Hindawi
Computational and Mathematical Methods in Medicine
Volume 2022, Article ID 6962359, 15 pages
https://doi.org/10.1155/2022/6962359

Research Article

Investigating Large-Scale Network with Unified Granger Causality Analysis

Zhenghui Hu, Fei Li, Minjia Cheng, and Qiang Lin

Key Laboratory of Quantum Precision Measurement, College of Science, Zhejiang University of Technology, Hangzhou 310023, China

Correspondence should be addressed to Qiang Lin; qlin@zjut.edu.cn

Received 6 December 2021; Accepted 22 January 2022; Published 2 March 2022

Academic Editor: Luca Faes

As the concept of integrating global neuron coupling effect is increasingly accepted, investigating causal connection increasingly requires the intervention of large-scale analysis. In this study, a large-scale brain network analysis was carried out by a description length guided framework, which involves a unified Granger causality analysis (uGCA) method and now integrates the concept of large-scale analysis. This will be helpful to make a more comprehensive determination for causal connection among the global brain regions. Distinct from the conventional GCA, which involves a two-stage scheme consisting of Akaike information criterion or Bayesian information criterion (AIC/BIC) and F-test to obtain a causal effect, a unified guided framework can ensure more reliable results while eliminating some confounding influences among network nodes. Then, we performed large-scale network simulation experiments involving 13 nodes; it was found that our proposal was more accurate and robust in guiding the causal connection investigation of large-scale networks. When it comes to the resting-state fMRI datasets, we studied a 90-node network selected from the Anatomical Automatic Labeling (AAL) template. Then, combining a K-means clustering method, we found that most brain nodes in the connection network obtained by uGCA methods were gathered into the corresponding functional brain regions and functionally related regions cooperated with each other. Compared to conventional GCA, their results were more consistent with clinical and anatomical priors. Moreover, in studies of several large-scale functional networks involving default mode network (DMN), dorsal attention network (DAN), and frontoparietal control network (FCN), the uGCA method more clearly revealed their empirical cooperation. As a brain with numerous nodes and massive connections, a unified large-scale analysis method is of great significance for the integration of causal connections in the whole brain network in the future.

1. Introduction

With the rise of a notion that the brain works as a union of complex neural circuits (functional integration) at different spatial scales, more information should be taken into account in describing brain region couplings; thus, more attention should be paid to investigating causal connection among brain regions. Moreover, dynamic coupling and synchronous oscillation of a large number of neurons exist in the brain; thus, aliasing and chaos may occur in some causal connectivities between neuron clusters at the space-time level. As a data-driven procedure, Granger causality analysis (GCA) does not require a biophysical model for investigating causal connections. In such a huge and complex system (the brain), GCA thus may be admirable to investigate these causal couplings. Meanwhile, due to the presence of intermediary external and potential nodes in the real-world, indirect connections and synchronous instantaneous connections will be hard to eliminate. Therefore, in order to describe the internal coupling between brain regions more unambiguously and accurately, investigating causal connection increasingly requires the intervention of large-scale analysis [1–3].

In conventional GCA, many efforts and achievements had been made, for example, conditional GCA, partial GCA, kernel GCA, and Geweke-GCA [4–7]. However, all these

developments still remain the framework of original GCA; it is actually a two-stage scheme that uses AIC/BIC to determine candidate models and then establishes causal connection through the F-test. Therefore, we collectively refer to them as conventional GCA. But in a purely mathematical sense, these two stages are generalized model selection issues; straight joint of two different mathematical theories will cause inherent relation to be discontinuous in the quantitative modeling process, namely, bringing in singularity. Meanwhile, a subjective selection of confidence level is one of the stages of conventional GCA, which will lead to a lack of uniformity of research results and bring in some performance issues. Another problem brought about by the F-test is a pairwise comparison that candidate models need to compare with each other through an intermediate model, which will increase algorithm complexity. And selection results by pairwise F -statistics sometimes depend on the initial targeted model and search path heavily, especially in a large-scale network. It is worth noting that selecting and using specific F-statistics has become very careful in current scientific research, and its statistical significance has also caused extensive discussion [8–11]. In general, some inherent issues still remain to be dealt with in conventional GCA which are the inconsistency of mathematical theories, the subjective selection of confidence level, and the algorithm complexity caused by a nested model. To alter the conventional GCA framework, thus we proposed a unified model selection approach for GCA based on the minimum description length (MDL) principle, called uGCA, and we had demonstrated its effectiveness and priority over the conventional GCA in our previous studies [12, 13].

Recur to the novel uGCA framework [12], which overcomes some inherent drawbacks of the conventional GCA; further, large-scale network analysis should be absorbed to improve it. Compared to conventional GCA, uGCA unifies these two generalized model selection issues into a description length guided framework. Specifically, it can integrate all candidate data into the same framework, so the established data model will be under the same context and can be used for large-scale network analysis more directly and effectively. This unified methodological framework is consistent with existing scientific theories and experiments, which will bring some advantages for future experimental research on mutiscale. As is well known, there are still many uncertainties in the study of causal connectivities in brain regions, we should try to avoid introducing new uncertain factors in the process of data processing and modeling, such as the definition of brain nodes is ambiguous, and causal connections of brain networks may exist on different time scales. Therefore, adhering to the principle of parsimony, our uGCA framework intends to integrate the whole process of causal investigating together and automatically returns to the most suitable descriptive model. Such a unified framework will be suitable for causal investigation on large-scale networks, which can deal with indirect or spurious connectivities more easily. In this paper, we will focus on the direct comparisons in synthetic data and fMRI data experiments between uGCA and conventional GCA to illustrate the priority of the uGCA to conventional GCA in large-scale network analysis.

The rest of this article is organized as follows. Firstly, a large-scale network investigation procedure using the uGCA framework has been stated. Then, we illustrate the advantages of several uGCA forms over the conventional two-stage GCA in a 13-node synthetic data experiment; results showed that most causal networks identified by uGCA were closer to its ground truth. Then, in resting-state fMRI datasets, direct comparisons between the uGCA and conventional GCA were carried out, which involved identifying large-scale causal connections in 90 region networks of the AAL template and several intrinsic functional networks, respectively. By contrast, causal connections identified by uGCA showed some clustering networks which were consistent with existing clinical and anatomical experience, especially in three large-scale functional networks. At last, the corresponding explanations are presented and we demonstrated the comparison between conventional two-stage GCA and our proposal.

2. Description Length and Unified Granger Causality Analysis

In the conventional GCA, it involved a two-stage scheme which actually both are generalized model selection issues. Therefore, inspired by coding theory, we considered a novel causal investigation method based on coding the candidate model, which can describe data models more succinctly. With the help of the MDL principle, which provides a generic solution for the model selection issue [14–17] and regards the probability distribution as a descriptive standpoint to choose the model with the shortest description of data, we propose a unified description length guided GCA method, namely, uGCA. Compared with the two-stage scheme of conventional GCA, the uGCA method fixes attention on unifying the model complexity term and error term into a unified description length guided framework then to determinate causal connection with their description length, in which it can avoid the inconsistency of applying several mathematical theories, the subjectivity of selecting confidence levels, and pairwise comparison of nested models. Essentially, uGCA framework still retained the foundation of the original GCA, Y Granger-cause X when Y provides predicted information for X; the difference is that uGCA altered two different mathematical theory procedures (AIC/BIC and F-test) of GCA into a single theory (description length guided framework with help of MDL). In this way, uGCA is essentially a more rigorous version than the conventional GCA.

2.1. Description Length Guided Causal Investigation. Firstly, considering two variables, X_N and Y_N, the description models associated with X_N

$$\begin{cases} X_t = \sum_{j=1}^{n1} a_{1i} X_{t-j} + \varepsilon_{1t}, \\ X_t = \sum_{j=1}^{n2} a_{2i} X_{t-j} + \sum_{j=1}^{n3} b_{2i} Y_{t-j} + \varepsilon_{2t}, \end{cases} \tag{1}$$

where ε_t is fitting residual. Distilling the sense of Granger causality, causal influence from Y to X within uGCA framework is defined by, that is, causal identification between two variables:

$$F_{Y\longrightarrow X} = L_X - L_{X+Y}, \tag{2}$$

where L_X denotes the shortest description length of restricted model in Equation (1), and L_{X+Y} denotes the shortest description length of unrestricted model in Equation (1) after adding Y_N. Causal influence from Y to X existed when $F_{Y\longrightarrow X} > 0$, or else, there is no causal effect from Y to X.

Accessing the concept of conditional Granger causality, the influence from Y to X can be identified while controlling the effect from another node Z to X. This joint representation is

$$\begin{cases} X_t = \sum_{j=1}^{m1} a_{3i} X_{t-j} + \sum_{j=1}^{m2} b_{3i} Z_{t-j} + \varepsilon_{3t}, \\ X_t = \sum_{j=1}^{m3} a_{4i} X_{t-j} + \sum_{j=1}^{m4} b_{4i} Y_{t-j} + \sum_{j=1}^{m5} c_{4i} Z_{t-j} + \varepsilon_{4t}. \end{cases} \tag{3}$$

Thus, if $F_{Y\longrightarrow X} > 0$, causal influence from Y to X conditioned Z is given by description length in (3):

$$F_{Y\longrightarrow X|Z} = L_{X+Z} - L_{X+Y+Z}, \tag{4}$$

where L_{X+Z} denotes the shortest length of description model in Equation (3) after joining Z_N. And L_{X+Y+Z} is the optimal length of description model in Equation (3) after joining X_N and Z_N. Same as above, causal influence from Z to X conditioned Y is

$$F_{Z\longrightarrow X|Y} = L_{X+Y} - L_{X+Y+Z}. \tag{5}$$

In this case, if both $F_{Y\longrightarrow X} > 0$ and $F_{Z\longrightarrow X} > 0$ exist, conventional GCA identifies causal effects by pairwise comparison of nested models, which has been illustrated in our previous study [12]. Intuitively, in this unified framework, all candidate models can be described as description length and then compared freely. Unlike the nested models in the conventional GCA scheme, uGCA can release the algorithm complexity due to freely choosing the number of comparison models. Thus, if both $F_{Y\longrightarrow X} > 0$ and $F_{Z\longrightarrow X} > 0$ exist,

$$F_{Y,Z\longrightarrow X} = \min\left(L_{X+Y}, L_{X+Z}\right) - L_{X+Y+Z}. \tag{6}$$

If $F_{Y,Z\longrightarrow X} > 0$, it means that both Y and Z have a direct effect on X. But there will be two cases to be dealt with when $F_{Y,Z\longrightarrow X} < 0$. One is that $F_{Y,Z\longrightarrow X} = (L_{X+Y} - L_{X+Y+Z}) < 0$ existed, it indicates Y impacts on X directly, and Z has an indirect causal effect on X. The other is $F_{Y,Z\longrightarrow X} = (L_{X+Z} - L_{X+Y+Z}) < 0$; it implies a direct causal effect from Z to X.

In general, uGCA provides the optimal description length of parametric model and then identify causal effect through them, which the whole causal investigation procedure only involves the description length. That is, the optimal description length only involves modeling the parametric model of candidate variables, like AIC/BIC. As in (6), their optimal description lengths can be compared when multipathways to the same node that exist, not like that conventional causal effect mediated by the pairwise nested models, in which F-statistics cannot be compared directly. Consequently, uGCA methods conveying a unified framework are more concise and have mathematical rigor, which is more in line with Occam's razor, or the principle of *parsimony*.

2.2. Large-Scale Network Analysis. Extending the large-scale network analysis into our uGCA framework to capture more information between brain regions coupling, more related nodes should be accommodated into causal investigation. Figure 1 shows the main process of large-scale network analysis: a facilitated classification for the target node associated network which contains target **T** and all its ancestors. All indirect connections can be grouped into two kinds of spurious connectivities that one sort of node **X** has only a single pathway to target **T**, and the other sort of node **Y** has more than two distinct pathways. In Figure 1(a), we see that direct parents of target **T** are P_1, P_2, P_3, and P_4.

First, through the bivariate uGCA procedure, we can obtain all the direct and indirect parent nodes of the target node **T**. In Figure 1(a), these nodes were grouped into a set $\mathbf{A_0} = (P_1 \cdots P_4 ; X_1 \cdots X_n ; Y_1 \cdots ; Y_n)$. Second, to further determine whether these connectivities obtained above were direct or indirect, an ergodic process was executed. Considering the indirect connectivities mediated by a single pathway, which just contains three nodes, we can eliminate them by accommodating the conditional uGCA. Theoretically, all single pathway nodes from set $\mathbf{A_0}$ were removed. Thus, the remaining set only contained direct parent node class **P** and indirect node class **Y** associated with several pathways, named $\mathbf{A_1} = (P_1 \cdots P_4 ; Y_1 \cdots ; Y_n)$, shown in Figure 1(b). For each node in $\mathbf{A_1}$, they impacted target node **T** through several direct parent nodes. Similarly, we eliminated these indirect connectivities by mediating several nodes, rather than a single node in the conditional procedure. In this process, the scale of conditioned nodes was selected by an ergodic process of the global network. Until now, we removed all indirect connectivities that were conditioned on several direct parent nodes, and then, we obtained a sparse network that only contained direct nodes to target node **T**. In Figure 1(c), node class **Y** is further deleted from $\mathbf{A_1}$, showing the remaining set as $\mathbf{A_2} = (T_1 \cdots . T_4)$.

In our uGCA guided large-scale network analysis, it was not restricted by pairwise comparison, which meant several different pathway selections can be compared freely. In this unified guided framework, description models for dataset were mapped into a unified space which guaranteed coherence of processing benchmark. Meanwhile, when faced with multiple candidate models for comparison, our uGCA framework showed advantages in a succinct modeling way, which can automatically regress to an optimal model and determine the causal connection [12].

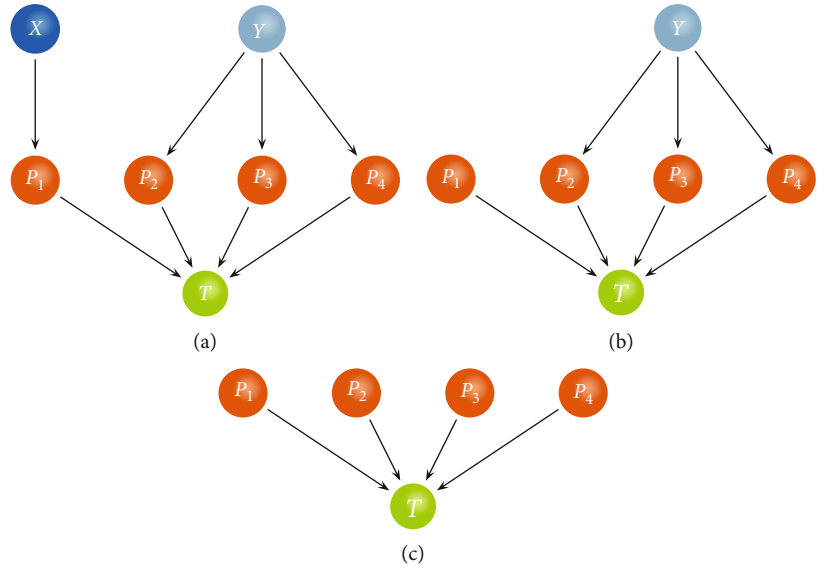


Figure 1: Large-scale network analysis procedure within uGCA. (a) All the direct and indirect parent nodes of the target node **T**. (b) Direct nodes **P** and indirect nodes **Y** with several pathways. (c) Only contains direct nodes **P**.

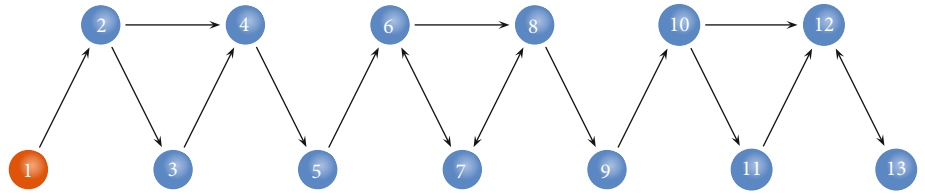


Figure 2: The causal connectivities between 13 nodes; the red one is the driving node.

2.3. Different Forms of Description Length Model in uGCA. The following is a causal investigation of which different forms of uGCA guided for linear AR model. Particularly, with the help of genetic solution of model selection in MDL, the shortest length of parametric model in uGCA (that is, the L in Equations (2) and (6)) is carried out. Variable $x^n = \{x_1, \cdots, x_n\}$ is given:

$$x_t = \beta_1 x_{t-1} + \beta_2 x_{t-2} + \cdots + \beta_k x_{t-k} + \varepsilon_t, \tag{7}$$

where $t = 1, \cdots, m$, and m is more than k to keep the solution determined. For describing x_t, it arrives at

$$f(x^n \mid x_t, \beta, \tau) = \frac{1}{(2\pi\tau)^{m/2}} e^{-(1/2\tau)\sum_t \left(x_t - \sum_k \beta_k x_{t-k}\right)}. \tag{8}$$

2.3.1. uGCA-TP: Crude Two-Part Coding Scheme. In uGCA-TP, this two-part coding scheme divided the descriptive model into a fitting error term and a parameter literal coding term [18]. Its parameter vector consists of data $\theta = (k, \xi)$ and $\xi = (\tau, \beta_1, \cdots, \beta_k)$, where $\xi \in R^{k+1}$, $\tau = \xi_0$ is the variance-parameter of zero-mean Gaussian distribution model for ε_t. Let RSS denote the residual sum of squares in the parameter estimation. Then, this description length is given as

$$L_{\text{uGCA-TP}} = m \ln \sqrt{2\pi\tau} + \frac{\text{RSS}}{2\tau} + \sum_{i=0}^{k} \ln \frac{|\xi_i|}{\delta} + \ln (k+1), \tag{9}$$

where δ is the precision, and it is optimal to choose $1/\sqrt{N}$ [16, 19, 20]. Particularly, $|\xi_i| / \delta < 1$ should be ignored.

2.3.2. uGCA-MIX: g-Prior for Parameter Space. For the uGCA-MIX, it provides some priors for describing parameter space, which is a mixture form represented as a member of the natural conjugate family of priors for an ordinary linear regression model, namely, *normal inverse-gamma* distributions [20]:

$$\omega(\beta, \tau) \propto \tau^{(-d+k+2)/2} e^{-(\beta-b)' c\Sigma(\beta-b)+a/2\tau}, \tag{10}$$

where $\Sigma = X'_k X_k = mS$ $(X'_k = \{x_{i,t-k}\})$ is a $k \times m$ matrix defined by the values of regressor variables [21]. [20] provided a special solution that $a = d = 0, b = (0, .., 0)$, and [22] christened a specification the g-prior. The value of Σ provided a closed-form expression for $\hat{c}$ in [16], namely, $1/\hat{c} = \max(F - 1, 0)$ where $F = (m-k)(X_t' X_t - \text{RSS})/(k \cdot \text{RSS})$. Thus, R^2 is the usual squared multiple correlation coefficient; the mixture form is given:

$$\mathrm{L}_{\mathrm{uGCA-MIX}} = \begin{cases} \dfrac{m}{2}\ln\dfrac{RSS}{m-k} + \dfrac{k}{2}\ln F + \ln m, & \text{if } R^2 \leqslant \dfrac{k}{n}, \\ \dfrac{m}{2}\log\dfrac{X_t{}' X_t}{m} + \dfrac{1}{2}\log m, & \text{otherwise.} \end{cases} \tag{11}$$

Finally, a simple approximation to this form is applied to derive the Stochastic Information Complexity (SIC) [16]:

$$\mathrm{SIC} = \frac{m-k-2}{2}\log\mathrm{RSS} + \frac{k}{2}\log m + \frac{1}{2}\log\det[\Sigma]. \tag{12}$$

In this context, mixture form adapts to behave like Bayesian model selection.

2.3.3. uGCA-NML: Minimax Solution for Inherent Redundancy. In the earlier two-part code scheme, it remains some inherent redundancy. Combining Fisher information to remove it, a sharper description length with stochastic complexity and universal process is derived for a class of parametric processes [23]. This description form is motivated by the maximum-likelihood estimate (MLE) which requires satisfying the central limit theorem [21, 24]. Thus, the nonintegrability of MLE is a key issue to be solved. Firstly, Fisher information is given by

$$|I(\beta,\tau)| = \frac{|S|}{2\tau^{k+2}}, \tag{13}$$

and the integral of its square root dealt by [21, 23, 24] is

$$\int_{\beta S\beta\leq R}\int_{r_0}^{\infty} |I(\beta,\tau)|^{1/2} d\tau d\beta = (2\mid S\mid)^{1/2}\left(\frac{R}{\tau_0}\right)^{k/2}\frac{V_k}{k}, \tag{14}$$

where $V_k R^{k/2} = 2(\pi R)^{k/2}/(k\sqrt{|S|}\,\Gamma(k/2))$ is the volume of a k-dimensional ball $B = \{\beta' S\beta \leq R\}$. Lower bound τ_0 is determined by precision of data written; $\widehat{\tau}_0 = \mathrm{RSS}/m$ and $\widehat{R} = (\widehat{\beta}' X'_{t-k} X_{t-k}\widehat{\beta})/m$ are given by MLE. Thus, description length in uGCA-NML arrives at

$$\mathrm{L}_{\mathrm{uGCA-NML}} = m\ln\sqrt{2\pi\tau} + \frac{\mathrm{RSS}}{2\tau} + \frac{k}{2}\ln\frac{m}{2} - \log\Gamma\left(\frac{k}{2}\right) + \frac{k}{2}\log\frac{\widehat{R}}{\tau_0} - 2\log k. \tag{15}$$

3. Experiment

3.1. Large-Scale Network Simulation. To verify the effectiveness of our uGCA method, we considered synthetic data experiments which contained 13 nodes in the network, seen in Figure 2. And this network was given by

$$\begin{cases} x(1,i) = 0.89x(1,i-1) - 0.29x(1,i-2) + \varepsilon_1, \\ x(2,i) = 0.79x(1,i-1) - 0.27x(1,i-2) + 0.88x(2,i-1) \\ -0.37x(2,i-2) + \varepsilon_2, \\ x(3,i) = 0.87x(2,i-1) - 0.37x(2,i-2) + 0.83x(3,i-1) \\ -0.28x(3,i-2) + \varepsilon_3, \\ x(4,i) = 0.75x(2,i-1) - 0.36x(2,i-2) + 0.73x(3,i-1) \\ -0.29x(3,i-2) + 0.82x(4,i-1) - 0.28x(4,i-1) + \varepsilon_4, \\ x(5,i) = 0.805x(4,i-1) - 0.35x(4,i-2) + 0.81x(5,i-1) \\ -0.37x(5,i-2) + \varepsilon_5, \\ x(6,i) = 0.65x(5,i-1) - 0.31x(5,i-2) + 0.77x(6,i-1) \\ -0.32x(6,i-2) + 0.71x(7,i-1) - 0.35x(7,i-2) + \varepsilon_6, \\ x(7,i) = 0.68x(6,i-1) - 0.34x(6,i-2) + 0.77x(7,i-1) \\ -0.31x(7,i-2) + 0.62x(8,i-1) - 0.36x(8,i-2) + \varepsilon_7, \\ x(8,i) = 0.66x(6,i-1) - 0.39x(6,i-2) + 0.64x(7,i-1) \\ -0.38x(7,i-2) + 0.69x(8,i-1) - 0.35x(8,i-2) + \varepsilon_8, \\ x(9,i) = 0.68x(8,i-1) - 0.25x(8,i-2) + 0.76x(9,i-1) \\ -0.37x(9,i-2) + \varepsilon_9, \\ x(10,i) = 0.71x(9,i-1) - 0.31x(9,i-2) + 0.75x(10,i-1) \\ -0.35x(10,i-2) + \varepsilon_{10}, \\ x(11,i) = 0.6(x(10,i-1) - 0.40x(10,i-2) + 0.78x(11,i-1) \\ -0.32x(11,i-2) + \varepsilon_{11}, \\ x(12,i) = 0.65x(10,i-1) - 0.38x(10,i-2) + 0.69x(11,i-1) \\ -0.36x(11,i-2) + 0.75x(12,i-1) - 0.33x(12,i-2) \\ \quad - 0.74x(13,i-1) + 0.38x(13,i-2) + \varepsilon_{12}, \\ x(13,i) = 0.67x(12,i-1) - 0.37x(12,i-2) + 0.73x(13,i-1) \\ -0.28x(13,i-2) + \varepsilon_{13}, \end{cases} \tag{16}$$

where ε_i denotes the noise terms. To ensure the effectiveness, synthetic data should be performed stationarity analysis and passed before being allowed to be further used. Then, firstly, it is necessary to test the anti-interference ability of uGCA against different noises in a large-scale network; thus, noise terms were ranged from 0.2 to 0.6.

Figure 3 illustrates causal networks obtained by several uGCA forms and conventional GCA. Obviously, the noise terms had few effects on all methods. For true connections (the real connection edges in Figure 2), both several uGCA forms and conventional GCA have an admirable property. But as shown in our previous research [13], uGCA-MIX had more chances of these false negatives in low noise level because of introducing some priors on estimated parameter distribution, as can be seen from the connection edge 1⟶2. Then, for the other two forms, uGCA-TP and uGCA-NML had a relatively stable performance in the true positive rate (TPR). But we found that it was not accurate enough to

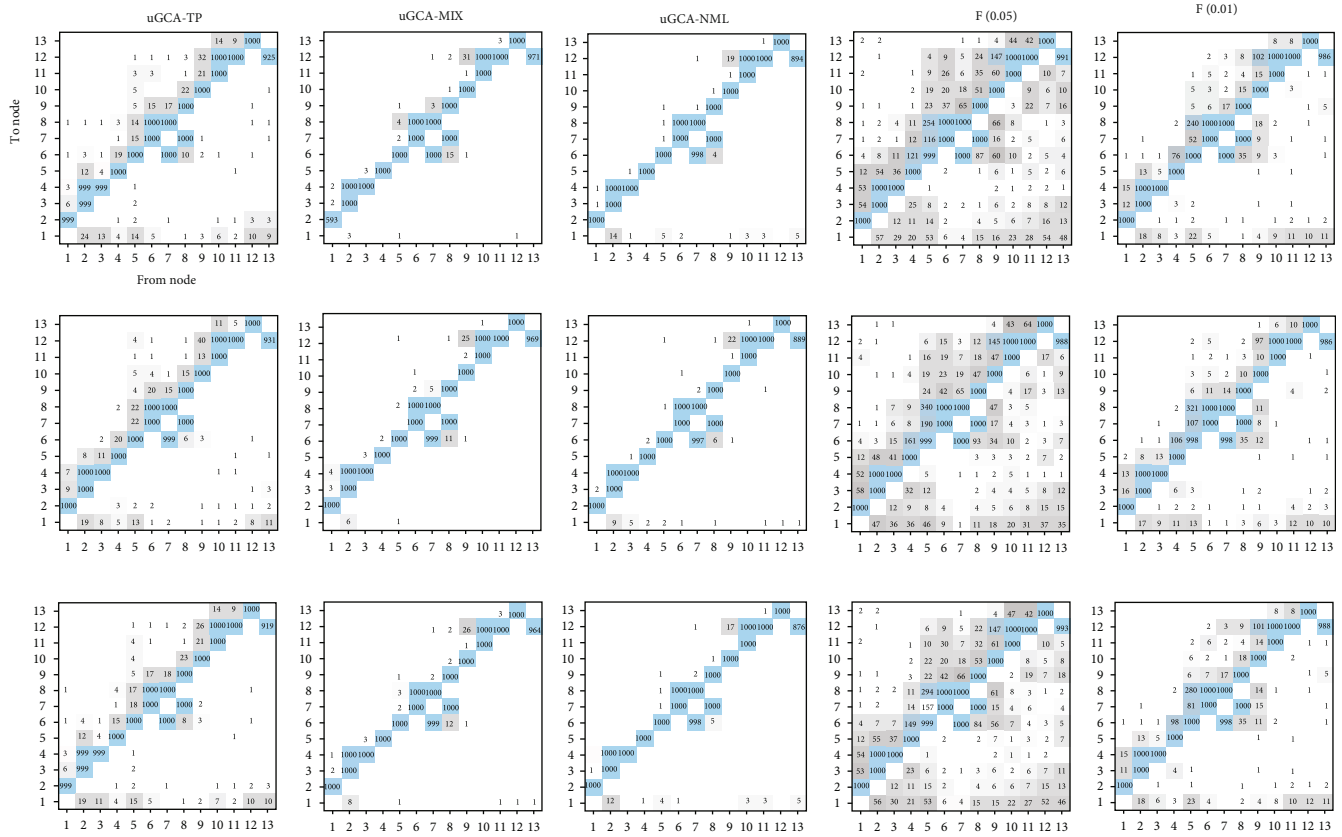


Figure 3: Causal connectivities obtained by several uGCA forms and conventional GCA. The top row represented results in low noise level (var = 0.2); the middle was middle noise level (var = 0.4); the bottom denoted high noise level (var = 0.6).

identify the causal connection 13⟶12; none of the methods can ensure a high TPR. As for false connectivities (the nonexistent connection edges in Figure 2), the advantages of uGCA methods have emerged distinctly. Specifically, uGCA-MIX and uGCA-NML had significantly higher true negative rates (TNRs). Even for uGCA-TP, its false positives were also stifled at a low level. However, poor identification in false connections was obvious for conventional GCA, at least for these two confidence levels in the experiment. Particularly for 4⟶6, 5⟶7, 5⟶8, and 9⟶12, it had quite a few false positives whatever its confidence level. Although results showed that increasing confidence level improved its TNR, the subjectivity of confidence level selection still needs to be solved. In other words, the ground truth is given in a synthetic data experiment, but in real data, its prior knowledge is usually absent, which leads to the lack of a uniform standard to choose a confidence level. In general, all methods had a good anti-interference ability for noise. The uGCA methods can identify true connections with a high TPR while ensuring high TNR to eliminate false connections, but conventional GCA can not guarantee a high TNR.

To further confirm the validity of the uGCA method in large-scale network analysis, we varied data length from 200 to 1000. Several uGCA forms and conventional GCA showed their own characteristic performance seen in Figure 4. Unlike varying the noise level, all methods were affected by varying data lengths and had different performances. For conventional GCA, it identified all true positives with high accuracy when data length was above 500. When data length increased from 200 to 1000, conventional GCA eliminated some spurious connections but led to an increase in some false positives. For example, causal effects of 4⟶6, 5⟶7, and 9⟶12 were eliminated to a low level while the rate of false positives of 5⟶8, 9,10,11⟶8 increased to a significant level, especially for 5⟶8. And lots of false connections can not be fully eliminated even varying data lengths from 200 to 1000. Distinct from conventional GCA, the uGCA method maintained a positive correlation with data length, although several forms had slight differences. For uGCA-TP form, except for the connection edge of 13⟶12, it ensured very high TPRs when data length was 300. Then, varying data length to 500, all true positives were fully identified. And uGCA-TP can also eliminate some false positives as its data length increased. As for uGCA-MIX, it achieved higher accuracy in identifying true positives within a shorter data length than uGCA-TP. Meanwhile, uGCA-MIX stifled false positives to a very low level; thus, it obtained a very sparse connection network. When data length was 1000, uGCA-MIX almost identified a connection network near to the ground truth (the synthetic network in Figure 2) for every synthetic data sample. Same as uGCA-MIX, uGCA-NML almost obtained a "ground truth network" in every synthetic sample when the data length was 1000. As for identifying false positives, uGCA-NML seemed to be more admirable than uGCA-MIX; it had the highest accuracy in eliminating these spurious connections. Similarly, except for the connection edge 13⟶12, it all acquired high accuracy

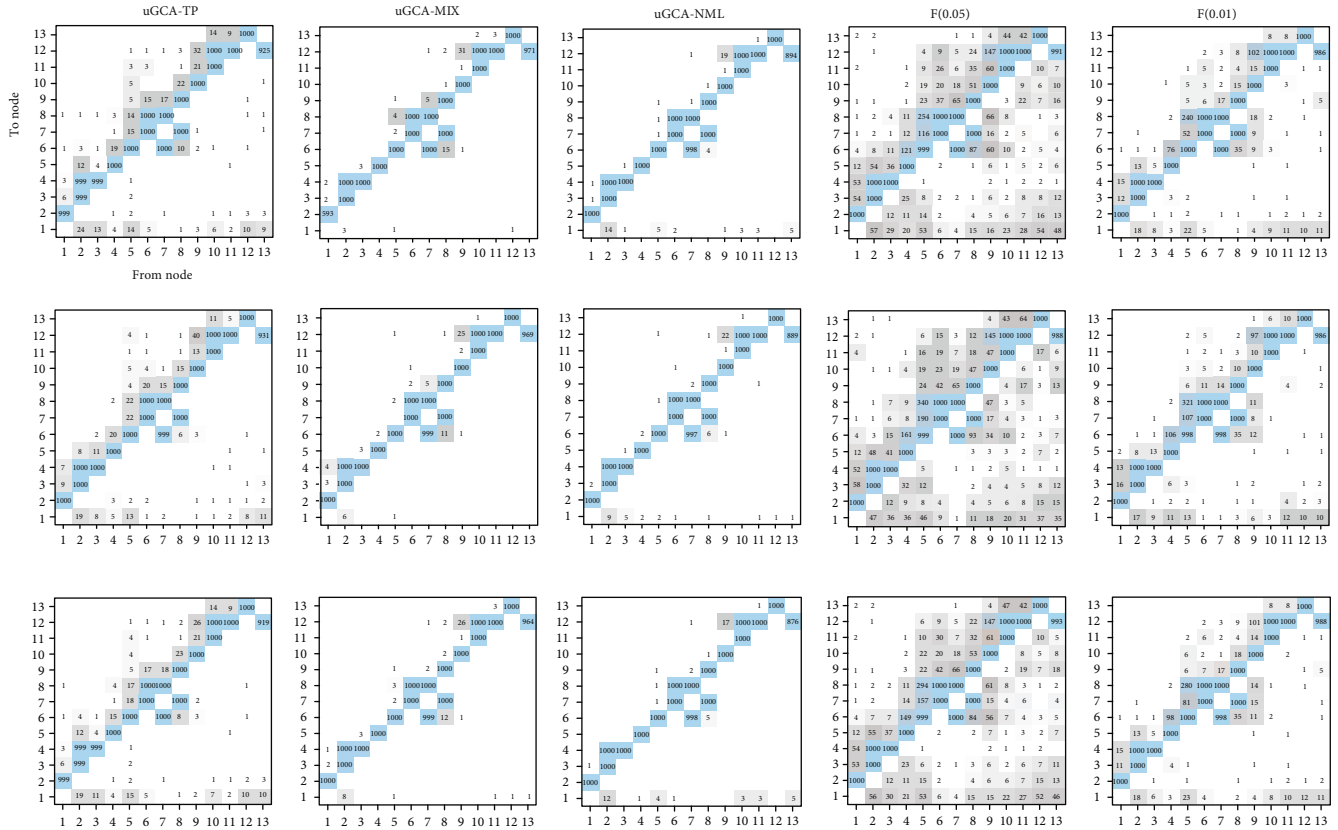


Figure 4: Causal connectivities obtained by uGCA and conventional GCA under different data length. The top row represented that data length was 200; data length with 500 was in the middle; the bottom denoted data length in 1000.

in identifying these true positives when data length was above 300 and fully obtained these connections when data length was 500. Therefore, due to the increase of data length, the performance of causal investigation in uGCA-NML had the most obvious improvement. To sum up, all the methods have a good anti-interference ability. Varying data length has no improvement for the performance of conventional GCA. By contrast, with the increase of data length, all uGCA forms have different degrees of improvement in all aspects which uGCA-NML is the most obvious. Specifically, uGCA-MIX can ensure relatively high TPR and TNR when its data length is short, and increasing the data length can further improve its performance. The uGCA-NML seems to rely on relatively long data length to ensure admirable identification ability, and it is not affected by noise. The uGCA-TP can be regarded as a conservative choice.

3.2. Resting-fMRI Data

3.2.1. Subjects. We downloaded freely available resting-state fMRI data from the website http://fcon_1000.projects.nitrc.org/indi/pro/nki.html, which all resting-state data on this NITRC website are freely available. The dataset comprised 37 male participants and 66 female participants (ages 20 ~ 23; mean age, 21.5 years; some left-handed).

3.2.2. AAL within uGCA. To further investigate the characteristics of uGCA method in large-scale network analysis, it is necessary to execute verification in real data. Within resting-state fMRI data, we applied several uGCA forms to identify causal connections of 90 brain regions in the Anatomical Automatic Labeling (AAL) template present by the Montreal Neurological Institute (MNI).

Figure 5 illustrates the causal connections obtained by uGCA-MIX in 90 regions; it was the most sparse network. At the same time, according to their causal connection networks, it seemed that there were also some functional clusters between 90 regions. In order to show the functional networks obtained by several uGCA forms more clearly, we presented $90 * 90$ connection matrices obtained by them, seen in Figure 6. Obviously, near the diagonal of these connection matrices, causal connections between some brain regions showed functional specialization. Meanwhile, the presence of causal connections elsewhere (not in the diagonal) indicated that these brain regions integrated to work together to cope with a specific task, scenario, stimulus, etc., which was consistent with the concept of functional integration of the brain.

Next, to further investigate the connection network obtained by these methods, we carried out a cluster analysis. Using the K-means algorithm, we clustered these connection networks into 26 clusters, as shown in Figure 7. In general, the causal connections of different forms of uGCA were similar to some extent, and the results of clustering were also similar to some extent, but there were still some differences in subtle points.

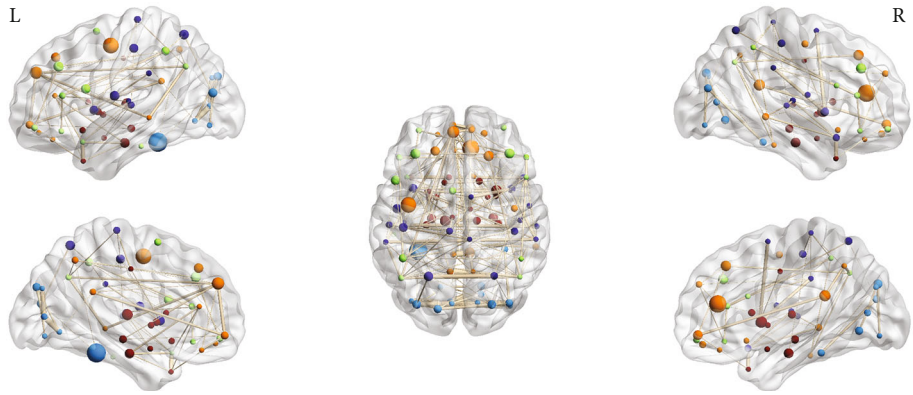


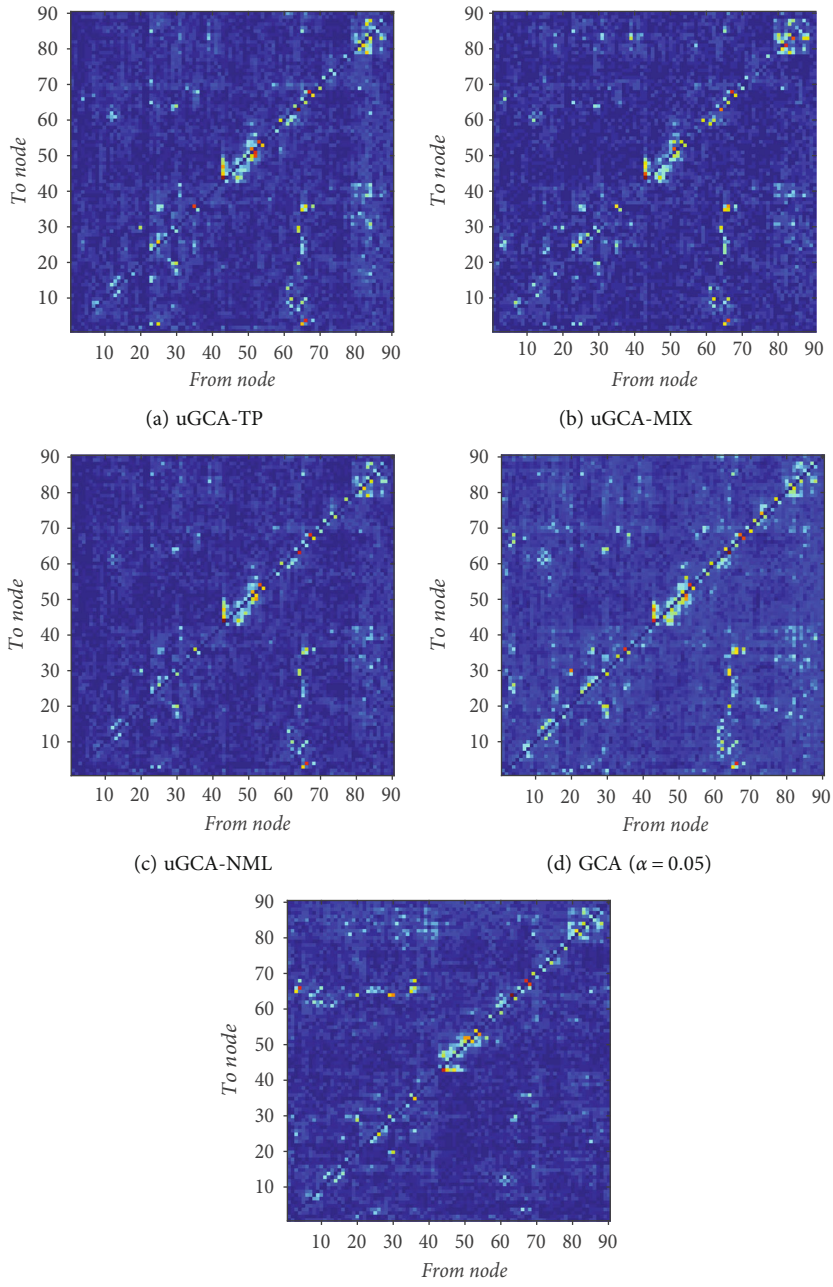
Figure 5: Causal connectivities between 90 regions obtained by uGCA-MIX form.

According to their connection matrices and clustering results, the relatively accurate functional specialization corresponding to each anatomical node was carried out. For example, we obtained a cluster of regions 45 to 58, named cluster-v, which was really the brain region associated with visual processing; the cluster of 63-70 nodes represented the brain region related to reading, self-cognition, and episodic memory, named cluster-r. The cluster containing regions 81-90 was a brain area associated with auditory processing, named cluster-a. The clustering results of these nodes were basically consistent with the current mainstream empirical views of their functional specialization regions [25–28]. Then, for nodes 1-44, they were clustered into several regions. According to the previous anatomical and clinical experiences, these brain regions represented the frontal lobe and limbic system, which were involved in a high level of cognitive processing and emotional processing. The internal and external connections were very complex, and the gap between samples was also very obvious; it is difficult to find more accurate clustering results between samples. Thus, according to Figure 6, we considered dividing nodes 1-44 into a cluster, named cluster-c. For the remaining nodes, they were more or less related to motion, named cluster-m.

At the same time, according to these results, it can be seen not only the functional specialization of 90 nodes but also their relevant connectivities among several large-scale clusters. For cluster-c, it had a close connection with other clusters, especially for cluster-r. For cluster-a, the interaction with the limbic system in cluster-c was considered to be due to the fact that large noise of fMRI machine working will generally affect the mood of subjects [26, 27, 29]. Finally, it seemed that visual processing cluster-v was not obviously connected with other clusters, which may be related to the resting state of subjects. Thus, combined with their connection matrices in Figure 5, these results showed that the clustering subnetworks of uGCA-TP and uGCA-NML were more consistent with the relevant anatomical and clinical experience [25, 28]. And the clustering subnetworks obtained by uGCA-MIX also had many similarities with the clustering results of the above two. However, there were some differences between several uGCA forms. The uGCA-MIX obtains a more sparse causal network, followed by uGCA-NML, and uGCA-TP obtains the network with the most connectivities. These were also consistent with the simulation experiment results. In contrast, conventional GCA had relatively poor clustering results at low confidence levels, and improved confidence levels appeared to improve. By the way, clustering analysis only adopted a relatively simple K-means method, and there was also some subjectivity in selecting clustering parameters, which meant that these results only had some reference values.

3.2.3. Three Large-Scale Functional Network within uGCA. Furthermore, we screened out the relevant nodes of several large-scale functional networks and then used several uGCA forms to perform causal connection analysis. According to previous research [29–32], we screened out 43 nodes, 19 nodes in the default mode network (DMN), 10 nodes in the dorsal attention network (DAN), and 14 nodes in the frontoparietal control network (FCN), seen in Table 1. Then, in Figure 8, three clusters can be clearly seen in the diagonal position of the connection matrix without using clustering methods, and they corresponded to the three functional networks of DMN, DAN, and FCN very accurately. For uGCA-TP, uGCA-NML, and conventional GCA, three corresponding clusters were obvious, and uGCA-MIX also can be seen some similar clusters in its diagonal position. In general, these methods also showed a consistent feature; the connection between DMN and FCN nodes was very dense. For DAN, there were some causal connectivities with FCN, but fewer connections to the DMN, which was consistent with relevant clinical and anatomical conclusions [25, 27–32].

In Figure 9, these properties of causal connections between these large-scale networks are even more pronounced. There are only a few connection edges between DMN and DAN regardless of the form of uGCA. For causal connections between DMN and FCN, their connection edges were very dense. The coupling effect of DAN and FCN was in between above. Moreover, we only showed those causal connection edges whose count was above 5% of sample size. Clearly, there were some differences in the characteristics of these causal networks obtained by different methods, but the overall performance was consistent, especially for uGCA-TP and uGCA-NML. It is worth noting that even though the uGCA results were very similar in different forms; there were still differences in the identification of causal connections among individuals. In other words, the causal connection between nodes itself has



(a) uGCA-TP

(b) uGCA-MIX

(c) uGCA-NML

(d) GCA ($\alpha = 0.05$)

(e) GCA ($\alpha = 0.01$)

Figure 6: Causal connections between 90 regions by uGCA and conventional GCA among 103 subjects: (a) uGCA-TP; (b) uGCA-MIX; (c) uGCA-NML; (d) conventional GCA ($\alpha = 0.05$); (e) conventional GCA ($\alpha = 0.01$).

dynamic attributes, and different forms of uGCA can be seen as describing the pseudodynamic causal connection between nodes from different aspects. For conventional GCA, the causal connections within three functional intrinsic networks are very dense, but it seemed too many connections between the networks, which was incompatible with the empirical knowledge of economical coupling among brain networks.

4. Discussion

First of all, for the uGCA method, when the large-scale network synthetic data model involving 13 nodes is involved, their performance advantages are significant compared with the conventional GCA. They also exhibit different characteristics, even if the overall performance is similar. To be specific, all uGCA methods have a certain anti-interference capability to noise. In addition to uGCA-MIX, this form does not seem to perform very well in the case of low noise, with some underreporting. But there is no denying that uGCA-MIX can guarantee high TPRs and TNRs in the case of insufficient data length. In contrast, uGCA-NML requires longer data to ensure high TPR and TNR, and it almost gets a "ground truth" network when the data length reaches 1000. The uGCA-TP is a conservative choice. It can guarantee high

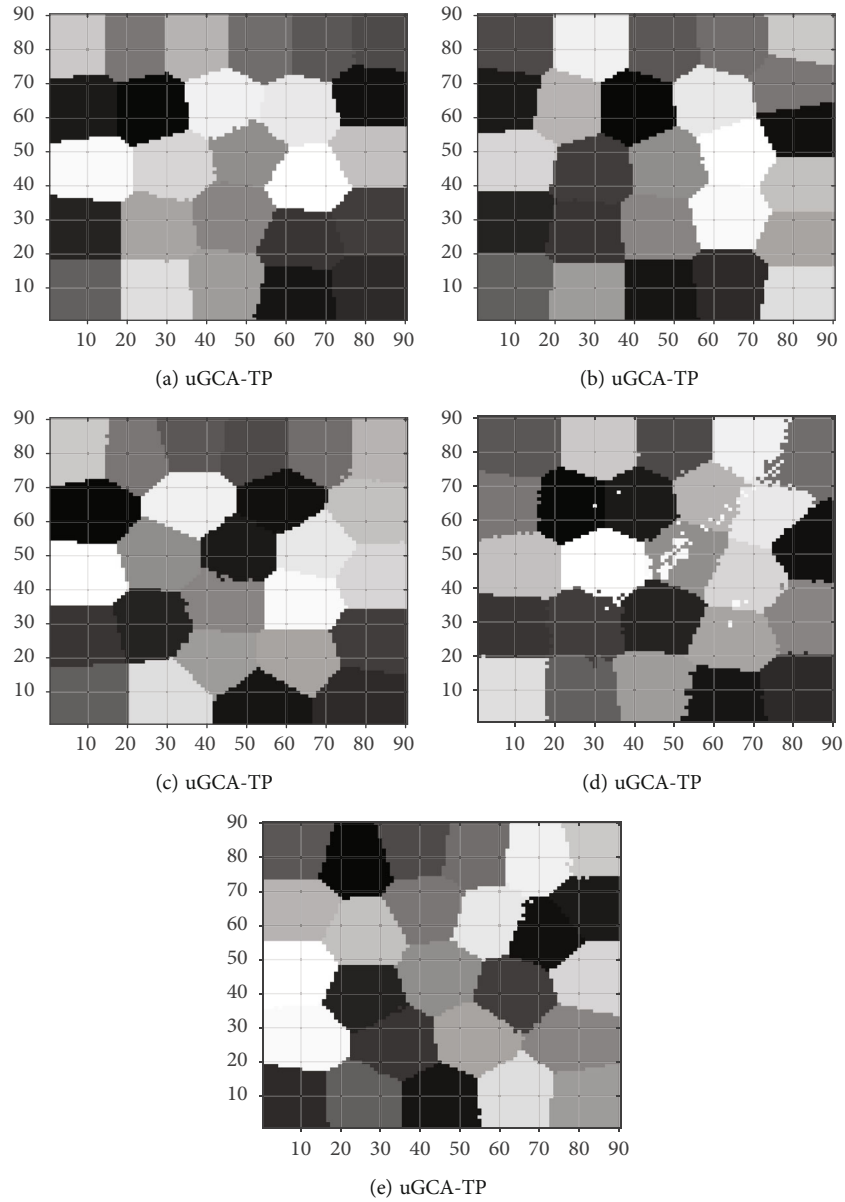


Figure 7: The clustering of causal connections between 90 regions: (a) uGCA-TP; (b) uGCA-MIX; (c) uGCA-NML; (d) conventional GCA ($\alpha = 0.05$); (e) conventional GCA ($\alpha = 0.01$). Maximum number of clusters is 26.

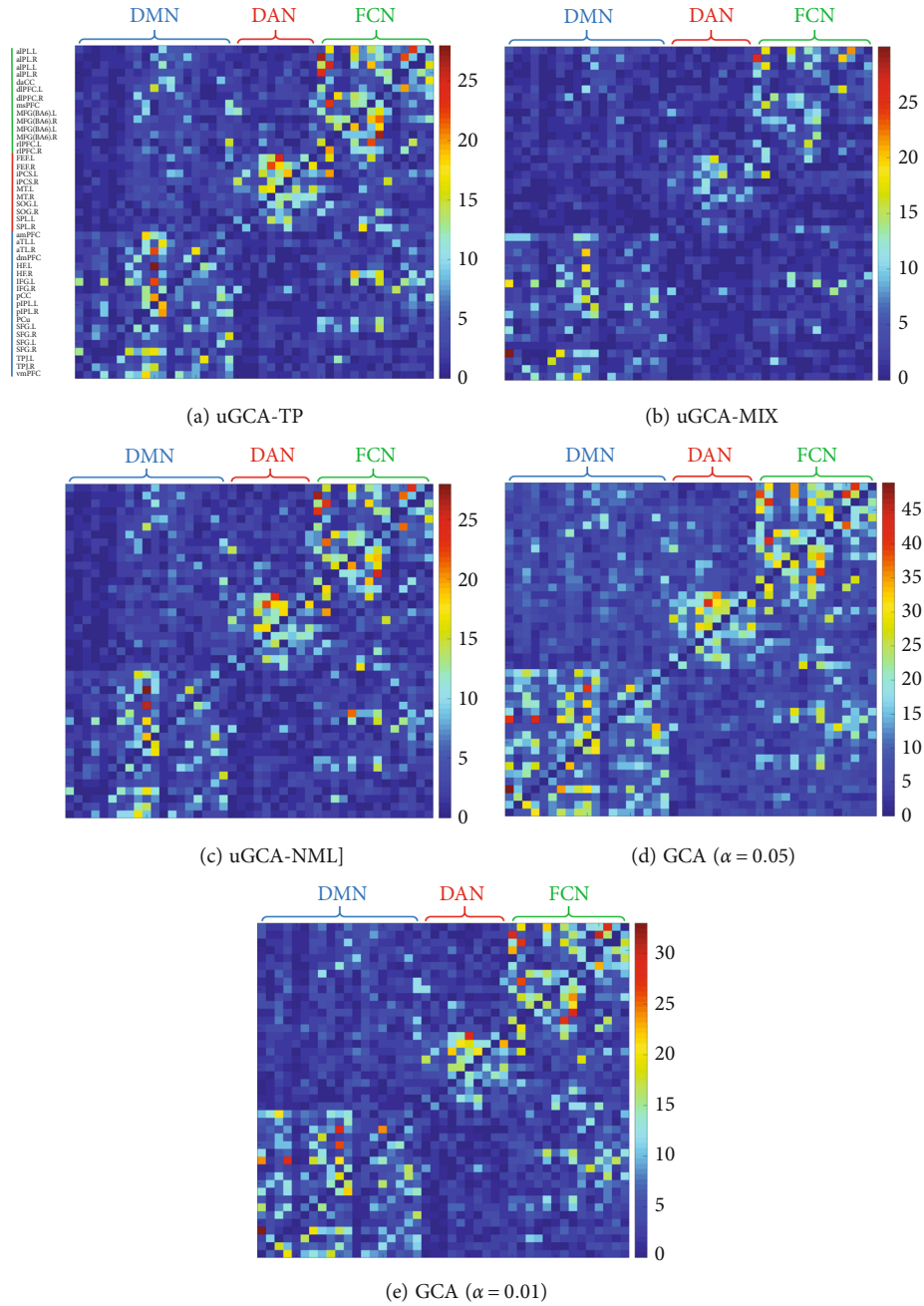
TPR, but there will be some false positives. Then, we used these methods for large-scale network analysis in the resting fMRI data, and the results showed that the characteristic performance demonstrated by these methods was consistent with their performances in the synthetic data. At the same time, the clustering analysis of the connection network can clearly see that some brain regions are clustered into several specific brain regions, and there were also collaborative works among these clustered brain regions. This was consistent with the current clinical and anatomical basis and further illustrated the need for a consistent causal investigation framework in large-scale networks. Further causal analysis in three large-scale functional networks was in line with the results of previous studies on causal connections between several networks and provided another strong evidence for the robustness verification of the uGCA method. By the way, both uGCA and conventional GCA methods adopted the same large-scale network analysis procedures in Section 2.2 among synthetic data and fMRI data experiments.

In addition, in order to eliminate the excessive false positives identified in conventional GCA, we have considered reducing these high false positives through a family-wise correction (e.g., Benjamini-Hochberg correction of the false discovery rate). However, considering that in this research, we focus on illustrating some problems in conventional GCA caused by the inconsistency of mathematical principles, subjective selection of confidence level, and algorithm complexity brought by the pairwise comparison. Different from those conventional two-stage schemes, we proposed the uGCA method to avoid these problems and formed a unified guiding framework for causality determination. Then, to make the comparison between uGCA and conventional GCA more

Table 1: Large-scale functional network location.

Macroanatomical region	Seed tag	Coordinates		
		x	y	z
Anterior medial prefrontal cortex	amPFC	-8	56	14
Left anterior temporal lobe	aTL.L	-52	-10	-20
Right anterior temporal lobe	aTL.R	52	-4	-16
Dorsal medial prefrontal cortex	dmPFC	-8	50	34
Left hippocampal formation	HF.L	-26	-8	-24
Right hippocampal formation	HF.R	24	-14	-22
Left inferior frontal gyrus	IFG.L	-42	26	-14
Right inferior frontal gyrus	IFG.R	50	32	-6
Posterior cingulate cortex	pCC	-2	-48	28
Left posterior inferior parietal lobule	pIPL.L	-50	-60	28
Right posterior inferior parietal lobule	pIPL.R	58	-60	28
Precuneus	PCu	-2	-60	50
Left superior frontal gyrus	SFG.L	-8	20	62
Right superior frontal gyrus	SFG.R	12	18	62
Left superior temporal sulcus	STS.L	-60	-28	-4
Right superior temporal sulcus	STS.R	50	-36	4
Left temporal parietal junction	TPJ.L	-44	-52	22
Right temporal parietal junction	TPJ.R	44	-58	18
Ventral medial prefrontal cortex	vmPFC	-2	44	-12
Left frontal eye fields	FEF.L	-24	2	62
Right frontal eye fields	FEF.R	24	-2	56
Left inferior precentral sulcus	iPCS.L	-36	0	28
Right inferior precentral sulcus	iPCS.R	42	6	26
Left middle temporal motion complex	MT.L	-44	-66	0
Right middle temporal motion complex	MT.R	54	-54	-6
Left superior occipital gyrus	SOG.L	-18	-66	50
Right superior occipital gyrus	SOG.R	26	-64	54
Left superior parietal lobule	SPL.L	-30	-48	52
Right superior parietal lobule	SPL.R	38	-46	54
Left anterior inferior parietal lobule	aIPL.L	-54	-48	48
Right anterior inferior parietal lobule	aIPL.R	50	-44	46
Left anterior insula	aINS.L	-30	20	-2
Right anterior insula	aINS.R	32	20	-4
Dorsal anterior cingulate cortex	daCC	6	30	40
Left dorsolateral prefrontal cortex	dlPFC.L	-38	32	30
Right dorsolateral prefrontal cortex	dlPFC.R	44	42	26
Medial superior prefrontal cortex	msPFC	-2	20	50
Left middle frontal gyrus BA6	MFG.L(BA6)	-28	14	58
Right middle frontal gyrus BA6	MFG.R(BA6)	26	16	48
Left middle frontal gyrus BA9	MFG.L(BA9)	-40	24	34
Right middle frontal gyrus BA9	MFG.L(BA9)	44	26	42
Left rostrolateral prefrontal cortex	rlPFC.L	-32	58	2
Right rostrolateral prefrontal cortex	rlPFC.R	32	58	8

Anatomical regions comprising the default mode network (19 nodes in the 1st grid), dorsal attention network (10 nodes in the 2nd grid), and frontoparietal control network (14 nodes in the 3rd grid) of the brain. BA = Brodmann's area. Coordinates (x, y, and z) are in MNI stereotaxic space.



(a) uGCA-TP

(b) uGCA-MIX

(c) uGCA-NML]

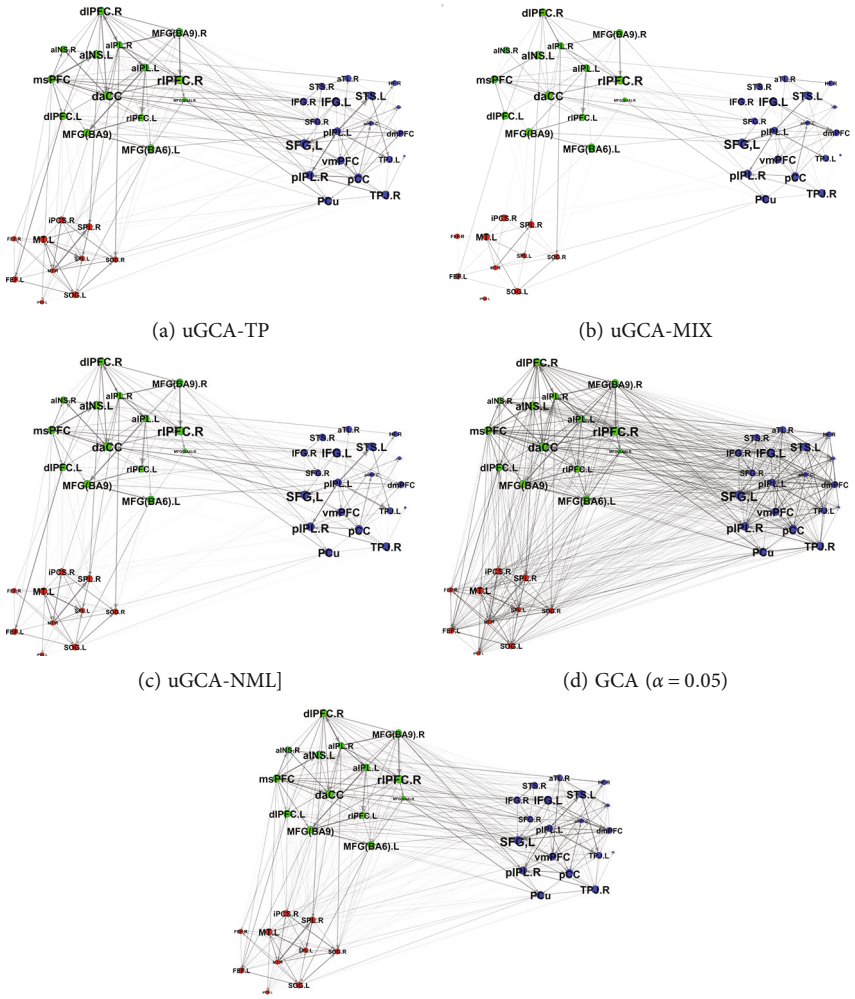
(d) GCA ($\alpha = 0.05$)

(e) GCA ($\alpha = 0.01$)

Figure 8: Causal connections obtained by several uGCA forms and conventional GCA. The Y-axis represents the driving node, and the X-axis represents the driven node.

direct, we only adopt different confidence levels of the F-test in conventional GCA instead of combining correction tools, which can more consistently and clearly show the original intention of this study. Also, applying correction tools brings another kind of mathematical theory, which is exactly what we tried to avoid in the original design of the uGCA method. Of course, it is meaningful to try to introduce relevant correction methods in the follow-up research and comparison.

As demonstrated above, uGCA's method has shown very good performance in both synthetic and real fMRI data. Compared with traditional GCA, the uGCA method integrates the whole causal analysis process into a unified framework, and all judgments and choices are mediated by the description length, which eliminates the inconsistency in mathematical principles, the subjective selection at the confidence level, and model complexity of pairwise models in the conventional process. In this study, we introduced large-scale network analysis into the uGCA framework, which can further improve the consistency of the resulting causal networks. In fact, the idea that the brain works together by integrating distributed networks is increasingly supported by relevant studies, so we should pay more attention to the problem of methodological rigor in large-scale network analysis. As we all know, with the increase of network nodes and



(a) uGCA-TP

(b) uGCA-MIX

(c) uGCA-NML]

(d) GCA ($\alpha = 0.05$)

(e) GCA ($\alpha = 0.01$)

FIGURE 9: Causal connectivities obtained by uGCA ($>0.05 * N$). N is the sample size. Blue nodes represent the default mode network, red nodes are dorsal attention network, and green nodes denote the frontoparietal control network.

connection edges, the resulting false or spurious connections can not be completely eliminated. Therefore, for causal analysis research involving large-scale networks, we should try our best to model all datasets of brain nodes to a unified space and then make a causal identification in this space. Otherwise, the process is not closed, the singularity occurs, and unexplained results occur.

Moreover, in this research, we have only preliminarily compared our proposal with most of the conventional GCA schemes, which involve two VAR models. Recently, Barnett and Anil have introduced state-space Granger causality in the literature, which can also accommodate the surrogate data analysis [33, 34]. Another study compared different methods for estimating GCA by showing that the state-space model was superior to the conventional double regression method [35]. In addition, a recent study introduced the application of different methods to overcome the limitations of the F-test and the chi-square test [36]. These methods make a good attempt to overcome the limitations of conventional GCA. Next, our method still needs to be considered for further testing and comparison with these GC estimation methods.

5. Conclusion

In this paper, we used the uGCA framework for the first time to investigate causality on large-scale networks. The unified guidance framework is of great significance for the definition of the final causal network, and different uGCA forms can describe the subtle differences of the connection characteristics of the network from different aspects. At the same time, several uGCA methods are suitable for studying causal networks in different scenarios, which provides universality for the uGCA framework. More importantly, in the future, causal network analysis in the brain will inevitably be closer to large-scale networks. Due to the complexity and uncertainty of causal analysis in large-scale networks, a rigorous and unified method is more needed to guide it.

Data Availability

We downloaded the freely available resting-state fMRI data (Oulu Kiviniemi, V.J./Veijiola, J.) from the website (FCP Classic Data Table (http://nitrc.org)), which all data on this NITRC website are freely available.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

Authors' Contributions

Zhenghui Hu and Fei Li conceived and designed the experiments. Fei Li performed the experiments and analyzed the data. Fei Li wrote the manuscript. Zhenghui Hu, Fei Li, Minjia Cheng, and Qiang Lin read and approved the final version of the manuscript.

Acknowledgments

This work is supported in part by the National Key Research and Development Program of China under Grant 2018YFA0701400, in part by the Public Projects of Science Technology Department of Zhejiang Province under Grant LGF20H180015.

References

[1] V. Menon, "Large-scale brain networks and psychopathology: a unifying triple network model," *Trends in Cognitive Sciences*, vol. 15, no. 10, pp. 483–506, 2011.

[2] S. L. Bressler and V. Menon, "Large-scale brain networks in cognition: emerging methods and principles," *Trends in Cognitive Sciences*, vol. 14, no. 6, pp. 277–290, 2010.

[3] X. Wang, R. Wang, F. Li, Q. Lin, X. Zhao, and Z. Hu, "Large-scale Granger causal brain network based on resting-state fMRI data," *Neuroscience*, vol. 425, pp. 169–180, 2020.

[4] M. Ding, Y. Chen, and S. L. Bressler, "17 granger causality: basic theory and application to neuroscience," in *Handbook of Time Series Analysis: Recent Theoretical Developments and Applications*, pp. 437–460, Wiley Online Library, 2006.

[5] S. Guo, A. K. Seth, K. M. Kendrick, C. Zhou, and J. Feng, "Partial Granger causality–eliminating exogenous inputs and latent variables," *Journal of Neuroscience Methods*, vol. 172, no. 1, pp. 79–93, 2008.

[6] D. Marinazzo, M. Pellicoro, and S. Stramaglia, "Kernel method for nonlinear granger causality," *Physical Review Letters*, vol. 100, no. 14, article 144103, 2008.

[7] Y. Chen, S. L. Bressler, and M. Ding, "Frequency decomposition of conditional granger causality and application to multivariate neural field potential data," *Journal of Neuroscience Methods*, vol. 150, no. 2, pp. 228–237, 2006.

[8] R. L. Wasserstein and N. A. Lazar, "The ASA statement onp-values: context, process, and purpose," *The American Statistician*, vol. 70, no. 2, pp. 129–133, 2016.

[9] D. J. Benjamin, J. O. Berger, M. Johannesson et al., "Redefine statistical significance," *Nature Human Behaviour*, vol. 2, no. 1, pp. 6–10, 2018.

[10] V. Amrhein, S. Greenland, and B. McShane, *Scientists Rise Up against Statistical Significance*, Nature Publishing Group, 2019.

[11] R. L. Wasserstein and N. A. Lazar, "Asa statement on statistical significance and p-values," in *The Theory of Statistics in Psychology*, pp. 1–10, Springer, 2020.

[12] F. Li, X. Wang, Q. Lin, and Z. Hu, "Unified model selection approach based on minimum description length principle in granger causality analysis," *IEEE Access*, vol. 8, pp. 68400–68416, 2020.

[13] Z. Hu, F. Li, X. Wang, and Q. Lin, "Description length guided unified granger causality analysis," *IEEE Access*, vol. 9, pp. 13704–13716, 2021.

[14] P. D. Grunwald, I. J. Myung, and M. A. Pitt, *Advances in Minimum Description Length: Theory and Applications, Ser. Neural Information Processing Series*, The MIT Press, Cambridge, MA, USA, 2005.

[15] M. Hansen and B. Yu, "Bridging aic and bic: an mdl model selection criterion," in *Proceedings of IEEE information theory workshop on detection, estimation, classification and imaging*, pp. 24–26, Santa Fe, NM, USA, 1999.

[16] M. H. Hansen and B. Yu, "Model selection and the principle of minimum description length," *Publications of the American Statistical Association*, vol. 96, no. 454, pp. 746–774, 2001.

[17] P. G. Bryant and O. I. Cordero-Brana, "Model selection using the minimum description length principle," *The American Statistician*, vol. 54, no. 4, pp. 257–268, 2000.

[18] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, no. 5, pp. 465–471, 1978.

[19] J. Rissanen, "A universal prior for integers and estimation by minimum description length," *Annals of Statistics*, vol. 11, no. 2, pp. 416–431, 1983.

[20] J. Rissanen, *Stochastic Complexity in Statistical Inquiry*, World Scientific, 1989.

[21] A. Barron, J. Rissanen, and B. Yu, "The minimum description length principle in coding and modeling," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2743–2760, 1998.

[22] A. Zellner, "On assessing prior distributions and Bayesian regression analysis with g-prior distributions," *Bayesian Inference and Decision Techniques*, vol. 6, pp. 233–243, 1986.

[23] J. J. Rissanen, "Fisher information and stochastic complexity," *IEEE Transactions on Information Theory*, vol. 42, no. 1, pp. 40–47, 1996.

[24] J. Rissanen, "Mdl denoising," *IEEE Transactions on Information Theory*, vol. 46, no. 7, pp. 2537–2543, 2000.

[25] G. E. Doucet, W. H. Lee, and S. Frangou, "Evaluation of the spatial variability in the major resting-state networks across human brain functional atlases," *Human Brain Mapping*, vol. 40, no. 15, pp. 4577–4587, 2019.

[26] F. Parente and A. Colosimo, "Functional connections between and within brain subnetworks under resting- state," *Scientific Reports*, vol. 10, no. 1, p. 3438, 2020.

[27] E. R. Kandel, J. H. Schwartz, T. M. Jessell et al., *Principles of Neural Science*, McGraw-hill New York, 2000.

[28] M. Bear, B. Connors, and M. A. Paradiso, *Neuroscience: Exploring the Brain*, Jones & Bartlett Learning, LLC, 2020.

[29] D. B. Dwyer, B. J. Harrison, M. Yucel et al., "Large-scale brain network dynamics supporting adolescent cognitive control," *Journal of Neuroscience*, vol. 34, no. 42, pp. 14096–14107, 2014.

[30] R. L. Buckner, J. R. Andrews-Hanna, and D. L. Schacter, "The brain's default network: anatomy, function, and relevance to disease.," in Wiley Online Library, 2008.

[31] R. N. Spreng, J. Sepulcre, G. R. Turner, W. D. Stevens, and D. L. Schacter, "Intrinsic architecture underlying the relations among the default, dorsal attention, and frontoparietal control networks of the human brain," *Journal of Cognitive Neuroscience*, vol. 25, no. 1, pp. 74–86, 2013.

[32] C. Gratton, H. Sun, and S. E. Petersen, "Control networks and hubs," *Psychophysiology*, vol. 55, no. 3, article e13032, 2018.

[33] L. Barnett and A. K. Seth, "Granger causality for state-space models," *Physical Review E*, vol. 91, no. 4, article 040101, 2015.

[34] T. Schreiber and A. Schmitz, "Improved surrogate data for nonlinearity tests," *Physical Review Letters*, vol. 77, no. 4, pp. 635–638, 1996.

[35] Y. Antonacci, L. Astolfi, and L. Faes, "Testing different methodologies for granger causality estimation: a simulation study," in *2020 28th European signal processing conference (EUSIPCO)*, pp. 940–944, Amsterdam, Netherlands, 2021.

[36] O. M. Cliff, L. Novelli, B. D. Fulcher, J. M. Shine, and J. T. Lizier, "Assessing the significance of directed and multivariate measures of linear dependence between time series," *Physical Review Research*, vol. 3, no. 1, article 013145, 2021.