Dynamic functional connectivity analysis based on time-varying partial correlation with a copula-DCC-GARCH model

Namgil Lee a, Jong-Min Kim b,∗

a Department of Information Statistics, Kangwon National University, Chuncheon, Gangwon 24341, South Korea
b Statistics Discipline, Division of Sciences and Mathematics, University of Minnesota-Morris, Morris, MN 56267-2134, USA

A R T I C L E   I N F O

Article history:
Received 2 September 2019
Received in revised form 12 June 2020
Accepted 23 June 2020
Available online xxx

Keywords:
Copula
Dynamic conditional correlation
Dynamic functional connectivity
Generalized AutoRegressive Conditional Heteroscedastic (GARCH)
Functional MRI
Partial correlation

A B S T R A C T

We suggest a time-varying partial correlation as a statistical measure of dynamic functional connectivity (dFC) in the human brain. Traditional statistical models often assume specific distributions on the measured data such as the Gaussian distribution, which prohibits their application to neuroimaging data analysis. First, we use the copula-based dynamic conditional correlation (DCC), which does not rely on a specific distribution assumption, for estimating time-varying correlation between regions-of-interest (ROIs) of the human brain. Then, we suggest a time-varying partial correlation based on the Gaussian copula-DCC-GARCH model as an effective method for measuring dFC in the human brain. A recursive algorithm is explained for computation of the time-varying partial correlation. Numerical simulation results demonstrate effectiveness of the partial correlation-based methods against pairwise correlation-based methods. In addition, a two-step procedure is described for the inference of sparse dFC structure using functional magnetic resonance imaging (fMRI) data. We illustrate the proposed method by analyzing an fMRI data set of human participants watching a Pixar animated movie. Based on twelve a priori selected brain regions in the cortex, we demonstrate that the proposed method is effective for inferring sparse dFC network structures and robust to noise distribution and a preprocessing step of fMRI data.

© 2020 Elsevier B.V. and Japan Neuroscience Society. All rights reserved.

1. Introduction

The functional connectivity (FC) in the brain refers to statistical dependency between functionally specialized areas of the brain (Friston, 1994). Whereas typical measures of FC such as correlation, coherence, and mutual information assume that the connectivity is static (Wang et al., 2014), recent studies demonstrated that FC in the human brain can change dynamically over time within time scales of seconds to minutes (Chang and Glover, 2010; Hindriks et al., 2016; Laumann et al., 2017; Liégeois et al., 2017).

In this study, we propose a statistical measure of dynamic functional connectivity (dFC) for functional magnetic resonance imaging (fMRI) data analysis without such restrictions as the Gaussian distribution assumption.

In recent years, numerous studies were carried out on dFC in the human brain based on neuroimaging data analysis; see, e.g., Hutchison et al. (2013) and Preti et al. (2017) for extensive reviews. The sliding window technique is the most basic approach to dFC, but an incorrect choice of window length can lead to spurious fluctuations and false detection of true dFCs (Hutchison et al., 2013; Leonardi and Van De Ville, 2015; Lindquist et al., 2014). Alternatively, model-based approaches are free from choice of the window length, and they outperformed the sliding window technique (Lindquist et al., 2014; Ryali et al., 2016). Especially, the family of Generalized AutoRegressive Conditional Heteroscedastic (GARCH) models has been developed for financial time series analysis (Bauwens et al., 2006; Engle, 2004). Among the multivariate GARCH models, the dynamic conditional correlation (DCC) GARCH model gained substantial attention for estimating time-varying correlations (Engle, 2002; Tse and Tsui, 2002), and it was applied to dynamic correlation analysis of resting-state fMRI by Lindquist et al. (2014).

On the other hand, a copula is a multivariate distribution function which can completely describe the dependence between random variables regardless of their marginal distributions (Nelsen, 1999; Sklar, 1973). Patton (2001) and Jondeau and Rockinger (2006) proposed the so-called copula-GARCH model, which we call the copula-DCC-GARCH model, which extends the DCC–GARCH model to copulas with time-varying parameters. Lee and Kim (2019) demonstrated that the copula-DCC–GARCH model outperformed the DCC–GARCH model in the cases of non-normal
skewed distributions through simulated experiments, and they applied the copula-DCC-GARCH model to an fMRI data set of 138 human participants watching a movie for their dFC structure.

This study proposes a time-varying partial correlation based on the copula-DCC-GARCH model as an effective approach for measuring dFC in the human brain. The partial correlation can measure the statistical dependence of each pair of regions-of-interest (ROIs) whose linear dependence to other ROIs has been removed (Whittaker, 1990). The partial correlation, and a similar concept, the partial coherence, has been widely used in the studies of FC in the brain (Marrelec et al., 2006; Smith et al., 2011). In this study, we suggest a time-varying partial correlation for dFC analysis in the human brain using fMRI data.

The main advantages of the proposed approach can be summarized as follows. First, the proposed time-varying partial correlation can be applied to dFC studies using a relatively large number of ROIs in the brain. We propose a recursive algorithm for computing the time-varying partial correlation efficiently, which enables us to study connectivity involving about 100 ROIs in the brain on laptop computers within a few minutes. Second, the proposed method does not have any restrictions on the distribution assumption on the data. The proposed method uses the copula-DCC-GARCH model so that dependencies between random variables can be completely described by the copula function. We remark that typical fMRIs or neuroimaging data sets have non-normal and skewed marginal distributions, in which cases the partial correlation cannot be applied due to its multivariate normal distribution assumption. Third, the proposed method is robust to artifacts in the data due to the property of partial correlation. Whereas the pairwise correlation-based methods such as the sliding window correlation and DCC measure both the direct and indirect dependences which are prone to external signal, the partial correlation-based methods can eliminate the external signal effect and measure only the direct dependence. In the same manner, the proposed method is robust to standard preprocessing steps for fMRI such as spatial smoothing or motion artifact removal, so it can be effectively applied to dFC analysis of fMRI data.

This paper is organized as follows. Section 2 describes the real fMRI data set and the preprocessing steps used in this study for analysis of dFC in the human brain. Section 3 explains the proposed time-varying partial correlation using the copula-DCC-GARCH model in a step-by-step manner. We also explain the recursive algorithm for efficient computation of the time-varying partial correlation. In Section 4, numerical simulation results based on three-dimensional simulated time series data are presented to compare the proposed method with other dynamic correlation estimation methods. Section 5 shows fMRI data analysis results using the proposed approach. Discussion and conclusions are presented in Sections 6 and 7.

2. Data

In this study, an fMRI data set of 155 human participants was obtained from the OpenfMRI database (Poldrack et al., 2013). The accession number of the data set is ds000228, and the data are freely available at https://openfmri.org/dataset/ds000228/ (Richardson et al., 2018).

2.1. Participants

The original data set consists of functional and anatomical MRIs of 155 human participants. We removed 22 participants during the preprocessing steps which will be explained in Section 2.3. The remaining 133 participants can be grouped according to their gender (‘F’ and ‘M’), age (‘3yo’, ‘4yo’, ‘5yo’, ‘6yo’, ‘8–12yo’, ‘Adult’), and handedness (‘Left’, ‘Right’, and ‘Ambiguous’). The ages of the participants in the ‘Adult’ age group are from 18 to 39. Table 1 summarizes the number of participants in each subgroup classified by their gender, age, and handedness.

2.2. Data acquisition

The participants were requested to watch a silent version of Disney Pixar’s animated movie ‘Partly Cloudy’ during the scanning of the fMRI brain images. The movie is a 5.6-min short film, and it was preceded by 10 s of rest. The participants were instructed to remain still while watching the movie.

The MRI data were acquired by using a 3-Tesla Siemens TIM Trio scanner. The total number of scans was 168 in one session with a repetition time (TR) of 2000 ms and an echo time (TE) of 30 ms. The number of slices was 32 with a slice thickness of 3.3 mm, a matrix size of 64 × 64, and a voxel dimension of 3 mm × 3 mm × 3.3 mm.

2.3. Preprocessing

The raw fMRI data were preprocessed by Statistical Parametric Mapping (SPM) version 12 from the Wellcome Trust Centre for Neuroimaging (Ashburner et al., 2017). We used the wrapper functions in the R package spm12r to use SPM version 12 (Muschelli, 2018). We performed slice-timing correction, spatial realignment to the mean volume, and spatial normalization to the Montreal Neurological Institute (MNI) brain template with all fMRI volumes of each participant. We did not perform spatial smoothing in the preprocessing steps in order to avoid erroneous correlations due to spatial smoothing (Alakorrokó et al., 2017).

We carried out additional preprocessing steps described in Satterthwaite et al. (2013) for removing in-scanner head motion artifacts. Recent studies showed that motion artifacts can influence functional connectivity significantly (Power et al., 2012; Satterthwaite et al., 2012; Van Dijk et al., 2012). First, we obtained six motion parameters consisting of three translations and three rotations during the realignment step. Then, framewise displacement (FD) was computed as a single time series by differentiating and summing the six motion parameters (Power et al., 2012). Each data frame was flagged when its FD was above a threshold of 1.75 mm, which was selected by the mean plus twice the standard deviation of the FD values over all the participants. Next, we carried out confound regression of voxelwise fMRI time series using the six motion parameters as regressors, and we performed spike regression for the motion censoring (scrubbing) using the frame flags. Finally, the fMRI time series were band-pass filtered using a Butterworth filter of order 5 with a frequency band of 0.01–0.08 Hz.

2.4. Selection of ROIs

We selected a total of twelve ROIs on the cortex to analyze dFC in the brain during the movie watching. First, we selected six Brodmann areas (BAs) (Brodmann, 2006) considering the visual-cognitive task of the experiment. The six BAs are the frontal eye fields (FEF), primary visual cortex (V1), inferior temporal gyrus (ITG), dorsal posterior cingulate cortex (dPCC), dorsal anterior cingulate cortex (dACC), and temporoparietal junction (TPJ). The FEF and V1 were selected considering the visual stimulus of the experiment. The three BAs, dPCC, dACC, and TPJ, were selected because previous studies showed that the Theory-of-Mind (ToM) brain regions including these three brain regions were activated while watching the same animated movie (Jacob et al., 2016; Richardson et al., 2018). The ITG was selected at the temporal lobe of the brain, considering the spatial distribution of the selected ROIs. Next, we identified the central locations of the six BAs on each of the right and left hemispheres, where the (x, y, z) -coordinates were defined
Table 1
Number of the participants in each subgroup classified by their age, gender, and handedness.

<table>
<thead>
<tr>
<th>Gender</th>
<th>3yo</th>
<th>4yo</th>
<th>5yo</th>
<th>7yo</th>
<th>8–12yo</th>
<th>Adult</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>M</td>
<td>F</td>
<td>M</td>
<td>F</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>Handed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ambi.</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>L</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>R</td>
<td>9</td>
<td>6</td>
<td>8</td>
<td>5</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>Total</td>
<td>10</td>
<td>6</td>
<td>8</td>
<td>6</td>
<td>15</td>
<td>16</td>
<td>12</td>
</tr>
</tbody>
</table>

Table 2
The Brodmann area (BA) numbers, abbreviated names, descriptions, and MNI coordinates of the selected twelve ROIs in the brain cortex.

<table>
<thead>
<tr>
<th>BA No.</th>
<th>Name</th>
<th>Description</th>
<th>MNI (x, y, z)-coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>BA 08</td>
<td>RFEF</td>
<td>Frontal eye fields (Right)</td>
<td>(22, 26, 45)</td>
</tr>
<tr>
<td>BA 08</td>
<td>LFEF</td>
<td>Frontal eye fields (Left)</td>
<td>(–23, 24, 44)</td>
</tr>
<tr>
<td>BA 17</td>
<td>RV1</td>
<td>Primary visual cortex (Right)</td>
<td>(11, –78, 9)</td>
</tr>
<tr>
<td>BA 17</td>
<td>LV1</td>
<td>Primary visual cortex (Left)</td>
<td>(–11, –81, 7)</td>
</tr>
<tr>
<td>BA 20</td>
<td>RITG</td>
<td>Inferior temporal gyrus (Right)</td>
<td>(48, –17, –31)</td>
</tr>
<tr>
<td>BA 20</td>
<td>LITG</td>
<td>Inferior temporal gyrus (Left)</td>
<td>(–47, –14, –34)</td>
</tr>
<tr>
<td>BA 31</td>
<td>RdPCC</td>
<td>Dorsal posterior cingulate cortex (Right)</td>
<td>(8, –48, 39)</td>
</tr>
<tr>
<td>BA 31</td>
<td>LdPCC</td>
<td>Dorsal posterior cingulate cortex (Left)</td>
<td>(–8, –49, 38)</td>
</tr>
<tr>
<td>BA 32</td>
<td>RdACC</td>
<td>Dorsal anterior cingulate cortex (Right)</td>
<td>(6, 33, 16)</td>
</tr>
<tr>
<td>BA 32</td>
<td>LdACC</td>
<td>Dorsal anterior cingulate cortex (Left)</td>
<td>(–5, 39, 20)</td>
</tr>
<tr>
<td>BA 39</td>
<td>RTPJ</td>
<td>Temporoparietal junction (Right)</td>
<td>(46, –59, 31)</td>
</tr>
<tr>
<td>BA 39</td>
<td>LTPJ</td>
<td>Temporoparietal junction (Left)</td>
<td>(–46, –60, 33)</td>
</tr>
</tbody>
</table>

From Eq. (3.1), the joint probability density function (PDF) can be written as

\[
f(x_1, x_2, \ldots, x_n) = c(F_1(x_1), F_2(x_2), \ldots, F_n(x_n)) = \prod_{i=1}^{n} f_i(x_i),
\]

where \( c(u_1, u_2, \ldots, u_n) = \frac{\partial^n C(u_1, u_2, \ldots, u_n)}{\partial u_1\partial u_2\cdots\partial u_n} \) is the PDF for the copula \( C(u_1, u_2, \ldots, u_n) \), and \( f_i(x_i) = \frac{\partial F_i(x_i)}{\partial x_i} \) is the marginal PDF. In this study, we chose the Gaussian copula from a wide variety of copulas because the PDF for the Gaussian copula can be expressed in closed form, and it can be parameterized explicitly with the time-varying correlations. The PDF for the Gaussian copula can be explicitly written as

\[
c_{\text{Gaussian}}(u_1, u_2, \ldots, u_n; \mathbf{R}_t) = \frac{1}{|\mathbf{R}_t|^{1/2}} \exp \left\{ -\frac{1}{2} \mathbf{z}^\top (\mathbf{R}_t^{-1} - I) \mathbf{z} \right\},
\]

where \( \mathbf{R}_t \) is the \( n \times n \) time-varying correlation matrix, \( \mathbf{z} = (z_1, \ldots, z_n)^\top \) with \( z_i = \Phi^{-1}(u_i) \), \( \Phi \) is the standard normal CDF, and \( I \) is the identity matrix.

The procedure for estimation of the time-varying correlation, \( \mathbf{R}_t \), is comprised of two levels. Let \( \mathbf{x}_t = (x_{1t}, \ldots, x_{nt})^\top \) denote the vector of the observed fMRI data at time \( t \). First, the marginal PDF \( F_t(\cdot) \) is modeled through univariate GARCH models. The ARMA(0,0)-GARCH(1,1) model for the univariate process \( x_{it} \) can be described by

\[
x_{it} = \mu_t + \epsilon_{it}, \quad \epsilon_{it} = \alpha_0 + \alpha_1 \epsilon_{it-1} + \beta_1 \eta_{it-1},
\]

where \( \eta_{it} \) is an independent random variable with a mean of zero and a variance of one, \( \alpha_0 > 0, \alpha_1 \geq 0, \beta_1 \geq 0 \), and \( \alpha_1 + \beta_1 < 1 \) (Tsay, 2005). The \( \eta_{it} \) is assumed to have the reparameterized Johnson’s SU distribution, \( JSU(\mu, \sigma, \nu, \kappa) \), where \( \mu \) is the mean, \( \sigma \) is the standard deviation, \( \nu \) is the skewness, and \( \kappa \) is the kurtosis of the distribution (Ghalanos, 2015; Stasinopoulos et al., 2018).

Second, before formulating the likelihood function via the copula framework in (3.2) and (3.3), the time-varying correlation matrix, \( \mathbf{R}_t \), is further modeled by the DCC-GARCH model of Engle (2002). We consider that the data time series \( \mathbf{x}_t \) can be decomposed into the conditional mean and the residuals as \( \mathbf{x}_t = \mu_t + \mathbf{e}_t \). We assume that \( \mu_t = \mathbf{c} \) is a constant vector so that connectivity

is modeled by the residual process $\epsilon_t$. The $\epsilon_t = (\epsilon_{1t}, \ldots, \epsilon_{nt})^T$ has a mean of zero and a conditional covariance matrix of $H_t = (h_{ij,t})$. The covariance matrix is separated into the product of the standard deviations and the correlation as

$$ H_t = D_t R_t D_t = \left( \rho_{ij,t} \sqrt{h_{ij,t}} h_{jj,t} \right), $$

(3.5)

where $D_t = \text{diag}(\sqrt{h_{11,t}}, \ldots, \sqrt{h_{nt,t}})$ is the diagonal matrix of standard deviations, and $R_t = (\rho_{ij,t})$ is the correlation matrix. Each of the conditional variances is modeled by the GARCH(1,1) model as

$$ h_{ij,t} = \alpha_{i0} + \alpha_{1i} \epsilon_{ij,t-1} + \beta_{i1} h_{ij,t-1}. $$

(3.6)

for $i = 1, \ldots, n$. The correlation matrix is modeled by $R_t = D_t^{-1} R_t D_t$, $Q_t = (1 - a - b) \hat{Q} + a \hat{Z}_{t-1}^2 + b Q_{t-1}$, $R_t = \text{diag}(Q_{11,t}^{-1/2}) \text{diag}(Q_{22,t}^{-1/2})$, where $\epsilon_t$ is the standardized residual, $\hat{Q}$ is the unconditional correlation matrix of $\epsilon_t$, and $a, b > 0$ are the scalars satisfying $a + b < 1$.

The parameters including the time-varying correlations in $R_t$ can be estimated by an efficient two-step maximum likelihood approach, where the conditional log-likelihood can be decomposed into the sum of the variance term and the correlation term (Ghalanos, 2015). The conditional likelihood can be formulated based on the copula framework in (3.2) and (3.3), and the time-varying correlations can be recursively updated as in (3.8) and (3.9).

### 3.2. Time-varying partial correlation

The copula-based dynamic conditional correlation matrix $R_t = (\rho_{ij,t})$ for the preprocessed fMRI data $x_t = (x_{1t}, x_{2t}, \ldots, x_{nt})^T$ at time $t$ can be obtained from the copula-DCC-GARCH model described in the previous section. In this section, we explain an efficient computational procedure for the time-varying partial correlation based on the copula-DCC-GARCH model.

Let $Y = x_{2t}, Z = x_{3t}$ be random variables and $T$ denote a random vector consisting of a subset of the remaining random variables in $x_t$. The partial covariance of $Y$ and $Z$ given $T$ is defined by the covariance between the residuals of the linear least squares regression as

$$ \text{cov}(Y, Z|T) = \text{cov}(Y - \hat{Y}(T), Z - \hat{Z}(T)), $$

(3.10)

where $\hat{Y}(T)$ and $\hat{Z}(T)$ are the linear least squares prediction on $Y$ and $Z$ with the predictor $T$ (Whittaker, 1990). The partial correlation is defined based on the partial covariance by

$$ \rho_{Y,Z|T} = \frac{\text{cov}(Y, Z|T)}{\sqrt{\text{var}(Y|T)\text{var}(Z|T)}} $$

(3.11)

where $\text{var}(Y|T) = \text{cov}(Y, Y|T)$ and $\text{var}(Z|T) = \text{cov}(Z, Z|T)$ are the partial variances.

The time-varying partial correlations can be computed from the copula-based dynamic conditional correlations, $\rho_{ij,t}$, by using the general recursive formula (Whittaker, 1990; Yule and Kendall, 1965). Let $X = x_{3t}, Y = x_{1t}, Z = x_{2t}$ be three random variables with $k \neq i, k \neq j$, and $T$ denoting a random vector consisting of a subset of the remaining random variables in $x_t$. The general recursive formula can be written as

$$ \rho_{Y,Z|T} = \frac{\pi_{Y,Z|T} - \pi_{Y,T|T} \pi_{X,T|T}}{\left(1 - \pi_{Y,T|T}^2 \right) \left(1 - \pi_{X,T|T}^2 \right)}^{1/2}. $$

(3.12)

Eq. (3.12) shows that the partial correlations given $(X, T)$ can be computed by using the partial correlations given $T$. The computational procedure is summarized in Algorithm 1.

### 4. Numerical simulation

We conducted extensive simulations to compare the proposed method with other methods for estimating dynamic correlations. Eight methods were compared in the simulation; four of them were pairwise correlation-based methods (items 1–4 below), and the other four methods were partial correlation-based methods (items 5–8 below).

1. SWC(w = 10): Sliding window correlation with a window length of $w = 10$.
2. SWC(w = 30): Sliding window correlation with a window length of $w = 30$.
3. SWC(w = 50): Sliding window correlation with a window length of $w = 50$.
4. DCC: The copula-DCC-GARCH model using the Gaussian copula (Jondeau and Rockinger, 2006; Patton, 2001).
5. SWPC(w = 10): The partial correlation computed based on SWC(w = 10).
6. SWPC(w = 30): The partial correlation computed based on SWC(w = 30).
7. SWPC(w = 50): The partial correlation computed based on SWC(w = 50).
8. DPC: The proposed partial correlation computed based on DCC.

We constructed a three-dimensional time series model in order to compare the pairwise correlation-based methods and the partial correlation-based methods, which is illustrated in Fig. 1. The graphical representation in the figure implies that, among the three time series variables $(y_{1t}, y_{2t}, y_{3t})$, $y_{2t}$ and $y_{3t}$ are independent conditionally on $y_{1t}$. In other words, the true partial correlation $\pi_{23|1} = \rho_{y_{2t}y_{3t}|y_{1t}} = 0$ was set to zero, and the other two true partial correlations, $\pi_{12|3}$ and $\pi_{13|2}$, were set to certain predefined values. In this study, we used two types of functions for the true partial correlations in consideration of typical fMRI experimental designs as follows.

(a) Step function: $\pi_{12|3} = \pi_{13|2} = 0.6 \sin(2\pi t(200)^{-1})$ for $t = 1, 2, \ldots, 800$. The step function assumes a rapid...

![Fig. 1](image-url)
sition of correlation states, which may occur from a block design in fMRI experiments consisting of several on-off periods.

(b) Sine function: \( \tau_{12t} = \tau_{23t} = 0.5 \) for \( 100k < t \leq 100(k+1) \) for \( k = 0, 2, 4, 6; 0 \) otherwise. The sine function assumes a slow transition of correlation states.

The two types of functions are depicted in the left columns of Fig. 3 and Figs. A.14 to A.20 in Appendix A as red dotted lines. For data generation, we first defined the \( 3 \times 3 \) precision matrix \( \Theta_t \) as

\[
\Theta_t = \begin{pmatrix}
1 & -\tau_{12t} & -\tau_{13t} \\
-\tau_{12t} & 1 & 0 \\
-\tau_{13t} & 0 & 1
\end{pmatrix}.
\]

Then, a covariance matrix was computed by normalizing the inverse precision matrix as \( \Sigma_t = diag(\Theta_t^{-1})^{-1/2} \Theta_t^{-1} diag(\Theta_t^{-1})^{-1/2} \).

Next, three-dimensional time series data, \( \{\tilde{y}_{1t}, \tilde{y}_{2t}, \tilde{y}_{3t}\}, t = 1, \ldots, 800 \), were generated from the trivariate normal distribution with a mean of zero and a covariance matrix of \( \Sigma_t \). Note that the marginal distributions of \( \tilde{y}_{2t}, i = 1, 2, 3 \), are the standard normal distribution, and the two variables \( \tilde{y}_{2t} \) and \( \tilde{y}_{3t} \) are conditionally independent given \( \tilde{y}_{1t} \).

Finally, in order to simulate non-normal distributions in the data, the marginal distributions of \( \tilde{y}_{it}, i = 1, 2, 3 \), were transformed to one of the four types of distributions:

(i) \( N(0, 1) \): The standard normal distribution.
(ii) \( t_4 \): The Student’s t-distribution with \( v = 4 \) degrees of freedom.
(iii) \( LN(0, 1) \): The log-normal distribution with location \( \mu = 0 \) and scale \( \sigma = 1 \).
(iv) \( \text{Beta}(5, 2) \): The beta distribution with shapes \( \alpha = 5 \) and \( \beta = 2 \).

The density functions of the four types of marginal distributions are illustrated in Fig. 2. That is, the final time series data were obtained by \( y_{it} = F(\Phi(y_{it})), i = 1, 2, 3, t = 1, \ldots, 800 \), where \( F(x) \) is the CDF of one of the four types of distributions, and \( \Phi \) is the standard normal CDF. We used the four types of distributions because they possess various values of skewness and kurtosis, which are indices of the non-normality of distributions. The \( N(0, 1) \) and \( t_4 \) distributions have a skewness of 0, and the \( LN(0, 1) \) and \( \text{Beta}(5, 2) \) distributions have skewnesses of 6.19 and -0.60, respectively. The kurtoses of \( N(0, 1), t_4, LN(0, 1), \) and \( \text{Beta}(5, 2) \) are 3, \( \infty \), 113.94, and 2.88.

We applied the eight methods for estimating the dynamic correlations from the simulated three-dimensional time series data. Fig. 3 illustrates an example of the dynamic correlations estimated by the eight methods in black straight lines when the data were generated using the step function and the normal distribution \( N(0, 1) \). It is apparent that the short window length \( (w = 10) \) for the sliding window correlation (SWC) and sliding window partial correlation (SWPC) yielded very noisy and unreliable estimates (the first and the fifth rows) compared to the other methods, which implies that the selection of an optimal window length is a critical issue for the sliding window methods. On the other hand, for the estimation of the true correlation between \( y_{1t} \) and \( y_{2t} \) (left column), the pairwise correlation-based methods and the partial correlation-based methods produced similar results fitting to the true step correlation function. However, for the estimation of the true correlation between \( y_{2t} \) and \( y_{3t} \) (right column), which is the zero function, only the partial correlation-based methods could produce estimates close to the true correlations (the sixth–eighth rows). Figs. A.14 to A.20 in Appendix A shows the other examples using the two correlation functions and the four types of distributions.

In order to evaluate the accuracy of the eight methods, we repeatedly generated time series data 50 times independently using one of the two true correlation functions and one of the four marginal distributions. The root mean squared deviation (RMSD) of an estimated dynamic correlation, \( \hat{\tau}_{it}(l) \), at the \( i \)th repetition is defined by

\[
\text{RMSD} \left( \{ \hat{\tau}_{it}(l) \} \right) = \frac{1}{T} \sum_{l=1}^{T} \left( \hat{\tau}_{it}(l) - \tau_{it} \right)^2,
\]

for \( i = 1, 2, \ldots, 50 \), where \( T \) is the total number of time points. Figs. 4–7 show the boxplots of the RMSDs comparing the eight dynamic correlations. In Fig. 4(a) and (b), where the time series data were generated using the \( N(0, 1) \) distribution, the DCC and DPC were yielded the smallest RMSDs for the estimation of \( \tau_{12t} \) between \( y_{1t} \) and \( y_{2t} \), which implies the effectiveness of the DCC and DPC for estimation of a rapid transition (the step function) or a slow transition (the sine function) of correlation states. The same results can be found for the estimation of \( \tau_{12t} \) (the step function or the sine function) using the other types of marginal distributions in Figs. 5–7.

On the other hand, in Fig. 4(a) and (b), while both the SWC and SWPC yielded less accurate estimates of \( \tau_{12t} \) (between \( y_{1t} \) and \( y_{2t} \) ) than those of the DCC and DPC, the medium window length of \( w = 30 \) obtained smaller RMSDs compared to the other window lengths. It is noteworthy that the longest window length of \( w = 50 \) obtained the smallest RMSDs than the other window lengths in the estimation of \( \tau_{23t} \) (between \( y_{2t} \) and \( y_{3t} \)). It implies that the sliding window technique requires a sophisticated decision of the window length, and an optimal length may not be fixed over different pairs of variables in the same data set. This phenomenon can be found in all the results using the other types of marginal distributions in Figs. 5–7.

The advantage of the partial correlation-based methods over the pairwise correlation-based methods can be found clearly in the estimation of \( \tau_{23t} \) (between \( y_{2t} \) and \( y_{3t} \)) in Figs. 4–7. In all the results, the RMSDs of the partial correlation-based methods were lower than those of their pairwise correlation-based counterparts. Due to the model assumption, the true partial correlation \( \tau_{23t} \) is the zero function, so we can conclude that the partial correlation-based methods produced reduced fluctuation in the estimates and achieved the lower RMSDs.

We note that the SWPC with the longest window length \( w = 50 \) often produced the lowest RMSDs for the estimation of \( \tau_{23t} \).
Fig. 3. Dynamic correlations estimated by the eight methods (black straight line) and the true correlation functions (red dotted line) for a sample of three-dimensional time series, \((y_{1t}, y_{2t}, y_{3t})\). The true correlation between \(y_{1t}\) and \(y_{2t}\) is the step function (left column), and the true correlation between \(y_{2t}\) and \(y_{3t}\) is the zero function (right column). The marginal distribution of each time series variable is the normal distribution, \(N(0, 1)\). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(between \(y_{2t}\) and \(y_{3t}\)) when the sine function was used as the \(\pi_{12t}\) in Figs. 4(b)–7(b). However, the SWPC with \(w = 50\) yielded much worse RMSDs in the estimation of \(\pi_{12t}\) in the figures, which was because long window lengths were not suitable for estimating continually changing dynamic correlations; see, e.g., the left column of Fig. 3 or A.14 in Appendix A.

It is noteworthy that the type of marginal distribution did not affect the RMSD results significantly except in the case of the log-normal distribution, \(LN(0, 1)\). That is, in Figs. 4, 5, and 7, the RMSD results for the \(N(0, 1), t_4\), and \(Beta(5, 2)\) distributions were not significantly different. Note that the three distributions have a similar skewness of 0, 0, and −0.60 while the skewness of \(LN(0, 1)\) is 6.19. The results imply that the skewness of the marginal distribution can significantly affect the estimation accuracies of dynamic correlation methods.

In Fig. 6, the RMSDs of the SWC and SWPC were higher in most of the cases using \(LN(0, 1)\) than in the cases using other distributions in Figs. 4, 5, and 7. On the other hand, we could find that the RMSDs of the DCC and DPCC were not sensitively different over the choices among the four types of marginal distributions, which implies that the copula-based approaches are robust to the high skewness in the marginal distributions.
5. Results

5.1. Analysis of a single participant

For illustration of the proposed time-varying partial correlations based on the copula DCC, we selected a participant from the group of age ‘4yo’ and gender ‘F’. We selected the twelve ROIs on the brain cortex, where six ROIs are on the right hemisphere and the other six are on the left hemisphere. Fig. 8 shows the preprocessed fMRI time series data obtained at the six ROIs selected on the right hemisphere of the brain.

In order to test the normality of the fMRI time series data, we conducted the Shapiro–Wilk normality test using the fMRI data of the same participant (age ‘4yo’, gender ‘F’). The test yielded p-values of 0.0033 (RFEF), 0.1632 (RV1), 0.1353 (RITG), 0.0001 (RdPCC), 0.4254 (RdACC), and 0.8772 (RTPJ) for the ROIs on the right hemisphere, and p-values of 0.1416 (LFEF), 0.0474 (LV1), 0.0534 (LITG), 0.0008 (LdPCC), 0.2219 (LdACC), and 0.2030 (LTPJ) for the ROIs on the left hemisphere. The p-values clearly demonstrate that the preprocessed fMRI data are non-normally distributed, which will hinder the use of standard statistical methods such as the partial correlation.

We computed the Spearman correlation coefficients as a measure of pairwise correlation between the selected twelve ROIs, which are summarized in Table 3. A Spearman correlation coefficient is the correlation coefficient between the ranks of the measured values, which is independent of any monotone transformations on the marginal distributions. It has been known that many dependence measures including the Spearman correlation coefficient are closely related to copulas (Cuadras, 2002; Fredricks and...
The root mean squared deviations (RMSDs) of the eight dynamic correlation estimation methods. Three-dimensional time series, \( (y_{1t}, y_{2t}, y_{3t}) \), were generated under the condition of (a) the log-normal distribution, \( \mathcal{N}(0, 1) \), and the step function between \( (y_{1t}, y_{2t}) \), and (b) the log-normal distribution, \( \mathcal{N}(0, 1) \), and the sine function between \( (y_{1t}, y_{2t}) \).

The root mean squared deviations (RMSDs) of the eight dynamic correlation estimation methods. Three-dimensional time series, \( (y_{1t}, y_{2t}, y_{3t}) \), were generated under the condition of (a) the beta distribution, \( \text{Beta}(5, 2) \), and the step function between \( (y_{1t}, y_{2t}) \), and (b) the beta distribution, \( \text{Beta}(5, 2) \), and the sine function between \( (y_{1t}, y_{2t}) \).

Table 3
Spearman correlation coefficients of the preprocessed fMRI data. Information on the selected twelve ROIs and their abbreviated names are summarized in Table 2. The values greater than the mean plus the standard deviation, which is 0.4223, are marked in bold font.
Nelsen, 2007; Schweizer and Wolff, 1981). In the table, we can find that the ROIs on the left and right hemispheres are highly dependent to each other; the four ROIs ((R/L)dPCC, (R/L)TPJ) are relatively highly dependent; and the three ROIs on the right hemisphere (RFEF, RdPCC, RTPJ) are relatively highly dependent. However, note that the dependence measured by the Spearman correlation coefficient does not remove spurious correlations by regressing out the influence from other ROIs.

Fig. 9 illustrates the dynamic conditional correlations (DCCs) and the dynamic (time-varying) partial conditional correlations (DPCCs) between the two ROIs LV1 and LdPCC. The figure shows that the DCCs have larger fluctuations than the DPCCs. The DCCs vary between −1 and 1, while the DPCCs vary between −0.1 and 0.5. Figs. A.21–A.29 in Appendix A illustrate the DCCs and DPCCs between each pair of five ROIs, which are LV1, LdPCC, LdACC, RTPJ, and LTPJ. Those five ROIs have been used in previous studies in search of the Theory-of-Mind (ToM) brain networks (Jacoby et al., 2016; Richardson et al., 2018).

Moreover, for dFC analysis, the significance of a partial correlation can be tested based on Fisher’s z-transformation. Fisher’s z-transformation of a sample (partial) correlation has an approximately normal distribution (Fisher, 1915), where the Fisher’s z-transformation of \( \pi_{Y,Z|I} \) is defined by

\[
z(\pi_{Y,Z|I}) = \frac{1}{2} \ln \left( \frac{1 + \pi_{Y,Z|I}}{1 - \pi_{Y,Z|I}} \right). \tag{5.1}
\]

The time-varying (partial) correlation \( \pi_{Y,Z|I} \) is significantly different from zero if the absolute value of the Fisher’s z-transformation is larger than a threshold value. In this study, the threshold value was determined adaptively by the false discovery rate (FDR) procedure of Strimmer (2008). The FDR procedure was employed to control the FDR of the multiple hypothesis test. In the FDR procedure, the empirical distribution of the z values is modeled by the mixture of a null distribution and an alternative distribution as

\[
f(z) = \eta f_0(z; \kappa) + (1 - \eta) f_A(z), \quad 0 \leq \eta \leq 1, \tag{5.2}
\]

where \( f_0 \) is the normal density function with a mean of zero and a standard deviation of \( \kappa \). The local FDR score is computed by

\[
fdr(z) = \Pr(\text{true zero}|z) = \frac{\eta f_0(z; \kappa)}{f(z)}. \tag{5.3}
\]

Typically, a (partial) correlation is not significant if \( \text{fdr}(z(\pi_{Y,Z}) / |\pi_{Y,Z}|) > 0.2 \).

\[ \kappa = 0.19, \tau_0 = 0.925 \]

\[ \begin{align*}
\text{Mixture} & \quad \text{Null Component} \quad \text{Alternative Component} \\
\text{Density} & \quad \text{latanh}(\text{DPCC})
\end{align*} \]

In Fig. 9, the threshold value is depicted by horizontal red dotted lines. The (partial) correlation values whose absolute values are larger than the threshold value are depicted by a grey shaded area. For the DCC on the left panel, the threshold value was larger than one, so all the DCCs were not significant. In contrast, for the DPCC on the right panel, two DPCC values were detected as significant values. Note that the Spearman correlation coefficient between the two regions LV1 and LDPC was –0.08, which is relatively close to zero.

On the other hand, the Spearman correlation coefficient between the two regions LDPC and LTPJ was 0.66, which is relatively large. Fig. A.26 in Appendix A shows the DCC and DPCC values between LDPC and LTPJ. In the figure, the DCC fluctuated severely between –0.5 and 1, but no significant correlation was detected by the FDR procedure. However, the DPCC value varied between 0 and 0.7, which is a smaller range, and many partial correlation values were found significant by the FDR procedure, which agrees with the Spearman correlation coefficient.

Fig. 10 illustrates the connectivity network of the twelve ROIs at a fixed time of \( t = 284 \) s. Two nodes were connected by a red straight line when the DPCC value was significantly different from zero, by the local FDR procedure. Note that the sparse connectivity network can dynamically change over time.

5.2. Sensitivity analysis

During the analysis of the single participant, we noted that the DCC fluctuated relatively largely compared to the DPCC. In this section, we use all the groups of participants to demonstrate that a preprocessing step of the fMRI data can affect the DCC significantly, and that the DPCC is relatively robust to the preprocessing step.

Specifically, we compared the time-varying (partial) correlations estimated in two scenarios: one is the fMRI data preprocessed without motion artifact removal, and the other is the fMRI data preprocessed with motion artifact removal. See Section 2 for the fMRI data preprocessing.
The sparse connectivity network obtained based on the DPCC values at the time point of 284 seconds for a participant of age ‘4yo’ and gender ‘F’. The number on each edge denotes the corresponding DPCC value. Two nodes were connected by a red line if the DPCC value was significantly different from zero according to the local FDR procedure. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The sensitivity of the DCC and DPCC was evaluated by using the root mean squared deviation (RMSD) of the (partial) correlations for each participant and each pair of ROIs. We used five ROIs which are LV1, LdPCC, LdACC, RTPJ, and LTPJ, so that a total of ten pairs of ROIs were considered. The RMSD is defined by

$$RMSD(i, p) = \sqrt{\frac{1}{T} \sum_{t=1}^{T} \left( \frac{r_{ip}^{(t)}}{\bar{r}_{ip}} - 1 \right)^2},$$  \hspace{1cm} (5.4)\n
Fig. 12. The estimated DCC (left panel) and DPCC (right panel) values between two ROIs, LV1 and LdPCC, for a participant of age ‘4yo’ and gender ‘F’. The red dotted line with circles is the DCC or DPCC values obtained from the fMRI data before performing motion artifact removal (smoothing) in the preprocessing step, and the black straight line is the DCC or DPCC values obtained from the fMRI data after performing smoothing. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 13. The root mean squared deviations of the RMSD values in Eq. (5.4) of the correlation (DCC, white box) and the partial correlation (DPCC, blue box), for all participants. The horizontal axis denotes each of the ten ROI pairs. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

where $i$ denotes a participant, $p$ denotes a pair of ROIs, and $T$ denotes the total number of observations (scans).

Fig. 12 illustrates the estimated DCC and DPCC values of the two ROIs, LV1 and LdPCC, for a participant of age ‘4yo’ and gender ‘F’. We can find a large fluctuation in the DCC on the left panel, but relatively small fluctuations in the DPCC on the right panel, in the case of the data with motion artifact removal, which is denoted by ‘smoothing’. On the other hand, both the DCC and DPCC values were relatively close to zero in the case of the data without motion artifact removal, i.e., before smoothing.

Fig. 13 shows the boxplots of the RMSD values of the DCC and DPCC, respectively. It is clearly shown that the RMSD of the DCC
is much larger than that of the DPCC, while there are almost no differences between different ROI pairs.

6. Discussion

Partial correlation has been well known as an effective method for brain functional connectivity analysis in fMRI studies. However, a strict assumption such as the multivariate normal distribution prohibited its application to neuroimaging data analysis. Moreover, the time-varying nature of the functional connectivity was discovered only recently, which can be crucial for analysis of real fMRI data.

The proposed time-varying partial correlation was developed based on the copula-DCC-GARCH model, which is an extension of the DCC-GARCH model to the copula framework. Therefore, the proposed method takes advantage of the partial correlations, the copula models, and the DCC-GARCH models. In this study, we conducted extensive numerical simulations to demonstrate the effectiveness of the proposed method by using three-dimensional time series data generated under various conditions of correlation functions and marginal distributions. The simulation results clearly showed the relative strengths of the proposed method compared to the other methods including the sliding window correlations, sliding window partial correlations, and the copula-DCC-GARCH model. For example, compared to the copula-DCC-GARCH model, the proposed method could eliminate spurious correlations and discover the conditional independence relationships accurately from three-dimensional time series data.

Typical fMRI studies require the prior selection of a few brain areas. However, it is often difficult to select suitable brain areas due to a lack of expert knowledge. Since the partial correlation removes unintended influence exerted by other brain areas, it is instead recommended to include as many brain areas as possible for studies of functional connectivity. However, the number of brain areas included in the partial correlation analysis cannot exceed the total number of observed time points. This is because a partial correlation is directly related with an off-diagonal element of the inverse correlation matrix, and the sample correlation matrix becomes a singular matrix (i.e., non-invertible) when the dimensionality is larger than the number of observations (Whittaker, 1990). The same argument also holds for the case of time-varying correlation of the copula-DCC-GARCH model in (3.8) and (3.9). Moreover, a larger number of brain areas included can lead to decreased accuracy of estimation due to a larger variance of the distribution of sample partial correlation coefficients (Hotelling, 1953). For these reasons, we considered total 12 ROIs in this study. We considered ROIs which have been analyzed in the previous studies (Jacoby et al., 2016; Richardson et al., 2018). In addition, we included ROIs on both of the left and right hemispheres, and we added the brain areas LITG and RITG considering the spatial distribution of the selected brain regions. In order to perform an extensive whole-brain connectivity analysis including a large number of ROIs, we recommend to incorporate a regularized estimation of large-scale (time-varying) covariance matrices, such as the shrinkage estimation methods (Kastner, 2019; Schäfer and Strimmer, 2005).

In this study, the sensitivity of the proposed measure of dFC was analyzed by using the whole group of participants. It is evident that the proposed time-varying partial correlation is more robust to the standard preprocessing step for motion artifact removal. We note that the motion artifact removal procedures have been suggested in recent studies (Power et al., 2012; Satterthwaite et al., 2013), but their robustness has not been studied for a wide variety of functional connectivity measures. We remark that the preprocessing step tested in this study is limited and relatively simple. There are more researches on the effect of preprocessing steps for dFC analysis of fMRI time series. Vergara et al. (2017) investigated the effect of preprocessing steps such as motion regression, spatial smoothing, group independent component analysis, and despiking for the classification of mild traumatic injury, where dFC was estimated using sliding window correlation. Xu et al. (2018) examined the impact of global signal regression on the performance of sliding window correlation based on resting state fMRI data. Nalci et al. (2019) found that nuisance regression such as the global signal regression could not eliminate the correlation between the nuisance regressors and dFC estimates.

For the real fMRI data analysis, we adopted a two-step procedure for inferring sparse connectivity networks, which consists of the estimation of time-varying partial correlations and the pruning of spurious correlations. Inferring sparse connectivity networks is crucial for giving a useful interpretation, so a reliable statistical method is very important. Schäfer and Strimmer (2005) compared the two-step procedure with the one-step procedure using Lasso regression by simulation experiments. And the simulation results demonstrated that the two-step approach produced more sparse networks while controlling their positive predictive value (PPV) above the desired level 1 − fdr > 0.8. Although their simulation experiments considering static connectivity should be modified suitably to the case of dynamic connectivity, the results sufficiently demonstrated that the two-step procedure is quite promising for producing a statistically reliable and meaningful sparse connectivity structures.

7. Conclusion

In this study, we proposed a time-varying partial correlation based on the copula-DCC-GARCH model. Under the assumption that the brain functional connectivity (FC) is not static, the proposed method can be an efficient alternative to the existing methods. The proposed method successfully incorporated the partial correlation coefficient into the copula-DCC-GARCH framework. The partial correlation coefficient has been used in studies of FC for reducing spurious correlations from other brain areas. We explained the recursive algorithm for computing the time-varying partial correlation efficiently for a large number of brain areas. In summary, the proposed method takes advantage of the partial correlations, the copula models, and the DCC-GARCH model, so it provides an improved approach to dFC analysis of fMRI data.

(1) Compared to the sliding window techniques, the proposed method does not need to choose the window length due to the property of the DCC-GARCH model.

(2) Since the proposed method is based on the copula model, it is suitable for fMRI data analysis where the marginal distributions are mostly skewed non-normal distributions. Note that the partial correlation could not be applied in the cases of skewed non-normal distributions due to its multivariate normal distribution assumption.

(3) The proposed method is robust to artifacts in fMRI data. Compared to the pairwise correlation-based methods including the copula-DCC-GARCH model, the partial correlation-based methods can eliminate the effect of external signal and measure the conditional independence without spurious correlations.

(4) The proposed method is robust to standard preprocessing steps for fMRI such as spatial smoothing or motion artifact removal, which is due to the property of the partial correlation.

We conducted extensive numerical simulations to compare the proposed method with other methods for estimating dynamic correlations under various conditions of correlation functions and marginal distributions. The other methods compared in the simu-
luation were the sliding window correlation, sliding window partial correlation, and the copula-DCC-GARCH model. The simulation results demonstrated the effectiveness of the proposed method in comparison with the other methods. Lastly, through group fMRI data analysis, we showed that the proposed method is robust to a standard preprocessing step used for motion artifact removal.

We showed that the DFC inferred by the proposed method can be visualized by sparse networks of brain regions. Standard approaches such as the FDR procedure for multiple hypothesis testing were used for the inference of large-scale network structures. The inferred brain networks demonstrated potential application of the proposed method to clinical purposes.

Acknowledgement

This study was supported by a National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No.2017R1C1B5076912).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://doi.org/10.1016/j.neures.2020.06.006

References