time-series model works most effectively only when the data can be explained primarily by their dynamical part.

Results of the analysis

In this article, we specified frequency-wise directed connectivity by employing RPC and its extension ERPC. As a result of applying the MAR model to the data from the random dot experiment, we can see the strong connectivity of V1 → V5 and V5 → PP for the periodic 60 s/cycle component, and the weak connectivity of PP ↔ V5 for the 30 s/cycle component. This weak connectivity might be related to the change of the sensory input from the control condition to MD, or vice versa because 30 s corresponds to the alternation of the experimental conditions. Some contributions of the correlations are quantified by the ERPC measure, but it is not easy to provide interpretation as long as we do not know what is cause of these

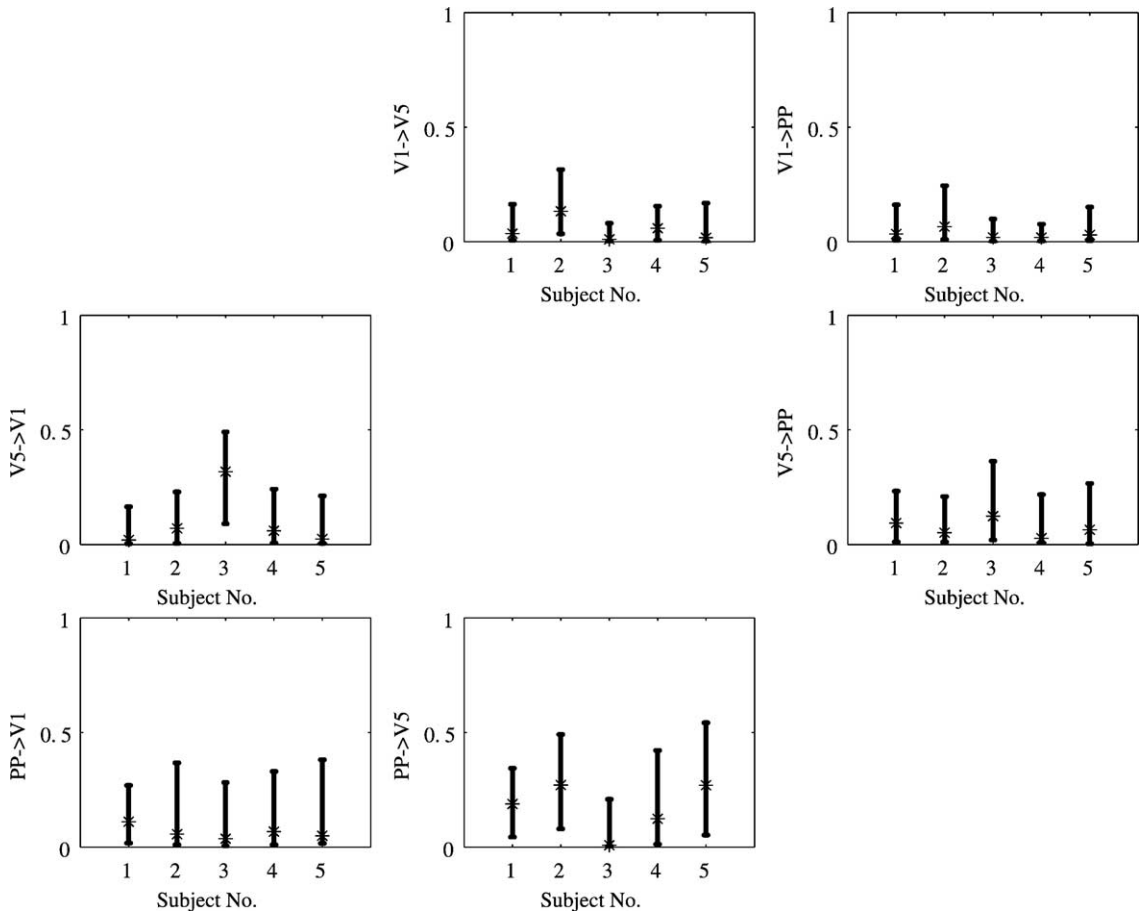

Fig. 11. Summarized results of the RPC analysis for the 30 s/period component for all the five subjects. On each panel, the RPC values and their 95% confidence intervals for the five subjects are represented by asterisks and error bars. The error bars are computed using 250 bootstrap samples.

correlations. The result from the analysis employing the ERPC with the MARX model, which is statistically more reliable, emphasizes the effect of the 60 s/cycle periodic component by introducing the box-car function as the input to V1. Furthermore, the effects of correlations are decreasing compared with those of the MAR model. Interestingly, the results obtained from the ERPC with the MARX model are similar across the five subjects. With regard to fast components of more than 0.1 Hz, we cannot find the information crucial to the establishment of connectivity. This might be a intrinsic limitation of the connectivity analysis of BOLD signals using the MAR(X) model because of the smoothing effect of hemodynamics response, or the low temporal resolution of fMRI data, or the averaging effect in choosing the representative time series of the ROIs as the first principle component (the effect of hemodynamics response and down-sampling to the causal inference of rapid neuronal signals is discussed in the simulation study of Goebel et al., 2003).

In conclusion, there exists some directed connectivity of the BOLD signal around the frequencies related to the block experimental design but difficult to find out the connectivity of fast components (higher than 0.1 Hz), which could be more closely related to neuronal connections. Since the components of connectivity specified in this analysis correspond to the period of the experimental design, this connectivity may be referred to as 'functionally directed connectivity'. Without any assumptions on the mechanics between neural signals and BOLD signals, it seems to be impossible to retrieve the connectivity directly related to neuronal phenomena from BOLD signals. One possible approach is to impose some assumptions on the structure of connectivity and the mechanics between neural signals and BOLD signals as the DCM does. The alternative could be to operationally find out some directed connectivity as in this paper, although the specified connectivity may not be directly related to effective connections of neuronal signals. The second kind of the analysis could be still of importance as it takes a step beyond localization to integration. For example, it might be interesting to compare the distributed (functional) network between two different groups.

In the multivariate time-series approach, the direction of connectivity is entirely determined by the idea of Granger causality; therefore, it is possible to misspecify connectivity that does not exist in the underlying anatomy. If we have anatomical knowledge of connectivity, this prior knowledge preferably would be incorporated into the analysis. The multivariate time-series approach combined with Bayesian estimation (introduced by Harrison et al., 2003) is more promising, although we used the maximum-likelihood estimate method in this paper.

In this analysis, we specified low-frequency connectivity of BOLD signals, fully taking advantage of the block experimental design. It could be one of future works to confirm validity of the frequency measures by applying them to the data of more complicated experimental designs such as event-related design

and randomized-block design. The statistical method to conduct group analysis of connectivity is also desired.

Acknowledgments

This research was supported by the Japan Society for the Promotion of Science (JSPS), grant number 13654075. This research was supported in part by the National Institute of Information and Communications Technology (NICI). The authors would like to thank Daisuke Saito at the National Institute for Physiological Sciences, Japan, for his cooperation in conducting this experiment. The authors also appreciate the referees for their helpful comments.

Appendix A. Maximum likelihood estimation

Let us define Θ as the collection of MAR coefficients as in Eq. (16). The log likelihood is represented as:

$$L(\Theta, C_\varepsilon) = \log P(\mathbf{Z}_1, \cdots, \mathbf{Z}_T; \Theta, C_\varepsilon)$$

$$= \sum_{t=p+1}^{T} \log P\left(\mathbf{Z}_t \mid \mathbf{Z}_{t-1}, \cdots, \mathbf{Z}_{t-p}; \Theta, C_\varepsilon\right) + \log P\left(\mathbf{Z}_1, \cdots, \mathbf{Z}_p; \Theta, C_\varepsilon\right) \quad (14)$$

Under the assumption of Gaussian innovation, the conditional distribution in Eq. (14) is given by:

$$P\left(\mathbf{Z}_t \mid \mathbf{Z}_{t-1}, \cdots, Z_{t-p}; \Theta, C_\varepsilon\right) = N(\bar{\mathbf{Z}}_t, C_\varepsilon) \quad (15)$$

where

$$\bar{\mathbf{Z}}_t = \sum_{i=1}^{p} A(i) Z_{t-i} \equiv \Theta' X_t$$

$$X_t \equiv \left(\mathbf{Z}'_{t-1}, \cdots, \mathbf{Z}'_{t-1}\right)'$$

$$\Theta' = (A(1), \cdots, A(p)). \quad (16)$$

The log-likelihood (Eq. (14)) can be rewritten in the form,

$$L(\Theta, C_\varepsilon) = C + \frac{T-p}{2} \log |C_\varepsilon^{-1}| - \frac{1}{2} \sum_{t=p+1}^{T} \left(\mathbf{Z}_t - \Theta' X_t\right)' C_\varepsilon^{-1} \left(\mathbf{Z}_t - \Theta' X_t\right), \quad (17)$$

where C is a constant term and the term $\log P(\mathbf{Z}_1, \cdots, \mathbf{Z}_p; \Theta, C_\varepsilon)$ has been neglected (when T is large, this term is small compared with the second and third terms in Eq. (17)). The maximum likelihood (ML) estimator can be obtained as the maximizer of $L(\Theta, C_\varepsilon)$, which results in

$$\hat{\Theta}' = \left[\sum_{t=p+1}^{T} \mathbf{Z}_t X'_t\right] \left[\sum_{t=p+1}^{T} X_t X'_t\right]^{-1} \quad (18)$$

$$\hat{C}_\varepsilon = \frac{1}{T-p} \sum_{t=p+1}^{T} \hat{\varepsilon}_t \hat{\varepsilon}'_t \quad (19)$$

where the innovation estimate is given by

$$\hat{\varepsilon}_t = \mathbf{Z}_t - \hat{\Theta}' X_t. \quad (20)$$

The jth row of $\hat{\Theta}'$ is

$$\hat{\theta}'_j = \left[\sum_{t=p+1}^{T} z_{j,t} X'_t\right] \left[\sum_{t=p+1}^{T} X_t X'_t\right]^{-1} \quad (21)$$

where $\theta_j = (A_{j1}(1), \cdots, A_{jd}(1), \cdots, A_{j1}(p), \cdots, A_{jd}(p))'$ is a vector of size $pd \times 1$. The variance of $\hat{\theta}'_j$ is given by

$$\mathrm{Var}\left(\hat{\theta}'_j\right) = \hat{\sigma}_j^2 \left[\sum_{t=p+1}^{T} X_t X'_t\right]^{-1} \quad (22)$$

where $\hat{\sigma}_j^2$ is the jth diagonal element of $\hat{C}_\varepsilon$. Note that the ML estimators of the coefficients (Eq. (14)) are equivalent to the ordinary least-squares estimators (for further details, see chapter 11 of Hamilton, 1994).

Selection of p

The parameters (p) are preferably determined by information criteria, such as AIC (Akaike, 1973), SIC (Schwarz, 1978) or AICc (Hurvich and Tsai, 1989). In this study, we chose to employ AIC because of its simple form and the wide use of these criteria in various applications. AIC is defined as follows

$$\mathrm{AIC} = -2L\left(\hat{\Theta}, \hat{C}_\varepsilon\right) + 2N_p \quad (23)$$

where N_p is the number of parameters in the model for fitting. AIC represents the asymptotically unbiased estimates of the Kullback–Leibler (KL) discrepancy between the true model and the fitted model. Therefore, the minimization of these criteria corresponds to choosing the nearest model to the true model in the sense of KL discrepancy. The AR order p is determined by calculating AIC for each p in some range and by choosing the minimizers.

Bootstrap method

The bootstrap method is one of the resampling techniques used on the data. Intuitively speaking, by resampling randomly from the histogram of the original data, a new sample can be obtained that shares the same probability distribution as the original data. This new sample can be used to calculate a new realization of some quantity.

As the analytical distribution of the values of the RPC is difficult to obtain, the parametric bootstrap method (Carstein, 1992) is applied to the MARX model as follows:

1. Obtain the parameter estimates $\hat{A}(i)$,$\hat{\mathbf{w}}$, $\hat{\mathbf{C}}_\varepsilon$ and the innovation estimates $\hat{\varepsilon}_{p+1}, \cdots, \hat{\varepsilon}_T$ from the original data $\mathbf{Z}_1, \cdots, \mathbf{Z}_T$.
2. Obtain a new sample of the innovation process $\varepsilon^*_{p+1}, \cdots, \varepsilon^*_T$ from the histogram of $\{\hat{\varepsilon}_{p+1}, \cdots, \hat{\varepsilon}_T\}$.

3. Generate a new sample of time series $\mathbf{Z}^*_1, \cdots, \mathbf{Z}^*_T$ as follows:

$$\mathbf{Z}^*_t = \mathbf{Z}_t \quad (t = 1, \cdots, p)$$

$$\mathbf{Z}^*_t = \sum_{k=1}^{p} \hat{A}(i)\mathbf{Z}^*_{t-k} + \hat{\mathbf{w}}S_t + \varepsilon^*_t \quad (t = p+1, \cdots, T)$$

4. Obtain new parameter estimates from the new sample and compute the RPC values.

These steps are iterated until a sufficiently large number of samples can be obtained. In our application, 250 samples were generated to construct the confidence interval.

References

Akaike, H., 1968. On the use of a linear model for the identification of feedback systems. Ann. Inst. Stat. Math. 20, 425–439.

Akaike, H., 1973. Information theory and an extension of maximum likelihood principle. 2nd International Symposium on Information Theory. Akadimiai Kiadi, Budapest, pp. 267–281.

Baccala, L.A., Sameshima, K., 2001. Partial directed coherence: a new concept in neural structure determination. Biol. Cybern. 84, 463–474.

Bernasconi, C., Konig, P., 1999. On the directionality of cortical interactions studied by structural analysis of electrophysiological recordings. Biol. Cybern. 81, 199–210.

Bernasconi, C., von Stein, A., Chiang, C., Koenig, P., 2000. Bi-directional interactions between visual areas in the awake behaving cat. Neuro-Report 11, 689–692.

Brovelli, A., Ding, M., Ledberg, A., Chen, Y., Nakamura, R., Bressler, S.L., 2004. Beta oscillations in a large-scale sensorimotor cortical network: Directional influences revealed by Granger causality. Proc. Natl. Acad. Sci. U. S. A. 101, 9849–9854.

Buchel, C., Friston, K.J., 1997. Modulation of connectivity in visual pathways by attention: cortical interactions evaluated with structural equation modelling and fMRI. Cereb. Cortex 7, 768–778.

Carstein, E., 1992. Resampling techniques for stationary time-series: some recent developments. In: Brillinger, D., Caines, P., Geweke, J., Parzen, E., Rosenblatt, M., Taqqu, M.S. (Eds.), New Directions in Time Series Analysis Part I. Springer-Verlag, New York.

Friston, K.J., 1994. Functional and effective connectivity in neuroimaging: a synthesis. Hum. Brain Mapp. 2, 56–78.

Friston, K.J., Holmes, A.P., Worsley, K.P., Poline, J.B., Frith, C.D., Frackowiak, R.S.J., 1995. Statistical parametric maps in functional imaging: a general linear approach. Hum. Brain Mapp. 2, 189–210.

Friston, K.J., Harrison, L., Penny, W., 2003. Dynamic causal modelling. NeuroImage 19, 1273–1302.

Gavrilescu, M., Stuart, G.W., Waites, A., Jackson, G., Svalbe, I.D., Egan, G.F., 2004. Changes in effective connectivity models in the presence of task-correlated motion: an fMRI study. Hum. Brain Mapp. 21, 49–63.

Geweke, J.F., 1982. Measurement of linear dependence and feedback between multiple time series. J. Am. Stat. Assoc. 77, 304–324.

Geweke, J.F., 1984. Measures of conditional linear dependence and feedback between time series. J. Am. Stat. Assoc. 79, 907–915.

Goebel, R., Roebroeck, A., Kim, D.S., Formisano, E., 2003. Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and Granger causality mapping. Magn. Reson. Imaging 21, 1251–1261.

Goncalves, M.S., Hall, D.A., 2003. Connectivity analysis with structural equation modelling: an example of the effects of voxel selection. NeuroImage 20, 1455–1467.

Granger, C.W.J., 1969. Investigating causal relations by econometric models and cross-spectral methods. Econometrica 37, 424–438.

Hamilton, J.D., 1994. Time Series Analysis. Princeton Univ. Press, New Jersey.

Harrison, L., Penny, W.D., Friston, K., 2003. Multivariate autoregressive modeling of fMRI time series. NeuroImage 19, 1477–1491.

Hosoya, Y., 1991. The decomposition and measurement of the interdependency between second-order stationary process. Probab. Theory Relat. Fields 88, 429–444.

Hosoya, Y., 2001. Elimination of third-series effect and defining partial measures of causality. J. Time Ser. Anal. 22 (5), 537–554.

Hurvich, C., Tsai, C., 1989. Regression and time series model selection in small samples. Biometrika 76, 297–307.

Kaminski, M., Ding, M., Truccolo, W.A., Bressler, S.T., 2001. Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance. Biol. Cybern. 85, 145–157.

Lahaye, P.J., Poline, J.B., Flandin, G., Dodel, S., Garnero, L., 2003. Functional connectivity: studying nonlinear, delayed interactions between BOLD signals. NeuroImage 20, 962–974.

Ljung, L., 1999. System Identification-Theory for the User. Second ed. Prentice Hall.

Lee, L., Harisson, L.M., Mechelli, A., 2003. The functional brain connectivity workshop: report and commentary. Netw.: Comput. Neural Syst. 14, R1–R15.

McIntosh, A.R., Gonzalez-Lima, F., 1994. Structural equation modeling and its application to network analysis in functional brain imaging. Hum. Brain Mapp. 2, 2–22.

Saito, Y., Harashima, H., 1981. Tracking of information within multichannel EEG record—causal analysis in EEG. In: Yamaguchi, N., Fujisawa, K. (Eds.), Recent Advances in EEG and MEG Data Processing. Elsevier, Amsterdam, pp. 133–146.

Schwarz, G., 1978. Estimating the dimension of a model. Ann. Stat. 6, 461–464.

Shumway, R.H., 2000. Time Series Analysis and Its Applications. Springer, New York.

Tanokura, Y., Kitagawa, G., 2004. Power contribution analysis for multivariate time series with correlated noise sources. Adv. Appl. Stat. 4 (1), 65–95.

Worsley, K.J., Liao, C.H., Aston, J., Petre, V., Duncan, G.H., Morales, F., Evans, A.C., 2002. A general statistical analysis for fMRI data. NeuroImage 15, 1–15.

Yamashita, O., 2004. Dynamical EEG inverse problem and causality analysis of fMRI data. PhD. Thesis.