NeuroImage 58 (2011) 323–329

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Comments and Controversies

Wiener–Granger Causality: A well established methodology

Steven L. Bressler [a,*], Anil K. Seth [b]

[a] Center for Complex Systems and Brain Sciences, Florida Atlantic University, Boca Raton, FL, USA
[b] Sackler Centre for Consciousness Science and School of Informatics, University of Sussex, Brighton, BN1 9QJ, UK

ARTICLE INFO

Article history:
Received 18 December 2009
Revised 10 February 2010
Accepted 13 February 2010
Available online 2 March 2010

Keywords:
Autoregressive model
Brain
Causality
Inference
Neuroscience
Time series

ABSTRACT

For decades, the main ways to study the effect of one part of the nervous system upon another have been either to stimulate or lesion the first part and investigate the outcome in the second. This article describes a fundamentally different approach to identifying causal connectivity in neuroscience: a focus on the predictability of ongoing activity in one part from that in another. This approach was made possible by a new method that comes from the pioneering work of Wiener (1956) and Granger (1969). The Wiener–Granger method, unlike stimulation and ablation, does not require direct intervention in the nervous system. Rather, it relies on the estimation of causal statistical influences between simultaneously recorded neural time series data, either in the absence of identifiable behavioral events or in the context of task performance. Causality in the Wiener–Granger sense is based on the statistical predictability of one time series that derives from knowledge of one or more others. This article defines Wiener–Granger Causality, discusses its merits and limitations in neuroscience, and outlines recent developments in its implementation.

Introduction

For most of its history, neuroscience has primarily been concerned with examining the physiological correlates of experimentally delivered stimuli and overt behavioral responses. More recently, there has been growing interest in studying the effect that one part of the nervous system has on another, either in the absence of identifiable behavioral events or in the context of task performance. Such effects are typically examined by stimulating or lesioning the first part and investigating the outcome in the second. In peripheral and spinal pathways, the interventional techniques of stimulation and ablation have proven to be powerful methods for inferring causal influences from one neuron or neuronal population to another. For the study of causal relations within the brain, interventional techniques also have utility, although that utility is diminished by the high levels of convergence and divergence in brain pathways, as well as the highly nested reciprocity of projections. This article deals with a different approach to the problem of causal influence in the brain, called Time Series Inference (TSI). This approach, although relatively recent in neuroscience, is showing promise as a valuable adjunct to more traditional interventional approaches.

TSI, unlike stimulation or ablation, does not require intervention in the nervous system. It is based on temporal relations existing between time series recordings of neural activity, which may be obtained noninvasively as electroencephalographic (EEG), magnetoencephalographic (MEG), or functional Magnetic Resonance (fMRI) data. Of course, the time series may also be obtained from invasive single-unit, multi-unit, local field potential, or electrocorticographic recordings. TSI depends on the statistical predictability of one time series by another time series. If the two time series represent neural activity from different neurons or neuronal populations, then inference about causal relations between those neurons or neuronal populations is possible. Another advantage of TSI methods is that they naturally accommodate stochastic processes, and thus are well suited to the ubiquitous variability that is found in neural time series data. Furthermore, under proper conditions, TSI methods can be effectively used to relate neural activity to cognitive function.

In this article, we will describe Wiener–Granger Causality (WGC) as the type of TSI most commonly employed in neuroscientific studies. TSI methods, in one form or another, are well established and widely used in many different fields of study. The use of TSI methods in neuroscience is relatively new, but already has appeared in different implementations. For example, the popular Dynamic Causal Modeling (DCM) (Friston et al., 2003) approach is based on TSI, although this fact is not commonly appreciated. We will focus here on WGC techniques, which, as described below, are based on a relatively small set of straightforward assumptions. DCM, which we will not discuss at length, involves TSI with additional and more complicated assumptions (see accompanying articles in this issue).

In summary, our purpose here is to detail the implementation of WGC via AutoRegressive (AR) modelling, discuss the assumptions required to apply WGC to neural time series data, describe some of the

* Corresponding author.
E-mail addresses: bressler@fau.edu (S.L. Bressler), a.k.seth@sussex.ac.uk (A.K. Seth).
URLs: http://www.ccs.fau.edu/~bressler (S.L. Bressler), http://www.anilseth.com (A.K. Seth).

1053-8119/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2010.02.059

limitations of its use, and provide some examples of applications to neural time series data. Finally, we will describe some potentially important extensions of currently employed techniques.

Wiener–Granger Causality

Time-domain WGC

The problem of defining 'causality' is non-trivial for complex systems, where, unlike simple systems observed in the every-day world, an intuitive understanding of cause and effect is lacking. In 1956 Norbert Wiener introduced the notion that one variable (or time series) could be called 'causal' to another if the ability to predict the second variable is improved by incorporating information about the first (Wiener, 1956). Wiener however lacked a practical implementation of his idea. Such an implementation was introduced in 1969 by the econometrician Clive Granger (1969) in the context of linear autoregressive models of stochastic processes.

The basic idea of Wiener–Granger Causality (WGC), or G-causality, is straightforward. Suppose that we have two variables **X**, and **Y**, and we try to predict $\mathbf{X}_{t+1}$ using only past terms of **X**. We also try to predict $\mathbf{X}_{t+1}$ using past terms of both **X** and **Y**. If the second prediction is significantly more successful, then the past of **Y** contains information useful for predicting $\mathbf{X}_{t+1}$ that is not in the past of **X**. In this case, **Y** is said to *G-cause* **X**. Note that these variables can be scalar time series, for which we use normal type (e.g., X) or vector time series, for which we use bold type (e.g., **X**).

Because the values of a variable at one time are predicted by values of other variables at earlier times, it is often said that WGC depends on 'temporal precedence'. However, it is not sufficient simply that events in the other variables temporally precede similar events in the first variable. For WGC to be significant, statistically significant predictability must be established. In other words, non-zero values for WGC can usually be obtained from any set of time series, but these values are meaningless unless it is determined that they are statistically significant.

Assuming that variables **X** and **Y** are stochastic and Wide-Sense Stationary (WSS, meaning that they have constant means and variances), WGC can be easily implemented using linear vector AR models,[1] though the concept is not tied to this implementation. Let us first define $\mathbf{X}_{t-1}^{(m)} = \mathbf{X}_{t-1} \oplus \ldots \oplus \mathbf{X}_{t-m}$, where $\mathbf{X}_{t-i}$ is a column vector consisting of the values of elements of **X** at time $t-i$, m is a fixed number of lagged observations, and the symbol $\oplus$ denotes concatenation of column vectors. Now consider the pair of (restricted and unrestricted) regression models:

$$\begin{aligned} \mathbf{X}_t &= A \cdot \mathbf{X}_{t-1}^{(m)} + \varepsilon_t \\ \mathbf{X}_t &= A' \cdot \left(\mathbf{X}_{t-1}^{(m)} \oplus \mathbf{Y}_{t-1}^{(m)} \right) = \varepsilon_t' \end{aligned} \tag{2.1}$$

where m is now referred to as the 'model order', A and A' contain the model coefficients, and ε_t and ε_t' are the residuals (also called 'prediction errors' or 'innovations') of the models. Note that the temporal duration of $\mathbf{X} \oplus \mathbf{Y}$ is the same as that of **X** and **Y** individually, and we assume that **X** and **Y** are both zero-mean. In practice, A (and A') and hence ε_t (and ε_t') can be derived by standard linear autoregression methods, including ordinary least squares and multivariate Yule–Walker equations (Kay, 1988).

In this framework, if the variability of the residual (ε_t') of the unrestricted model is significantly less than the variability of the residual (ε_t) of the restricted model, then there is an improvement in the prediction of **X** due to **Y**. Then, following Geweke (1984), for scalar X, the WGC from **Y** to X can be defined as:

$$\mathcal{F}_{\mathbf{Y} \to X} = \ln \frac{\mathrm{var}(\varepsilon_t)}{\mathrm{var}(\varepsilon_t')}. \tag{2.2}$$

Eq. (2.2) has several implications. First, for scalar X and Y it is possible both for Y to G-cause X and for X to G-cause Y, a feedback stochastic process. This relation generalizes to vector **X** and **Y** (see below, and Geweke (1984)). Second, $\mathcal{F}$ can never be negative. Third, statistical significance can be determined via the F-statistic (Greene, 2002):

$$\mathcal{F} = \frac{\frac{RSS_{\mathrm{r}} - RSS_{\mathrm{ur}}}{m}}{\frac{RSS_{\mathrm{ur}}}{T-2m-1}} \tag{2.3}$$

where RSS_{r} and RSS_{ur} are the Residual Sum of Squares of the restricted ($\sum_{t=m+1}^{T} \varepsilon_t^2$) and unrestricted ($\sum_{t=m+1}^{T} \varepsilon_t'^2$) models, respectively, and T is the total number of observations used to estimate the unrestricted model. The F-statistic approximately follows an F distribution with degrees of freedom m and $(T-2m-1)$. A significant F-statistic[2] may be reasonably interpreted as evidence that the unrestricted model provides a better prediction than does the restricted model, and in that case **Y** is said to G-cause X.

Geweke (1984) showed that a conditional WGC may be calculated by including in Eq. (2.1) another WSS variable **Z**:

$$\begin{aligned} \mathbf{X}_t &= A \cdot \left(\mathbf{X}_{t-1}^{(m)} \oplus \mathbf{Z}_{t-1}^{(m)} \right) + \varepsilon_t \\ \mathbf{X}_t &= A' \cdot \left(\mathbf{X}_{t-1}^{(m)} \oplus \mathbf{Y}_{t-1}^{(m)} \oplus \mathbf{Z}_{t-1}^{(m)} \right) + \varepsilon_t' \end{aligned} \tag{2.4}$$

where vectors **X**, **Y**, and **Z** represent zero-mean stationary stochastic processes. Then, the WGC from **Y** to scalar X, conditional on **Z**, is defined as:

$$\mathcal{F}_{\mathbf{Y} \to X \mid \mathbf{Z}} = \ln \frac{\mathrm{var}(\varepsilon_t)}{\mathrm{var}(\varepsilon_t')}. \tag{2.5}$$

Geweke made several other important contributions to the WGC concept (Geweke, 1982; Geweke, 1984). First, he showed that the total interdependence between two variables could be decomposed in terms of their reciprocal causality plus an 'instantaneous feedback' term. Second, he showed that under fairly general conditions $\mathcal{F}$ can be decomposed additively by frequency (see Spectral WGC section). Last, and less appreciated, he pointed out that it is possible to calculate WGCs among multivariate *sets* of variables, represented as vector time series. Considering again Eq. (2.4), there is no problem in allowing **Y** and **Z** to be vector time series: the residuals ε_t will still be scalar and Eq. (2.5) can be used without modification. However, if the 'predictee' variable **X** is a vector, then the residuals will also be vectors and Eq. (2.5) cannot be used directly. A natural approach suggested by Geweke, and developed further by Barrett et al. (in press), is to replace $\mathrm{var}(\varepsilon_t)$ with the determinant of the covariance matrix of the residuals (the generalized variance), though see Ladroue et al. (2009) for an alternative.

Computation of WGC requires specification of model order m. Too low m can lead to a poor representation of the data, whereas too high m can lead to problems of model estimation. A principled means to specify m is to minimize a criterion that balances the variance accounted for, against the number of coefficients to be estimated. Two

[1] Since X and Y may be scalar as well as vector time series, we use the inclusive term 'AR' (AutoRegressive) to refer to models in which the variables may be of either type.

[2] Statistical significance can also be calculated using the χ^2 statistic, which is monotonically related to the F-statistic (Geweke, 1982). The maximum likelihood estimator $\hat{\mathcal{F}}_{\mathbf{Y} \to X \mid \mathbf{Z}}$ will have a χ^2-distribution under the null hypothesis $\mathcal{F}_{\mathbf{Y} \to X \mid \mathbf{Z}} = 0$, and a non-central χ^2-distribution under the alternative hypothesis $\mathcal{F}_{\mathbf{Y} \to X \mid \mathbf{Z}} > 0$ (Granger, 1969; Geweke, 1982).

suitable criteria are the Akaike information criterion (Akaike, 1974) and the Bayesian information criterion (Schwartz, 1978).

Spectral WGC

As mentioned above, an important advance in developing WGC methodology was to provide a spectral (frequency domain) decomposition of WGC (Geweke, 1982; Geweke, 1984). In Geweke's spectral formulation, the total time-domain WGC is equal to the sum of spectral WGC components over all frequencies from zero to the Nyquist frequency. The spectral WGC at a given frequency ω, for scalar X and Y, is given by:

$$\mathcal{F}_{Y\to X}(\omega) = \ln \frac{S_{XX}(\omega)}{\tilde{H}_{XX}(\omega)\mathrm{var}(\varepsilon_t')\tilde{H}^*_{XX}(\omega)} \tag{2.6}$$

where $S_{XX}(\omega)$ is the autospectrum of X_t, $\tilde{H}_{XX}(\omega)$ is the (X, X) element of the normalized form of the transfer matrix $\tilde{\boldsymbol{H}}(\omega)$ (Geweke, 1982), $\tilde{H}^*_{XX}(\omega)$ is the complex conjugate of $\tilde{H}_{XX}(\omega)$, and ω denotes frequency. It should be noted that $\tilde{\boldsymbol{H}}(\omega)$ is the inverse of the normalized spectral coefficient matrix $\tilde{\boldsymbol{A}}(\omega)$, which in turn is derived as the Fourier transform of the time-domain coefficient matrix jointly representing the unrestricted models for X and Y (see Ding et al. (2006) for details). The form of Eq. (2.6) provides an important intuition: the causal influence depends on the relative sizes of the total power ($S_{XX}(\omega)$) and the intrinsic power ($H_{XX}(\omega)\mathrm{var}(\varepsilon_t')H^*_{XX}(\omega)$). Since the total power is the sum of the intrinsic and causal powers, the spectral WGC is zero when the causal power is zero, i.e. the intrinsic power equals the total power, and increases as the causal power increases (Ding et al., 2006). A conditional form of spectral WGC has been described by Chen et al. (2006).

Spectral WGC is important in neurophysiological studies because causal influences between neuronal populations often depend on oscillatory synchrony. In such cases, spectral WGC, like spectral coherence, is band-limited so that both the coherence and WGC spectra consist of band-limited peaks (Brovelli et al., 2004). It is an empirical question whether a peak in the coherence spectrum has corresponding peaks in the directional WGC spectra, and, if so, in what proportion. Nonetheless, because the coherence is directly related to the total interdependence (Ding et al., 2006), and the total interdependence is the sum of three measures (the two directional WGCs and the 'instantaneous feedback' WGC), coherence peaks are expected to have corresponding peaks in one or more of these WGC spectra.

Several other measures related to spectral WGC have been described. The Directed Transfer Function (DTF) is a frequency-domain measure of causal influence based on the elements of the transfer matrix $\boldsymbol{H}(\omega)$ of a multivariate AR model. The DTF has both normalized (Kaminski and Blinowska, 1991) and non-normalized (Kaminski, 2007) forms. Another measure, called the Partial Directed Coherence (PDC) (Baccala and Sameshima, 2001), is constructed from elements of $\boldsymbol{A}(\omega)$, the Fourier-transformed matrices of the coefficient matrices of a multivariate AR model. The PDC is normalized, but in a different way from the normalized DTF: the PDC represents the outflow from a variable j to i relative to all outflows from j, whereas the normalized DTF represents the inflow to i from j relative to all inflows to i. The comparison of WGC, DTF, and PDC is discussed in Eichler (2006) and Schelter et al. (2005).

Assumptions and challenges

In this section we identify some key challenges for WGC analysis, both in general and with regard to the commonly employed neuroimaging methods of fMRI, EEG, and MEG. In each case we outline recent work that may address, at least in part, these challenges.

Stationarity and data length

The measurement of WGC from neural time series data depends on estimating AR models of stochastic processes. It is usually required that the stochastic process be WSS so that model estimation is tractable. AR model estimation on nonstationary data is known to produce spurious regression results (Granger and Newbold, 1974). In practice, WSS can be assessed by verifying that the autocorrelation function is sharply declining, by ensuring that the process does not contain a 'unit root' (Hamilton, 1994),[3] and/or by applying the so-called KPSS test on the null hypothesis of stationarity (Kwiatkowski et al., 1992). For data that are nonstationary, a widely used approach is to *difference* the data (i.e., $X_t = X_t - X_{t-1}$), repeatedly if necessary. A problem with the use of differencing is that it may easily change the interpretation of any resulting WGCs, especially in the spectral domain where differencing acts as a high-pass filter. Alternatively, nonstationary data could indicate that causal relations are varying over time. In this case, it may make sense to use methods which are sensitive to time-varying WGC. One simple approach is to analyze shorter time-windows each of which may be locally stationary (Ding et al., 2000). More complex but potentially more widely applicable alternatives include spectral factorization of wavelet transformations (Dhamala et al., 2008), and adaptive recursive least-squares modelling (Hesse et al., 2003).

WGC analysis depends on having a total number of observations (time points, T) that is adequate to estimate the AR model coefficients. The observations may come from a single period of stationary time series if that period is of sufficient duration to provide the requisite number of observations. In this way analysis of resting states is possible if sufficiently long stationary rest periods are available. By assuming that the neural time series recorded in a stationary rest period are generated by an underlying stationary stochastic physiological process, an appropriately estimated AR model may be taken to represent that process. Alternatively, an AR model can be estimated using observations that come from multiple repetitions of relatively short periods (trials) if it can be assumed that each period is an independent realization of a single stationary stochastic process (Ding et al., 2000).

Following the latter approach, WGC has been measured from neural time series recorded during brief intervals of a cognitive task by using repeated task trials as observations for model coefficient estimation. Brovelli et al. (2004), for example, applied WGC analysis to brief intervals of local field potential time series simultaneously recorded from multiple somatosensory and motor cortical locations in monkeys maintaining hand pressure as part of a visual discrimination task. Oscillatory neural activity in somatosensory and motor cortices was synchronized in the beta (14–30 Hz) frequency range during hand pressure. WGC analysis revealed beta-frequency causal influences directed from somatosensory to motor cortex that were significantly stronger than in the reverse direction, suggesting that sensory and motor cortices are dynamically bound in a functional loop during maintenance of steady motor outflow.

Linearity and parametric estimation

In its standard application, WGC captures only linear features of time series data, whereas many target systems are known to be nonlinear. However, nonlinear systems often have extensive linear regimes, and, in many neuroscience applications, especially those dealing with large-scale interactions, linear approximations are found to work extremely well (McIntosh and Gonzalez-Lima, 1994). More generally, WGC is not tied to linear AR models (Freiwald et al., 1999;

[3] A unit root occurs when a root of the characteristic equation for an AR model equals one.

Marinazzo et al., 2010). Specific implementations of nonlinear WGC include a kernel-based method (Ancona et al., 2004; Marinazzo et al., 2008) which is suggested to be well suited to short and noisy time series. A recent result has indicated that, for Gaussian variables, there is nothing additional to account for by nonlinear extensions to WGC because a stationary Gaussian AR process is necessarily linear (Barnett et al., 2009). Nonlinear TSI methods outside the WGC framework are discussed briefly in the Relation to transfer entropy and other nonlinear methods section.

As previously noted, estimation of WGC using parametric AR models requires choosing a model order m that balances variance accounted for against model complexity. A problem with this trade-off is that resulting AR models can sometimes fail to capture complex spectral features of data that require higher-order AR models (Mitra and Pesaran, 1999). An alternative, nonparametric approach has recently been described (Dhamala et al., 2008) in which pairwise and conditional WGC are computed directly from Fourier and wavelet transformations, bypassing the step of (linear) AR model estimation. This nonparametric method is useful also because it provides a measure of time-varying WGC (if wavelets are used), and because it supplies a natural WGC analysis of point-process data, such as spike trains, based on their Fourier transforms (Nedungadi et al., 2009).

Latent variables

All brain connectivity models involve the step of 'structural model selection', in which a relevant set of neuronal variables is selected for analysis (Roebroeck et al., 2009). In practice, this step is likely to exclude some relevant variables, which, for all TSI-based analysis methods, can lead to the detection of apparent causal interactions that are actually spurious (Pearl, 1999).

One recent response to this challenge has been what is called 'partial G-causality' (Guo et al., 2008a). The idea is that latent variables may give rise to detectable correlations among the residuals of the corresponding vector AR model (i.e., cross-correlations among the ε_t of Eq. (2.4)). By analogy with the concept of partial correlation (Kendall and Stuart, 1979), an additional term based on these correlations can mitigate the confounding influence of the latent variables. Consider that, in addition to Eq. (2.4), the regressions of the *conditional* variable $\mathbf{Z}_t$ are:

$$\begin{aligned}\mathbf{Z}_t &= B\cdot\left(\mathbf{X}_{t-1}^{(m)}\oplus\mathbf{Z}_{t-1}^{(m)}\right)+\eta_t\\ \mathbf{Z}_t &= B'\cdot\left(\mathbf{X}_{t-1}^{(m)}\oplus\mathbf{Y}_{t-1}^{(m)}\oplus\mathbf{Z}_{t-1}^{(m)}\right)+\eta'_t\end{aligned}\tag{3.1}$$

so that the roles of the 'predictee' and conditioning variables are reversed. Then (for scalar X) the partial G-causality of $\mathbf{Y}$ to X given $\mathbf{Z}$ is defined as:

$$\mathcal{F}^{P}_{\mathbf{Y}\to X|\mathbf{Z}} = \ln\frac{\sum(\varepsilon_t|\eta_t)}{\sum(\varepsilon'_t|\eta'_t)}\tag{3.2}$$

where $\sum(\varepsilon|\eta)$ is defined as $\mathrm{cov}(\varepsilon,\varepsilon)-\mathrm{cov}(\varepsilon,\boldsymbol{\eta})\,\mathrm{cov}(\boldsymbol{\eta},\boldsymbol{\eta})^{-1}\,\mathrm{cov}(\varepsilon,\boldsymbol{\eta})^{\top}$, the 'partial covariance' (Barnett et al., 2009), and the superscript $^{\top}$ denotes matrix transpose.

Importantly, whereas partial correlation removes the influence of *known* exogenous variables from a fundamentally bivariate measure (correlation), $\mathcal{F}^P$ attempts to control for the influence of *unknown* variables on a multivariate measure (WGC) indirectly via their influence on ε_t. Therefore, $\mathcal{F}^P$ can only fully control for latent variables in the unlikely case that they have equivalent effects on all measured variables. However, numerical investigations show that even when this condition is not met, $\mathcal{F}^P$ nonetheless can deliver substantially improved results as compared to $\mathcal{F}$ (Guo et al., 2008a). It is also noteworthy that $\mathcal{F}^P$ can also be decomposed additively by frequency (Guo et al., 2008b).

fMRI

The application of WGC to fMRI is enormously promising given the dominance of fMRI within neuroimaging (Roebroeck et al., 2005). However, fMRI data is subject to several potential sources of artifact. In particular, variability in the shape and latency of Hemodynamic Response Functions (HRFs) in different brain regions and different subjects (Aguirre et al., 1998) may lead to mis-attribution of the direction of causal influence by WGC in some circumstances (David et al., 2008). The implications of this problem, especially with regard to the pros and cons of WGC versus DCM, are comprehensively covered by other authors in this issue. Here we emphasize four points. First, although the sources of inter-regional and inter-subject variability in the fMRI are only poorly understood, it is clear that any attribution of this variability entirely to HRF variability is erroneous. Recent simulation results indicate that WGC is surprisingly resilient to hemodynamic variability within normal physiological ranges (Deshpande et al., 2011-this issue). Second, the comparison of WGC values from fMRI Blood Oxygen Level Dependent (BOLD) time series in different experimental conditions may be robust to HRF variation, even if characterization of the underlying neural processes *per se* (Roebroeck et al., 2005) is not. This is because, since HRFs are not expected to vary between conditions, it is unlikely that HRF variation across brain regions affects the comparison of conditions. Third, adequate estimates of HRF shape and latency can be used to deconvolve the fMRI BOLD signal to recover measures of the underlying neural processes (David et al., 2008; Chang et al., 2008; Vakorin et al., 2007), which can then be analyzed using WGC to determine causal influences among neuronal populations. Fourth, knowledge of the HRFs is not necessary for WGC analysis to uncover causal relations among event-related BOLD time series (Tang et al., 2009).

WGC analysis of fMRI BOLD time series may serve as a useful complement to ablation and/or stimulation studies for characterizing the function of neural systems. For example, Bressler et al. (2008) found top-down directed influences from high-level frontal and parietal regions to visual occipital cortex in visual spatial attention that are consistent with the results of transcranial magnetic stimulation of the same high-level regions (Ruff et al., 2009). In another study, Sridharan et al. (2008) used WGC analysis of fMRI BOLD time series to identify the right fronto-insular cortex as a critical brain region that drives changes in the brain's central-executive and default-mode networks.

Other issues that may affect the application of WGC specifically to fMRI include: (i) the large number of possible variables (voxels); (ii) the likelihood of missing rapid causal influences because of the slow dynamics of the BOLD signal; and (iii) the necessity of finding appropriate statistical methods for group-level analyses. The first issue can be addressed in several ways, for example by using dimensionality reduction techniques such as principal components analysis prior to WGC (Zhou et al., 2009) or by using sparse regression techniques in combination with pruning of 'unlikely' connections (Valdes-Sosa et al., 2005). Furthermore, using Geweke's formulation of WGC for multivariate sets (see above), it is possible for causal influences to be measured between sets of simultaneous time series drawn from multiple single voxels, or principal components, in individual brain regions (Barrett et al., in press).

Although the second issue is more difficult to deal with, modelling studies have shown that fast neural exchanges can still be identified via a combination of short TR times, and by carrying out causal inference, not on the individual $\mathcal{F}_{X\to Y}$ and $\mathcal{F}_{Y\to X}$, but rather on their difference: $\mathcal{F}_{X\to Y}-\mathcal{F}_{Y\to X}$ (Roebroeck et al., 2005). This strategy controls for the loss of information that arises from the low-pass filtering introduced by the HRF, but at the expense of missing reciprocal causal connections. One caveat is that it is difficult to measure the conditional WGC in this case.

As for the third issue, group-level statistical methods are often needed in fMRI time series analysis. In the case of WGC mapping at the voxel level, group-level statistical methods are used to ascertain the threshold of significance. When WGC is measured at the brain-region level based on voxel-level statistical distributions (Bressler et al., 2008), the group-level statistical significance threshold may be obtained directly from the distribution of the *F*-statistic defined in Eq. (2.3).

EEG and MEG

The application of WGC analysis to MEG or EEG data can be extremely informative due to the submillisecond time resolution offered by these imaging methods. The demonstrated usefulness of spectral WGC for discriminating causal influences at different frequencies in MEG and EEG data (Astolfi et al., 2007; Gow et al., 2008), as in local field potential data (Brovelli et al., 2004; Bressler et al., 2007), is a particular benefit of WGC analysis that is not shared by some other TSI-based analysis methods. Relating causal connectivity in either the time or frequency domains to cognitive function is facilitated if WGC can be applied to source-localized MEG or EEG signals. However, both MEG and EEG suffer from uncertainties in the transformation from sensor space to source space. Although source localization methods are continually improving (Pantazis and Leahy, 2006), generally accepted methods are still lacking and the existence of unique solutions cannot be guaranteed by any method (Koles, 1998).

A further problem in applying WGC analysis to MEG and EEG data is that it is difficult to ensure that the many preprocessing steps involved in MEG and EEG data processing (even prior to source localization) do not introduce causal artifacts by disrupting the fine-grained timing relations on which WGC analysis depends. For example, bandpass filtering may cause severe confounds in WGC analysis by introducing temporal correlations into MEG or EEG time series (Seth, 2010; Florin et al., 2009).

Discussion

Causal inference in neural systems

Any discussion of causal inference should be prefaced by one fundamental question: What do we expect from a measure of causality? One useful perspective is that causal measures in neuroscience should reflect *effective connectivity*, namely the directed influences that neuronal populations in one brain area exert on those in another (Friston, 1994). As applied to neuroimaging data (e.g., by way of DCM), effective connectivity analysis aims at identifying, from recorded neural time series data, the underlying physiological influences exerted among neuronal populations in different brain areas. Effective connectivity can be distinguished from *functional connectivity*, which is based on correlations at zero-lag or across multiple time lags (Friston, 1994; Daunizeau et al., 2011-this issue).

One may, however, envisage a different but complementary goal for a causality measure, namely to reflect directed dynamical connectivity — causal connectivity — *without* requiring that the resulting networks univocally recapitulate the underlying physiological processes. A causal connectivity perspective can be justified in several ways. First, modelling studies have demonstrated that the same underlying (physical) network structure can give rise to multiple distinct dynamical connectivity patterns (Seth, 2005, 2008; Lungarella and Sporns, 2006) depending on how the system interacts with its environment. Second, in practice it is always infeasible to measure *all* relevant variables. Therefore effective connectivity measures will always be provisional (unless somehow validated by intervention techniques), whereas descriptions of causal connectivity stand as valid descriptions of dynamical relations among measured variables regardless of omitted elements.[4] Third, neural dynamics can modify the underlying neural structural connectivity, for example via Hebbian processes, implying that dynamical and structural connectivity patterns are engaged in continual interaction and mutual specification, and therefore that any assumptions by effective connectivity analysis of invariant underlying physiological influences may be unrealistic.

Our aim is not to argue against effective connectivity analysis. On the contrary, descriptions of effective connectivity are greatly to be desired (Horwitz, 2003). Nevertheless, it may also make sense to consider the more liberal level of causal connectivity analysis (Seth, 2005, 2008) as shedding useful new light on the general relations that exist between structural and dynamical levels of description of complex neural systems.

Exploratory versus confirmatory statistics

WGC is often described as an example of exploratory statistics, in the sense that it is data-driven, making few, if any, prior assumptions about causal connectivity (Roebroeck et al., 2009). Exploratory statistics can be contrasted with confirmatory statistics, in which a causal model is specified in advance and data are used to confirm or disconfirm the plausibility of this model (McIntosh and Gonzalez-Lima, 1994). From this perspective, DCM is often considered to be an example of confirmatory statistics (Roebroeck et al., 2009). However, it may be more productive to think of both WGC and DCM as occupying different and modifiable positions on a spectrum from 'purely exploratory' to 'purely confirmatory' methods. WGC analysis is confirmatory in the sense that it makes a 'model-based' assumption that the underlying data-generating mechanism can be effectively modelled as an AR process (see Relation to transfer entropy and other nonlinear methods section below). And recent implementations of DCM may be considered to be exploratory since they include a model selection stage in which different causal models are compared based on their log evidence (Friston, 2011-this issue). By increasing the range of candidate models in a DCM approach, or by introducing priors into the AR equations, WGC and DCM could be brought even closer together. In the limit, as Friston (this issue) points out, both WGC and DCM are based on TSI and can be subsumed within a single theoretical framework.

In their most common instantiations, DCM and WGC are distinguished in several important ways with regard to neuroimaging applications. We remark on three (see Roebroeck et al. (2009) and commentaries in this issue for a more comprehensive discussion). First, DCM incorporates an explicit model of the neuronal causes of observed data whereas WGC derives inferences directly from data (this being the case even following modality-specific preprocessing such as hemodynamic deconvolution). Second, unlike WGC, DCM does not incorporate stochastic processes, though there are recent extensions of DCM in this direction (Daunizeau et al., 2011-this issue). Third, whereas DCM is based on Bayesian model inversion, WGC makes use of a classical frequentist approach in which conclusions are drawn on the basis of distributions of sampling statistics. More generally, as noted above in the Causal inference in neural systems section, DCM aims at identifying *effective* connectivity whereas WGC may be more broadly useful for inferring *causal* connectivity.

Relation to transfer entropy and other nonlinear methods

The concept of transfer entropy (TE) (Schreiber, 2000) is extremely similar to that of WGC. TE is an information-theoretic

[4] In this regard, analysis based on partial WGC ($\mathcal{F}^p$) must always be considered to be incomplete since it must always be performed against a background of unmeasured environmental variables. On the other hand, it may be argued that the causal connectivity relations among a set of *measured* variables are properly given by the conditional WGC ($\mathcal{F}$).

measure of directed (time-asymmetric) information transfer between jointly dependent processes. While WGC is framed in terms of prediction, TE is framed in terms of resolution of uncertainty: The TE from **Y** to **X** is the degree to which **Y** disambiguates the future of **X** beyond the degree to which **X** already disambiguates its own future. A possible advantage of TE over WGC is that TE is 'model agnostic' inasmuch as it depends only on estimation of multivariate entropies directly from probability distributions.[5] As such, TE may make fewer assumptions about the data than the standard WGC implementation. On the other hand, the estimation of TE by state-space partitioning is problematic, especially in multivariate situations (Kaiser and Schreiber, 2002). Moreover, unlike WGC, the distribution of the sample statistic is not known, rendering significance testing difficult without recourse to computationally expensive nonparametric resampling techniques (Theiler et al., 1992). It has recently been shown that, for Gaussian variables, WGC and TE are in fact entirely equivalent (Barnett et al., 2009). By unifying information-theoretic and autoregressive approaches to causal inference, this result allows insights from one domain to be translated easily into the other. For example, as mentioned above, it now appears that under Gaussian assumptions there is nothing additional for nonlinear WGC to account for.

WGC and TE are both part of a large repertoire of nonlinear TSI techniques that rely on the concept of reduction of uncertainty in future states. Although less often employed than WGC, alternatives within this repertoire may have advantages in particular applications. For example, Schiff et al. (1996) describe a nonlinear prediction technique based on time-delayed coordinate embedding, their emphasis, however, being more on identifying generalized synchrony than causality *per se*. A similar approach has been applied to multichannel EEG, revealing correlated patterns of nonlinear interdependence (Breakspear and Terry, 2002). Lungarella et al. (2007) provide a useful comparative analysis of some of these techniques, including WGC and TE.

Resources

Because WGC is conceptually straightforward, it is a relatively attractive methodology, and a variety of software resources have been developed which implement WGC analysis. The BSMART toolbox (http://www.brain-smart.org) is a GUI-based package specifically targeted toward the WGC analysis of neural time series data (Cui et al., 2008). The GCCA toolbox (http://www.anilseth.com) is a MATLAB (Natick, MA) based toolbox that provides a core set of WGC functions applicable both to neuroimaging data and to time series from other sources (Seth, 2009). Both packages are freely available under the GNU software license and have been used successfully by multiple groups in neuroimaging analysis (Saalmann et al., 2009; Hermer-Vazquez, 2008; Sridharan et al., 2008; Gaillard et al., 2009; Chang et al., 2008).

Conclusions

We have described Wiener–Granger Causality (WGC) analysis as a robust form of Time Series Inference (TSI), having an intuitive definition, a straightforward description, and readily-available means of implementation. WGC, and related measures, are finding increasing utility in the neuroscientific laboratory. For understanding the neural basis of cognitive functions, such as the top-down cortical mechanism of visual selective attention, WGC analysis is providing useful results that complement those from stimulation techniques such as transcranial magnetic stimulation and electrical microstimulation, and from lesion techniques, such as reversible focal cryogenic blockade.

[5] In fact, as noted above, the nonparametric implementation of WGC by Dhamala et al. (2008) bypasses AR model estimation, and thus may also be considered 'model agnostic'.

Acknowledgments

AKS is supported by EPSRC Leadership Fellowship EP/G007543/1 and by a donation from the Dr. Mortimer and Theresa Sackler Foundation. The authors thank Ms. Wei Tang, Dr. Lionel Barnett, Dr. Adam Barrett, Dr. Alard Roebroeck, Dr. Abraham Snyder, and our reviewers for helpful comments.

References

Aguirre, G.K., Zarahn, E., D'esposito, M., 1998. The variability of human, BOLD hemodynamic responses. Neuroimage 8, 360–369.
Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Automatic Control AC-19, 716–723.
Ancona, N., Marinazzo, D., Stramaglia, S., 2004. Radial basis function approaches to nonlinear Granger causality of time series. Phys. Rev. E 70, 056221.
Astolfi, L., Cincotti, F., Mattia, D., Marciani, M.G., Baccala, L.A., de Vico, Fallani F., Salinari, S., Ursino, M., Zavaglia, M., Ding, L., Edgar, J.C., Miller, G.A., He, B., Babiloni, F., 2007. Comparison of different cortical connectivity estimators for high-resolution eeg recordings. Hum. Brain Mapp. 28, 143–157.
Baccala, L.A., Sameshima, K., 2001. Partial directed coherence: a new concept in neural structure determination. Biol. Cybern. 84, 463–474.
Barnett, L., Barrett, A.B., Seth, A.K., 2009. Granger causality and transfer entropy are equivalent for Gaussian variables. Phys. Rev. Lett. 103, 238701.
Barrett, A.B., Barnett, L., Seth, A.K., in press. Multivariate Granger causality and generalized variance. Phys Rev E.
Breakspear, M., Terry, J.R., 2002. Topographic organization of nonlinear interdependence in multichannel human EEG. Neuroimage 16, 822–835.
Bressler, S.L., Richter, C.G., Chen, Y., Ding, M., 2007. Cortical functional network organization from autoregressive modeling of local field potential oscillations. Stat. Med. 26, 3875–3885.
Bressler, S.L., Tang, W., Sylvester, C.M., Shulman, G.L., Corbetta, M., 2008. Topdown control of human visual cortex by frontal and parietal cortex in anticipatory visual spatial attention. J. Neurosci. 28, 10056–10061.
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 (Electronic publication 2004 Jun 21).
Chang, C., Thomason, M.E., Glover, G.H., 2008. Mapping and correction of vascular hemodynamic latency in the BOLD signal. Neuroimage 43, 90–9102.
Chen, Y., Bressler, S.L., Ding, M., 2006. Frequency decomposition of conditional Granger causality and application to multivariate neural field potential data. J. Neurosci. Methods 150, 228–237.
Cui, J., Xu, L., Bressler, S.L., Ding, M., Liang, H., 2008. BSMART: a Matlab/C toolbox for analysis of multichannel neural time series. Neural Netw. 21, 1094–1104.
Daunizeau, J., David, O., Stephan, K.E., 2011. Dynamic causal modelling: a critical review of the biophysical and statistical foundations. Neuroimage. 58, 312–322 (this issue).
David, O., Guillemain, I., Saillet, S., Reyt, S., Deransart, C., Segebarth, C., Depaulis, A., 2008. Identifying neural drivers with functional MRI: an electrophysiological validation. PLoS Biol. 6, 2683–2697.
Deshpande, G., Sathian, K., Hu, X., 2011. Effect of hemodynamic variability on Granger causality analysis of fMRI. Neuroimage. 52, 884–896 (this issue).
Dhamala, M., Rangarajan, G., Ding, M., 2008. Analyzing information flow in brain networks with nonparametric Granger causality. Neuroimage 41, 354–362.
Ding, M., Bressler, S., Yang, W., Liang, H., 2000. Short-window spectral analysis of cortical event-related potentials by adaptive multivariate autoregressive modeling: data prepocessing, model validation, and variability assessment. Biol. Cybern. 83, 35–45.
Ding, M., Chen, Y., Bressler, S.L., 2006. Granger causality: basic theory and application to neuroscience. In: Schelter, B., Winterhalder, M., Timmer, J. (Eds.), Handbook of Time Series Analysis: Recent Theoretical Developments and Applications. Wiley-VCH, Berlin, pp. 437–460.
Eichler, M., 2006. On the evaluation of information ow in multivariate systems by the directed transfer function. Biol. Cybern. 94, 469–482.
Florin, E., Gross, J., Pfeifer, J., Fink, G.R., Timmermann, L., 2009. The effect of filtering on Granger causality based multivariate causality measures. Neuroimage 47, Supplement 1, S101.
Freiwald, W., Valdes, P., Bosch, J., Biscay, R., Jimenez, J., Rodriguez, L., Rodriguez, V., Kreiter, A., Singer, W., 1999. Testing non-linearity and directedness of interactions between neural groups in the macaque inferotemporal cortex. J. Neurosci. Methods 94, 105–119.
Friston, K., 1994. Functional and effective connectivity in neuroimaging: a synthesis. Hum. Brain Mapp. 2, 56–78.
Friston, K., Harrison, L., Penny, W., 2003. Dynamic causal modeling. Neuroimage 19, 1273–1302.
Friston, K., 2011. Dynamic causal modeling and Granger causality Comments on: the identification of interacting networks in the brain using fMRI: Model selection, causality and deconvolution. Neuroimage. 58, 303–305 (this issue).
Gaillard, R., Dehaene, S., Adam, C., Clémenceau, S., Hasboun, D., Baulac, M., Cohen, L., Naccache, L., 2009. Converging intracranial markers of conscious access. PLoS Biol. 7, e61-e61.
Geweke, J., 1982. Measurement of linear dependence and feedback between multiple time series. J. Am. Stat. Assoc. 77, 304–313.
Geweke, J., 1984. Measures of conditional linear dependence and feedback between time series. J. Am. Stat. Assoc. 79, 907–915.

Gow, D.W., Segawa, J.A., Ahlfors, S.P., Lin, F.H., 2008. Lexical influences on speech perception: a Granger causality analysis of MEG and EEG source estimates. Neuroimage 43, 614–623.
Granger, C., 1969. Investigating causal relations by econometric models and cross-spectral methods. Econometrica 37, 424–438.
Granger, C., Newbold, P., 1974. Spurious regressions in econometrics. J. Econometrics 2, 111–120.
Greene, W.H., 2002. Econometric Analysis, fifth ed. Prentice-Hall, Upper Saddle River, NJ.
Guo, S., Seth, A.K., Kendrick, K., Zhou, C., Feng, J., 2008a. Partial Granger causality: eliminating exogenous inputs and latent variables. J. Neurosci. Methods 172, 79–93.
Guo, S., Wu, J., Ding, M., Feng, J., 2008b. Uncovering interactions in the frequency domain. PLoS Comput. Biol. 4, e1000087.
Hamilton, J.D., 1994. Time Series Analysis. Princeton University Press, Princeton, NJ.
Hermer-Vazquez, L., 2008. Tracing 'driver' versus 'modulator' information ow throughout large-scale, task-related neural circuitry. J Comb. Optim. 15, 242–256.
Hesse, W., Möller, E., Arnold, M., Schack, B., 2003. The use of time-variant EEG Granger causality for inspecting directed interdependencies of neural assemblies. J. Neurosci. Methods 124, 27–44.
Horwitz, B., 2003. The elusive concept of brain connectivity. Neuroimage 19, 466–470.
Kaiser, A., Schreiber, T., 2002. Information transfer in continuous processes. Physica D 166, 43–62.
Kaminski, M., 2007. Multichannel data analysis in biomedical research. In: Jirsa, V.K., McIntosh, A.R. (Eds.), Handbook of Brain Connectivity. Springer, Berlin, pp. 327–355.
Kaminski, M., Blinowska, K.J., 1991. A new method of the description of the information ow in the brain structures. Biol. Cybern. 65, 203–210.
Kay, S., 1988. Modern Spectral Estimation: Theory and Application Prentice-Hall. Englewood Cliffs, NJ.
Kendall, M.G., Stuart, A., 1979. The advanced theory of statistics. Inference and Relationship, vol. 2. Charles Griffin and Co., Buckinghamshire, England.
Koles, Z.J., 1998. Trends in EEG source localization. Electroencephalogr. Clin. Neurophysiol. 106, 127–137.
Kwiatkowski, D., Phillips, P., Schmidt, P., Shin, Y., 1992. Testing the null hypothesis of stationarity against the alternative of a unit root. J. Econometrics 54, 159–178.
Ladroue, C., Guo, S., Kendrick, K., Feng, J., 2009. Beyond element-wise interactions: identifying complex interactions in biological processes. PLoS One 4, e6899.
Lungarella, M., Ishiguro, K., Kuniyoshi, Y., Otsu, N., 2007. Methods for quantifying the causal structure of bivariate time series. Int. J. Bifurcation Chaos 17, 903–921.
Lungarella, M., Sporns, O., 2006. Mapping information flow in sensorimotor networks. PLoS Comput. Biol. 2, e144-e144.
Marinazzo, D., Pellicoro, M., Stramaglia, S., 2008. Kernel method for nonlinear Granger causality. Phys. Rev. Lett. 100, 144103.
McIntosh, A., Gonzalez-Lima, F., 1994. Structural equation modeling and its application to network analysis in functional brain imaging. Hum. Brain Mapp. 2, 2–22.
Mitra, P.P., Pesaran, B., 1999. Analysis of dynamic brain imaging data. Biophys. J. 76, 691–708.
Nedungadi, A., Rangarajan, G., Jain, N., Ding, M., 2009. Analyzing multiple spike trains with nonparametric Granger causality. J. Comput. Neurosci. 27, 55–64.
Pantazis, D., Leahy, R.M., 2006. Imaging the human brain with magnetoencephalography: basics and current development. In: Lazakidou, A. (Ed.), Handbook of Research on Informatics in Healthcare and Biomedicine. Idea Group Publishing, Hershey, PA, pp. 294–302.
Pearl, J., 1999. Causality: Models, Reasoning, and Inference. Cambridge University Press, Cambridge, UK.
Roebroeck, A., Formisano, E., Goebel, R., 2005. Mapping directed inuence over the brain using Granger causality and fmri. Neuroimage 25, 230–242.
Roebroeck, A., Formisano, E., Goebel, R., 2009. The identification of interacting networks in the brain using fMRI: model selection, causality and deconvolution. Neuroimage.
Ruff, C.C., Driver, J., Bestmann, S., 2009. Combining TMS and fMRI: from 'virtual lesions' to functional-network accounts of cognition. Cortex 45, 1043–1049.
Saalmann, Y.B., Balsters, J.H., Wright, M.J., Ramnani, N., 2009. Learning rules: investigating prefrontal-cerebellar connectivity with Granger causality. Neuroimage 47, S39–S41.
Schelter, B., Winterhalder, M., Eichler, M., Peifer, M., Hellwig, B., Guschlbauer, B., Lucking, C.H., Dahlhaus, R., Timmer, J., 2005. Testing for directed inuences among neural signals using partial directed coherence. J. Neurosci. Methods 152, 210–219.
Schiff, S.J., So, P., Chang, T., Burke, R.E., Sauer, T., 1996. Detecting dynamical interdependence and generalized synchrony through mutual prediction in a neural ensemble. Phys. Rev. E 54, 6708–6724.
Schreiber, T., 2000. Measuring information transfer. Phys. Rev. Lett. 85, 461–464.
Schwartz, G., 1978. Estimating the dimension of a model. Ann. Stat. 6, 461–464.
Seth, A.K., 2005. Causal connectivity of evolved neural networks during behavior. Netw. Comput. Neural Systems 16, 35–54.
Seth, A.K., 2008. Causal networks in simulated neural systems. Cogn. Neurodyn. 2, 49–64.
Seth, A.K., 2010. A MATLAB toolbox for Granger causal connectivity analysis. J. Neurosci. Methods 186, 262–273.
Sridharan, D., Levitin, D.J., Menon, V., 2008. A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. Proc. Natl. Acad. Sci. U. S. A. 105, 12569–12574.
Tang, W., Bressler, S.L., Sylvester, C.M., Shulman, G.L., Corbetta, M., 2009. fMRI-based Granger causality is an effective measure of effective connectivity, vol. 188.15.
Theiler, J., Eubank, S., Longtin, A., Galdrakian, B., Farmer, J., 1992. Testing for nonlinearity: the method of surrogate data. Physica D 58, 77–94.
Vakorin, V.A., Krakovska, O.O., Borowsky, R., Sarty, G.E., 2007. Inferring neural activity from bold signals through nonlinear optimization. Neuroimage 38, 248–260.
Valdes-Sosa, P.A., Sanchez-Bornot, J.M., Lage-Castellanos, A., Vega-Hernandez, M., Bosch-Bayard, J., Melie-Garcia, L., Canales-Rodriguez, E., 2005. Estimating brain functional connectivity with sparse multivariate autoregression. Philos. Trans. R. Soc. B 360, 969–981.
Wiener, N., 1956. The theory of prediction. In: Beckenbach, E. (Ed.), Modern Mathematics for Engineers. McGraw-Hill, New York.
Zhou, Z., Ding, M., Chen, Y., Wright, P., Lu, Z., Liu, Y., 2009. Detecting directional influence in fMRI connectivity analysis using PCA based Granger causality. Brain Res. 1289, 22–29.