



CORRESPONDENCE

On the interpretability and computational reliability of frequency-domain Granger causality [version 1; referees: 2 approved]

Luca Faes ¹, Sebastiano Stramaglia^{2,3}, Daniele Marinazzo⁴

¹BIOtech, Department of Industrial Engineering, University of Trento, Trento, Italy

²Istituto Nazionale di Fisica Nucleare, Sezione di Bari, Bari, Italy

³Dipartimento di Fisica, Università degli Studi Aldo Moro, Bari, Italy

⁴Department of Data Analysis, Ghent University, Ghent, Belgium

v1 First published: 20 Sep 2017, 6:1710 (doi: [10.12688/f1000research.12694.1](https://doi.org/10.12688/f1000research.12694.1))
 Latest published: 20 Sep 2017, 6:1710 (doi: [10.12688/f1000research.12694.1](https://doi.org/10.12688/f1000research.12694.1))

Abstract

This Correspondence article is a comment which directly relates to the paper “A study of problems encountered in Granger causality analysis from a neuroscience perspective” (Stokes and Purdon, 2017). We agree that interpretation issues of Granger causality (GC) in neuroscience exist, partially due to the historically unfortunate use of the name “causality”, as described in previous literature. On the other hand, we think that Stokes and Purdon use a formulation of GC which is outdated (albeit still used) and do not fully account for the potential of the different frequency-domain versions of GC; in doing so, their paper dismisses GC measures based on a suboptimal use of them. Furthermore, since data from simulated systems are used, the pitfalls that are found with the used formulation are intended to be general, and not limited to neuroscience. It would be a pity if this paper, even if written in good faith, became a wildcard against all possible applications of GC, regardless of the large body of work recently published which aims to address faults in methodology and interpretation. In order to provide a balanced view, we replicate the simulations of Stokes and Purdon, using an updated GC implementation and exploiting the combination of spectral and causal information, showing that in this way the pitfalls are mitigated or directly solved.

Open Peer Review

Referee Status:

	Invited Referees	
	1	2
version 1 published 20 Sep 2017	 report	 report
1 Luiz Antonio Baccala , University of São Paulo, Brazil		
2 Karin Schiecke , Friedrich Schiller University Jena, Germany		

Discuss this article

Comments (0)

Corresponding author: Luca Faes (faes.luca@gmail.com)

Author roles: **Faes L:** Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Software, Supervision, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Stramaglia S:** Conceptualization, Investigation, Methodology, Project Administration, Supervision, Validation, Visualization, Writing – Review & Editing; **Marinazzo D:** Conceptualization, Data Curation, Investigation, Methodology, Project Administration, Resources, Software, Supervision, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

How to cite this article: Faes L, Stramaglia S and Marinazzo D. **On the interpretability and computational reliability of frequency-domain Granger causality [version 1; referees: 2 approved]** *F1000Research* 2017, **6**:1710 (doi: [10.12688/f1000research.12694.1](https://doi.org/10.12688/f1000research.12694.1))

Copyright: © 2017 Faes L *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. Data associated with the article are available under the terms of the [Creative Commons Zero "No rights reserved" data waiver](#) (CC0 1.0 Public domain dedication).

Grant information: The work was funded by the Department of Industrial Engineering, University of Trento, Italy (40300314). *The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

First published: 20 Sep 2017, **6**:1710 (doi: [10.12688/f1000research.12694.1](https://doi.org/10.12688/f1000research.12694.1))

Correspondence

Granger causality (GC)¹ is an extremely popular statistical tool used to analyze directed interactions from multivariate time series measured from coupled dynamical systems. A particularly appealing aspect of the notion of GC is that it can be formulated in the frequency domain, and is thus eligible for the analysis of signals that are rich of oscillatory content such as those commonly encountered in neuroscience and physiology^{2,3}. The spectral formulation of GC is obtained by elaborating in the frequency domain the parameters of the linear vector autoregressive (VAR) model that fit the observed multivariate time series. A main approach to do this was developed by Geweke^{4,5}, yielding bivariate and conditional frequency domain measures of the so-called Granger-Geweke causality (GGC). An alternative framework stems from the works of Kaminski *et al*⁶ and Baccalà *et al*^{7,8}, who derived measures like the directed coherence (DC) and the partial DC (PDC), quantifying the total and direct directed influence of one time series over another in a fully multivariate setting (see references 9–11 for comprehensive treatments).

In their recent work¹², Stokes and Purdon performed a critical evaluation of frequency-domain GC computed within the Geweke framework, evidencing in two simulation studies some computational and interpretational problems associated with the GGC measures. Specifically, they showed that – even when the systems generating the observed data belong to the finite-order VAR model class – spectral GGC cannot be reliably estimated and cannot recover the functional oscillatory structure underlying the data. These observations led the authors to conclude that the notion of causality quantified by GGC, and by other Granger causality measures in general, often yield counterintuitive and misleading results, thus being incompatible with the objectives of many neuroscience studies.

We definitely agree that GC and lag-based data-driven methods in general cannot provide measures of “causality” as intended in other applications (see references 9 and 13 for a thoughtful distinction between data-driven and model-based approaches). We also share the view that the assumptions of linearity and stationarity, as well as the presence of unobserved variables, noise or inappropriate sampling may pose theoretical and practical problems which can severely impair both the formulation and the computation of spectral GC measures – this has been stated by Stokes and Purdon¹² and in previous studies^{2,3,14}. On the other hand we think that, based on the way simulated data have been analyzed and interpreted by Stokes and Purdon¹², frequency domain GC methods have been dismissed based on a suboptimal (even though frequently applied) formulation of GGC, and based on the lack of direct consideration of the DC/PDC framework.

In this contribution, we repeat the simulations of Stokes and Purdon¹², and suggest that the negative conclusions based on the results of such simulations are overstated. We show that spectral GGC estimates can be obtained with a high computational

reliability if proper estimation approaches are employed, and the interpretation of frequency domain causality measures can be meaningfully performed if spectral and causal information are properly combined. The codes for running our analyses are based on existing Matlab[®] toolboxes^{11,15,16} and are provided as supplementary data alongside this article.

The first simulation of Stokes and Purdon¹² shows that, due to the modeling choices required to compute spectral GGC, this measure cannot be reliably estimated even for simple systems. By generating 100 realizations of this simulation with the same parameters and data length we confirm that, by applying the standard method of fitting separate full and reduced VAR models, spectral GGC estimates display a strong bias (Figure 1A) or a very large variability (Figure 1B), depending on the choice of the model order. As explained by Stokes and Purdon¹², this tradeoff between bias and variance arises from the incorrect representation of the reduced model as a VAR process of finite order. Exactly for this reason however, the problem can be overcome employing the state-space (SS) approach¹⁶, which allows to compute GGC in closed form from the SS parameters of any observed VAR process. Here we show that this approach yields highly accurate spectral estimates of GGC, which closely follow the expected profiles over the coupled directions and have negligible magnitude over the uncoupled direction (Figure 1C); the higher reliability of the SS estimator compared with the standard VAR-based method is evident also looking at single process realizations (Figure 1D).

The second simulation of Stokes and Purdon¹² shows that, due to the independence of GGC measures from the intrinsic dynamics of the “receiver” process, the spectral GGC profiles linking this process to its putatively causal “transmitter” process are often misleading, because different systems can have identical causality functions but different receiver dynamics. In Figure 2 we confirm this result both in terms of GGC and using the DC, a spectral causality measure taken from the VAR framework⁷ that for bivariate processes like the one simulated here is analytically related to the spectral GGC¹⁵. However, this invariance property is in our view absolutely reasonable, because the DC has a clear-cut interpretation as the relative amount of spectral power that, at each frequency, arrives to the receiver starting from the transmitter¹¹. Nevertheless, the DC is also useful to fully recover the functional oscillatory structure of the observed processes, because it shapes the receiver spectrum to reveal the portion of its spectral power that is “causally” due to the transmitter; this is depicted in the spectral decomposition of Figure 2.

In conclusion, while thanking Stokes and Purdon¹² for pointing out some weaknesses of GGC measures, we think that proper formulations can provide meaningful results of directed dynamical influence, whose interpretation still is bound to the knowledge and good faith of those who write and read related scientific literature.

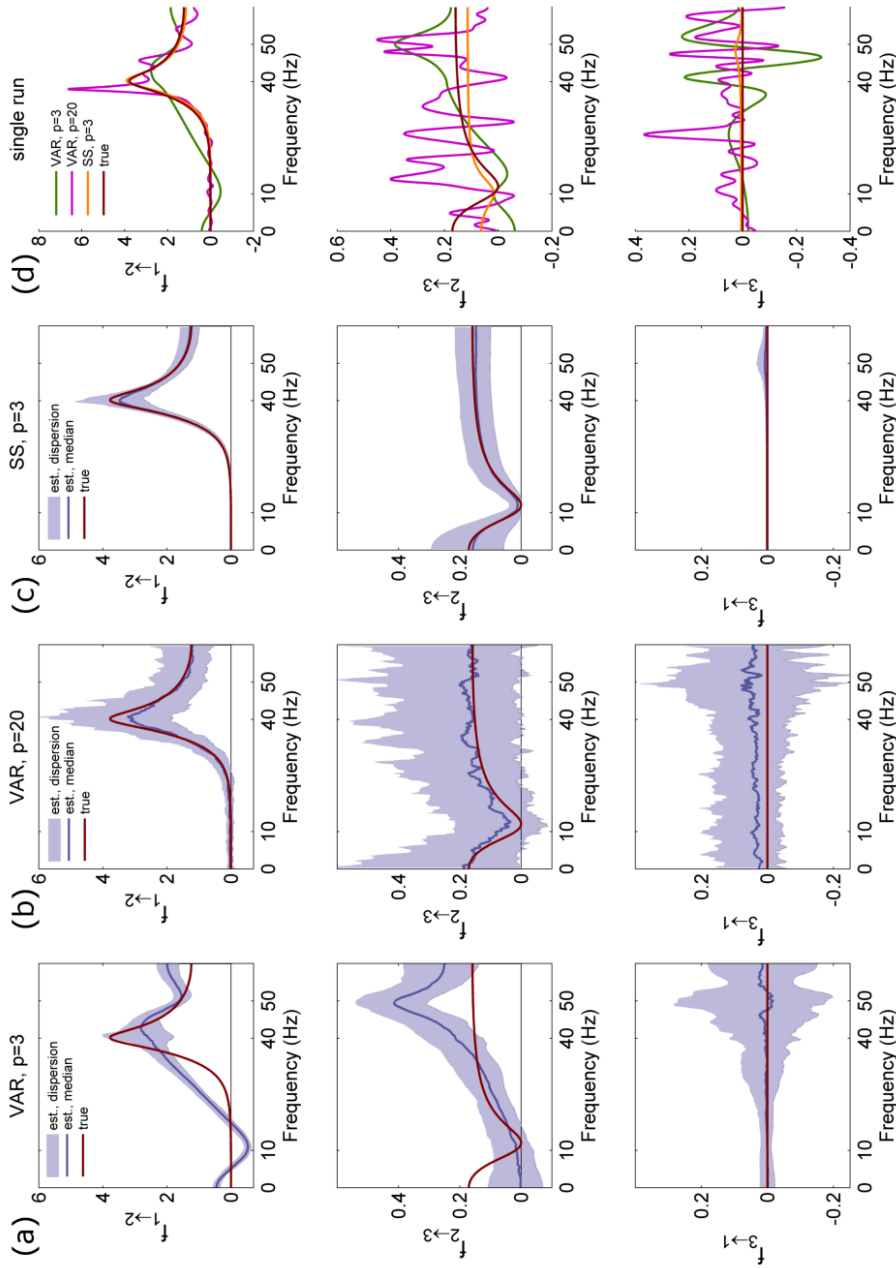


Figure 1. Comparison of conditional frequency-domain Granger-Geweke causality (GGC) profiles computed for the three-node system of Stokes and Purdon¹² (Example 1, where nodes 1, 2, and 3 resonate respectively at 40 Hz, 10 Hz, and 50 Hz, and where unidirectional causality is imposed from node 1 to node 2, and from node 2 to node 3). GGC is computed along the two coupled directions ($f_{1,2}$, $f_{2,3}$) and along a direction with no coupling ($f_{3,1}$). Columns report the distribution of GGC estimates (median and 5th-95th percentiles) computed using classical vector autoregressive (VAR) estimation of full and reduced models performed with the true model order $p=3$ (a) and with an increased order $p=20$ (b), and using state space (SS) estimation (c), as well as estimates obtained for a single simulation run (d); in each plot, the true causality values computed from the original model parameters are reported in red. Results evidence the lower bias and variability of spectral GGC computed using the SS method compared to the classical VAR approach.

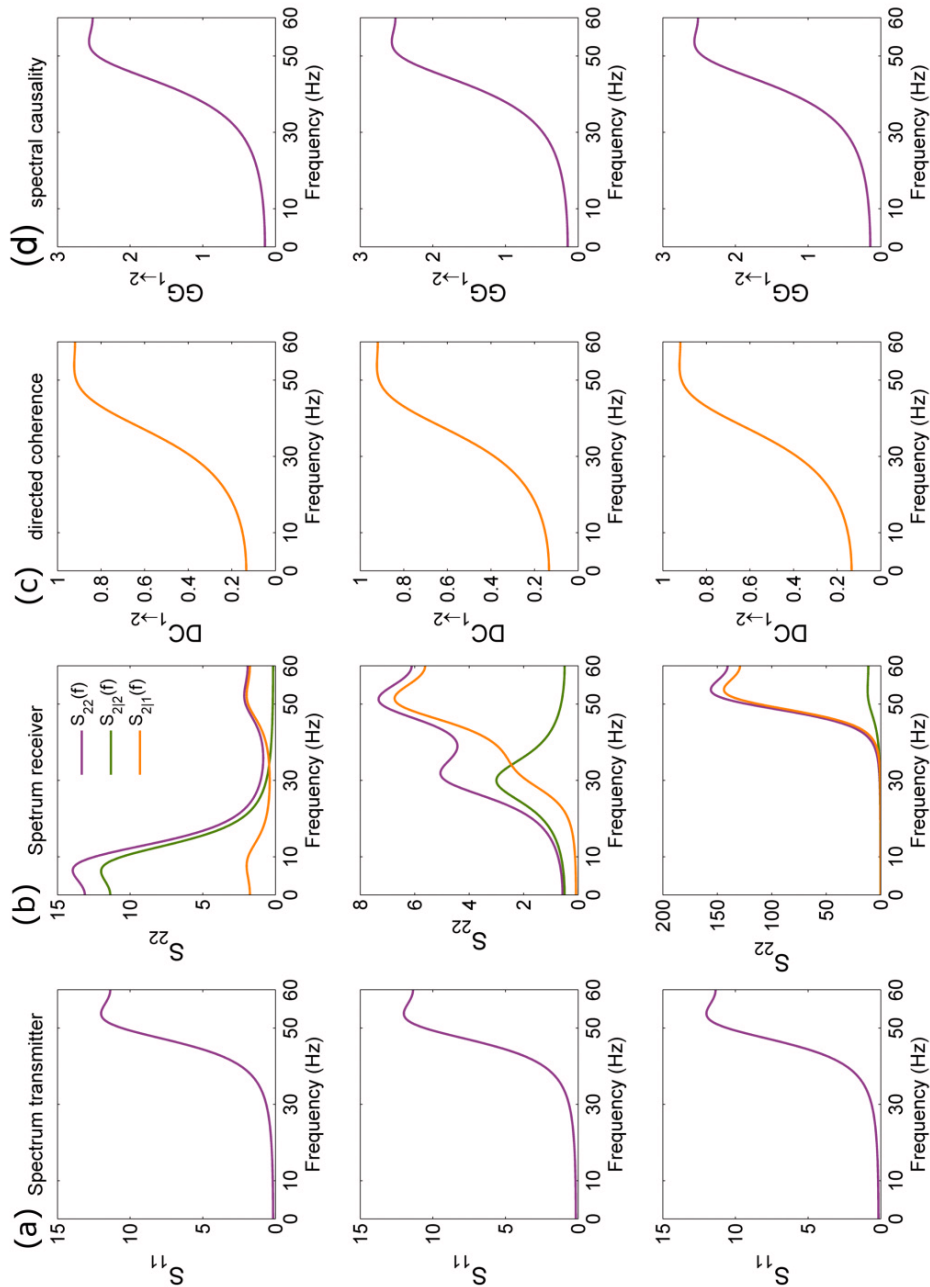


Figure 2. Theoretical profiles of spectral and causality measures computed for the two-node system of Stokes and Purdon¹² (Example 2, where unidirectional causality is imposed from node 1 to node 2). The system is studied setting a resonance frequency of 50 Hz for the transmitter and of 10 Hz (top row panels), 30 Hz (mid row panels) and 50 Hz (bottom row panels) for the receiver. The fact that the power spectral density (PSD) of the transmitter ($S_{11}(f)$, **a**) is the same for the three cases includes, together with the unaltered coefficients determining the causal effects, the same profile for the directed coherence $DC_{1 \rightarrow 2}(f)$ (**c**) and the spectral GGC measure $GG_{1 \rightarrow 2}(f)$ (**d**). However, the different causal contribution of the transmitter on the receiver is revealed by the partial spectrum $S_{21}(f)=S_{22}(f)DC_{1 \rightarrow 2}(f)$ (orange line in **b**), which quantifies the portion of the overall PSD of the receiver ($S_{22}(f)$, purple line in **b**) that is causally explained by the transmitter; the non-explained part ($S_{212}(f)$, green line in **b**) reflects the autonomous dynamics of the receiver.

Dataset 1. Codes to compute frequency-domain Granger causality in linear stochastic processes

<http://dx.doi.org/10.5256/f1000research.12694.d178159>

The package contains Matlab scripts and functions that allow to reproduce the simulations of Stokes and Purdon, and to compare for these simulations: (i) the standard vector autoregressive estimator and the updated state space estimator of the spectral Granger-Geweke causality measure; (ii) the spectral Granger-Geweke causality measure with the directed coherence measure.

Data availability

Dataset 1: Codes to compute frequency-domain Granger causality in linear stochastic processes. The package contains Matlab scripts and functions that allow to reproduce the simulations of Stokes and Purdon, and to compare for these simulations:

(i) the standard vector autoregressive estimator and the updated state space estimator of the spectral Granger-Geweke causality measure; (ii) the spectral Granger-Geweke causality measure with the directed coherence measure.

DOI, [10.5256/f1000research.12694.d178159](https://doi.org/10.5256/f1000research.12694.d178159)¹⁷.

Competing interests

No competing interests were disclosed.

Grant information

The work was funded by the Department of Industrial Engineering, University of Trento, Italy (40300314).

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

- Granger CWJ: **Investigating Causal Relations by Econometric Models and Cross-spectral Methods.** *Econometrica*. 1969; **37**(3): 424–438.
[Publisher Full Text](#)
- Seth AK, Barrett AB, Barnett L: **Granger Causality Analysis in Neuroscience and Neuroimaging.** *J Neurosci*. 2015; **35**(8): 3293–3297.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Porta A, Faes L: **Wiener–Granger causality in network physiology with applications to cardiovascular control and neuroscience.** *Proc IEEE*. 2016; **104**(2): 282–309.
[Publisher Full Text](#)
- Geweke J: **Measurement of Linear Dependence and Feedback between Multiple Time Series.** *J Am Stat Assoc*. 1982; **77**(378): 304–313.
[Publisher Full Text](#)
- Geweke JF: **Measures of Conditional Linear Dependence and Feedback between Time Series.** *J Am Stat Assoc*. 1984; **79**(388): 907–915.
[Publisher Full Text](#)
- Kamiński M, Ding M, Truccolo WA, et al.: **Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance.** *Biol Cybern*. 2001; **85**(2): 145–157.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Baccalá LA, Sameshima K, Ballester G, et al.: **Studying the Interaction Between Brain Structures via Directed Coherence and Granger Causality.** *Appl Signal Process*. 1998; **5**(1): 40–48.
[Reference Source](#)
- Baccalá LA, Sameshima K: **Partial directed coherence: a new concept in neural structure determination.** *Biol Cybern*. 2001; **84**(6): 463–474.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Bressler SL, Seth AK: **Wiener-Granger causality: a well established methodology.** *Neuroimage*. 2011; **58**(2): 323–329.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Chicharro D: **On the spectral formulation of Granger causality.** *Biol Cybern*. 2011; **105**(5–6): 331–347.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Faes L, Erla S, Nollo G: **Measuring connectivity in linear multivariate processes: definitions, interpretation, and practical analysis.** *Comput Math Methods Med*. 2012; **2012**: 18, 140513.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Stokes PA, Purdon PL: **A study of problems encountered in Granger causality analysis from a neuroscience perspective.** *Proc Natl Acad Sci U S A*. 2017; **114**(34): E7063–E7072.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Valdes-Sosa PA, Roebroeck A, Daunizeau J, et al.: **Effective connectivity: influence, causality and biophysical modeling.** *Neuroimage*. 2011; **58**(2): 339–361.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Florin E, Gross J, Pfeifer J, et al.: **Reliability of multivariate causality measures for neural data.** *J Neurosci Methods*. 2011; **198**(2): 344–358.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Faes L, Nollo G: **Measuring frequency domain granger causality for multiple blocks of interacting time series.** *Biol Cybern*. 2013; **107**(2): 217–232.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Barnett L, Seth AK: **Granger causality for state-space models.** *Phys Rev E Stat Nonlin Soft Matter Phys*. 2015; **91**(4): 040101.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Faes L, Stramaglia S, Marinazzo D: **Dataset 1 in: On the interpretability and computational reliability of frequency-domain Granger causality.** *F1000Research*. 2017.
[Data Source](#)

Open Peer Review

Current Referee Status:  

Version 1

Referee Report 06 November 2017

doi:10.5256/f1000research.13747.r26209



Karin Schiecke 

Bernstein Group for Computational Neuroscience Jena, Institute of Medical Statistics, Computer Sciences and Documentation (IMSID), Jena University Hospital, Friedrich Schiller University Jena, Jena, Germany

The correspondence is a comment directly related to “A study of problems encountered in Granger causality analysis from a neuroscience perspective¹. I completely agree with the authors that there are interpretation issues of Granger causality (and by the way also a lot of wrong use/wrong adaptation in the field) and therefore the whole concept has to be applied with care. But – like the authors clearly stated in their comment – Stokes and Purdon uses 1) a version of frequency-domain Granger causality which is at least sub-optimal and 2) the results they were able to show by means of simulated data should be a general problem not only occurring in the case of neuroscience data. Therefore, I explicitly support the comment of Faes, Stramaglia and Marinazzo.

The comment is well written. Basic information concerning the concept of Granger causality and in particular the frequency-domain approach are given (including the coverage of related literature). The main idea behind the paper of Stokes and Purdon and the reason why the authors felt the need to comment on this paper are clearly stated. Then, the simulations of Stokes and Purdon are repeated using an updated Granger causality implementation and a combination of spectral and causal information. Technical details are given, results are correctly presented and well described, and conclusions are drawn.

I do not have any major or minor concerns.

References

1. Stokes PA, Purdon PL: A study of problems encountered in Granger causality analysis from a neuroscience perspective. *Proc Natl Acad Sci U S A*. 2017; **114** (34): E7063-E7072 [PubMed Abstract](#) | [Publisher Full Text](#)

Is the rationale for commenting on the previous publication clearly described?

Yes

Are any opinions stated well-argued, clear and cogent?

Yes

Are arguments sufficiently supported by evidence from the published literature or by new data and results?

Yes

Is the conclusion balanced and justified on the basis of the presented arguments?

Yes

Competing Interests: No competing interests were disclosed.

Referee Expertise: computational neuroscience, complex systems, nonlinear dynamics, connectivity analysis

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Referee Report 02 November 2017

doi:10.5256/f1000research.13747.r26204



Luiz Antonio Baccala

Department of Telecommunications and Control Engineering, Polytechnic School, University of São Paulo, São Paulo, Brazil

I am glad that Faes et al. have not fallen prey to these considerations and have mounted a fair and elegant, albeit brief, critical appraisal of the PNAS Stokes and Purdon (S/P) paper. Faes et al. have restricted themselves to nailing just three obvious crucial facts that may go unnoticed by the casual reader for whom the authoritativeness of a vehicle like PNAS may stand as a certificate of validity.

The first point is that the S/P paper fails to mention how their VAR models were computed. My own AsympPDC package (The AsympPDC Package 3.0 is directly downloadable from http://www.lcs.poli.usp.br/~baccala/pdc/asymp_package_v3.zip. A preliminary version is also available through [7]. Visit <http://www.lcs.poli.usp.br/~baccala/pdc/> for future version updates.) provides five different methods: the simplest naive and most popular approach, the least squares method, is the worst performer thanks to error propagation alone. It is important to stress that accurate VAR model estimation is crucial to whichever approach to GC is chosen. Faes et al. correctly supply an alternative method of estimation where no theoretically meaningless 'negative' GGC value is observed. The Faes et al. paper has the added obvious merit of including the actual value of GGC computed from the actual true model absentmindedly lacking in the original S/P paper.

The second point raised almost in passing by Faes et al. is that the S/P paper may possibly induce its readers to completely disregard frequency domain causality descriptions while S/P mostly glosses over the alternative DC/PDC framework. This is a huge oversight since currently only DC/PDC estimators have statistical theoretically rigorous computable confidence intervals and objective null hypothesis threshold tests¹⁻² which may be applied using the freely downloadable AsympPDC package. This fact alone sets the DC/PDC methodologically apart. In today's internet era of information no more than a google away, this omission is unforgivable.

The third point concerns the two time series case of Example 2. Indeed here is a point that Faes et al

correctly argue that S/P fail to grasp. GC was conceived by Granger ³ to decompose pairwise relationships into exposing factors that aid predictability. This was later shown to be equivalent to detecting and characterizing the presence of feedback by Sims ⁴. In this case, GGC, DC and PDC address the so called connection 'detection' problem by focusing on the connection. If some given influence is present and if an influenced subsystem is affected, how it responds resonantly or otherwise is its own business.

Do the valid Faes et al. criticisms imply the S/P paper is worthless? Despite its many additional shortcomings, the S/P paper has the important merit of stripping bare some of the field's reigning confusion, enough to call for added discussion. This only stresses the relevance of the present criticisms and the urgent need for clarification.

Finally I would like to put forward some thoughts that may explain the present state of conceptual disarray regarding causality. The first issue is ignorance about time series estimation — of the 'a little knowledge is dangerous thing' kind. One cannot expect correct and reasonable results by just downloading some package, pressing some buttons without knowing the precise limitations of each available tool - time series analysis still has some elements of an art. The second problem has to do with the notions of Causality - it is tempting to ask the methodology to provide a glimpse on actual mechanisms. Sometimes, on physical grounds, this desire may be fulfilled, but true (mechanical) causality determination requires observer intervention ⁵⁻⁶; what GC does is to allow the exclusion of some tentative mechanisms. Thirdly, even when it reflects actual mechanisms, but more than two time series are simultaneously examined, different descriptions, as is the case for the complementarity between DC and PDC, they reflect different system properties so that one must break the GC concept into more general ideas: Granger Connectivity and Granger Influentiability ⁵⁻⁶. Last but not least, popular descriptions of large scale connectivity sometimes qualified as 'effective' versus 'functional' have further added to the present state of confusion due to their shear inconsistent application throughout the literature (see a discussion in ⁵⁻⁷).

I think we should thank Faes et al. for pointing out some of these problems. Perhaps this is a good opportunity to start clearing up these issues once and for all.

References

1. Baccalá LA, de Brito CS, Takahashi DY, Sameshima K: Unified asymptotic theory for all partial directed coherence forms. *Philos Trans A Math Phys Eng Sci*. 2013; **371** (1997): 20120158 [PubMed Abstract](#) | [Publisher Full Text](#)
2. Baccala LA, Takahashi DY, Sameshima K: Directed Transfer Function: Unified Asymptotic Theory and Some of Its Implications. *IEEE Trans Biomed Eng*. 2016; **63** (12): 2450-2460 [PubMed Abstract](#) | [Publisher Full Text](#)
3. Granger C: Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*. 1969; **37** (3). [Publisher Full Text](#)
4. Sims, C: Money, income, and causality. *The American Economic Review*. 1972. 540-552
5. Baccalá L, Sameshima K: Causality and Influentiability: The Need for Distinct Neural Connectivity Concepts. 2014; **8609**: 424-435 [Publisher Full Text](#)
6. Sameshima K, Baccala LA: Methods in Brain Connectivity Inference through Multivariate Time Series Analysis. *CRC Press*. 2014.
7. Baccalá L. A, Sameshima K: A neural link centered approach to brain connectivity. *Cognitive Science: Recent Advances and Recurring Problems, Vernon Press*. 2017. 101-112

Is the rationale for commenting on the previous publication clearly described?

Yes

Are any opinions stated well-argued, clear and cogent?

Yes

Are arguments sufficiently supported by evidence from the published literature or by new data and results?

Yes

Is the conclusion balanced and justified on the basis of the presented arguments?

Yes

Competing Interests: No competing interests were disclosed.

Referee Expertise: I work in statistical signal analysis and estimation with interest in applications to multivariate time series contexts as applied to neuroscience and acoustical imaging applications. I have been co-proposer of partial directed coherence (PDC) mentioned in the text.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research