

Dynamic effective connectivity network based on change points detection

Le Zhao, Weiming Zeng^{*}, Yuhu Shi, Weifang Nie

Lab of Digital Image and Intelligent Computation, Shanghai Maritime University, Shanghai 201306, China

ARTICLE INFO

Keywords:

Alzheimer's disease (AD)
Dynamic brain connectivity
Fused lasso regression
Change points detection
Granger causality analysis

ABSTRACT

Human brain networks can be modeled as a system of interconnected brain regions which are recorded by time-dependent observations with functional magnetic resonance imaging (fMRI). In order to spot trends, detect anomalies, and interpret the temporal dynamics, it is essential to understand the connections among distinct brain regions, and how these connections evolve over time. However, the change points of dynamic reorganization in brain connectivity are unknown, which may occur frequently during the scanning session. In this paper, we introduce a fused lasso regression approach to detect the number and position of rapid connectivity changes of subject and subsequently estimate the brain effective connectivity networks within each state phase lying between consecutive change points by conditional Granger causality method from fMRI time series data. The performance of the method is verified via numerical simulations and the obtained classification accuracy with support vector machine (SVM) was 86.24% in 140 subjects from Alzheimer's Disease Neuroimaging Initiative (ADNI). Compared with static EC model and conventional dynamic EC model based on sliding window technique, the experimental results show that the fused lasso achieved better classification effect, which probably due to better dynamic description. The result shows that the dynamic effective connectivity based on change points detected by fused lasso method is a better feature for classification.

1. Introduction

In recent years, functional magnetic resonance imaging (fMRI) has been widely used in clinical and scientific research as a fast, non-invasive and repeatable imaging technology, which measures the functional activity of brain neurons based on blood oxygen level dependent (BOLD) effect [1]. Functional brain network (FBN) consists of brain regions and connections that reflect dependency between intrinsic BOLD signals in distributed regions. Namely, nodes of the network are modeled as regions of interests (ROIs) and edges are quantified as functional connectivity (FC) or effective connectivity (EC) [2]. Functional connectivity measures the time-domain correlation between spatially distant neurophysiological signals, that is, the existence and strength of connectivity relationship between different brain regions. It is essentially a statistical notion, that is usually evaluated through correlation or coherence analysis. However, the FBN constructed by FC is an undirected graph which may not adequately reveal the causal effects of neural activity among brain regions. Effective connectivity measures the information transfer pattern of neuronal interactions, which is closer to the real brain function mechanism. The common methods to compute EC conclude Granger causality analysis (GCA), Bayesian networks (BN),

dynamic causal modeling (DCM) and transfer entropy [3–5]. Khazaee et al. [6] found that the directed network formed by effective connectivity can obtain higher classification accuracy compared with the undirected network formed by functional connectivity in identifying mild cognitive impairment (MCI) and AD. Studies have shown that brain network has become a biological marker of cognitive research and disease prediction [7–9].

Traditional FBN studies assume that brain networks are temporal stationary throughout the fMRI scan. However, recent studies show that FBN changes dynamically with time, and the time-varying characteristics of FBN contain a lot of useful information [10,11]. The brain is constantly variable even at rest and show moment-to-moment (second apart) changes in connectivity. It is believed that information on the temporal dynamics of brain connectivity changes in both strength and direction contributes to a more integrated and comprehensive understanding of the functional organization of the brain. Indeed, studies have shown that quantification of dynamic connectivity offers potential value for better diagnosis, prognosis, and even treatment in physical and mental disorders such as AD [12]. For example, Rangaprakash et al. [13] integrated static and dynamic EC modeling with strength and variance of directional connectivity to identify disease foci and associated paths.

^{*} Corresponding author at: College of Information Engineering, Shanghai Maritime University, Shanghai 201306, China.

E-mail address: zengwm86@163.com (W. Zeng).

Fu et al. [14] explored the potential association of static and dynamic functional network connectivity abnormalities with AD and SIVD and discovered dynamic FBN to be a more important biomarker since its progressively altered patterns can better track cognitive impairment.

In practical experiments, low signal-to-noise ratio, noise caused by non-neural interference (e.g., physiological noise from heartbeat and respiration, and machine instability) can cause changes in the BOLD signal over time, leading to changes in network connectivity metrics. Also, due to the nature of psychological processes, particularly resting-state fMRI without any experimental design, the duration of FBN states and the exact timing of changes are habitually unknown [15]. Meanwhile, evidence indicated that the dynamics of brain networks in task-state fMRI are usually not entirely governed by the boundaries of task events/blocks [16,17]. Thus, there is a great need for establishing an efficient method for mining the dynamic behavior of brain networks that does not require a priori knowledge. In recent years, the exploration of algorithms based on the dynamic connectivity of fMRI data and their applications in diseases have developed rapidly, including sliding window method [18,19], time-varying vector autoregressive models (TV-VAR) [20,21], and dynamic connectivity detection (DCD) algorithm [22]. However, these methods have their own shortcomings. The step size and window length are key parameters affecting the sliding window technique, and their selection will have a significant impact on the analysis results. TV-VAR model develops a method that can identify changes of causal influence (directed) connectivity, but it needs additional clustering algorithms to get the change points and it does not account for the temporal structure. DCD model explores a time-varying sparse covariance matrix [23], revealing dynamic interdependent networks of entities. However, it fails when brain activity patterns change too frequently, and the resulting network is undirected.

The variability of undirected networks in brain regions has been frequently addressed in past studies [24,25], while dynamic effective connectivity has been less explored. In this work, we propose a data-driven change points model to automatically assess the number and location of temporal changes. The time-varying effective networks are obtained by detecting changes in brain networks over time at the individual level through connectivity, i.e., finding time points when most or all unidimensional connectivity signals change together in some specific way. In detail, first, the connectivity metric at each time point was estimated from fMRI data of each subject, and the change points of the connectivity time series were detected under the fused lasso method [26]. Second, we used conditional GCA [27] to infer directed brain networks for state phases specified with the estimated adjacent change points. The performance of the proposed method was evaluated through extensive simulation studies that involved dynamic networks with various changes in connectivity. Finally, to further demonstrate its validity, the values of the dynamic effective connectivity network constructed by resting-state fMRI data from Alzheimer's disease (AD) patients and healthy controls (HC) are applied as the original feature set for feature selection, and then classified using support vector machines (SVM). The results showed that the fused lasso method achieved better classification results than static and sliding time window methods, indicating that the dynamic effective connectivity based on change points detected by fused lasso method is a better feature for classification.

2. Materials and methods

To characterize changes in brain networks, dynamic connectivity may be a scalable brain measure that is accessible to neuroscientists and informatics researchers alike. In this paper, we first figure out the change points of brain network connectivity for each subject, and then estimate segmented effective connectivity networks [28]. It means that the brain is in an approximately consistent connectivity state during the corresponding time interval.

2.1. Change points estimation

In order to determine the change points of brain networks during fMRI data acquisition, we describe a time-series pattern of connection which means a time-varying connectivity vector between any two brain regions at each time during the experimental period. These pairwise correlated temporal processes are utilized to identify the change points of the network. As a result, the multidimensional fMRI time series is managed to be divided into diverse segments according to significant changes in connectivity. The brain network is approximated within each segment, with significant differences between adjacent segments.

2.1.1. Connectivity Metric constructed by fMRI subsequences

The fMRI data of a single subject is described by a matrix, with the number of rows T being the number of sampling points and the number of columns p referring to the number of brain regions studied. Each ROI corresponds to the average BOLD signal of all voxels in the brain region. Each column represents the fMRI observations of neural activity for a ROI, and each row represents the observation of all ROIs at each time point. Then, a series of overlapping fMRI subsequences are generated with the length of h (that is, h observations). As a result, $n = \lfloor (T-h)/s \rfloor + 1$ overlapping fMRI subsequences with step size s were generated. Each fMRI subsequence is denoted as a matrix $X_\tau = [X_{\tau 1}, X_{\tau 2}, \dots, X_{\tau p}] \in \mathbb{R}^{h \times p}$, $\tau \in \{1, \dots, n\}$.

It is assumed that the whole scanning process contains $K+1$ different state phases spaced by K unknown brain network change points with $1 < a_1 < \dots < a_{K-1} < a_K < T$. The k -th state stage consists of time interval between a_{k-1} and a_k , where the brain network is relatively stationary. The set of n subsequences' centroids is denoted as $\{\tau_1, \tau_2, \dots, \tau_n\}$. For each subsequence, the Pearson correlation coefficient matrix of the brain regions at the center time point τ is expressed as $\tilde{\rho}_\tau = \text{cor}(X_{\tau i}, X_{\tau j}) = \frac{\text{cov}(X_{\tau i}, X_{\tau j})}{\sigma_{x_i} \sigma_{x_j}} \in \mathbb{R}^{p \times p}$, $i, j \in \{1, \dots, p\}$, $\tau \in \{\tau_1, \dots, \tau_n\}$. The correlation profile/connectivity from time points 1 to τ_1 , and from τ_n to T is expanded and equal to correlation profile at time 1 and T , respectively. The values of lower triangle in ρ_τ (size: $p \times p$), $\tau \in \{1, \dots, T\}$ at each time point are drawn into a row vector to form a matrix Y (size: $T \times \frac{p(p-1)}{2}$), namely, the connectivity metric. Each column of Y represents a collection of a connection across all time points during an experiment and each row represents all connections at a time point. Then, we approximate the multivariate pairwise dependent time series via a piecewise constant function under a fused lasso approach to detect temporal changes in connectivity. These change points are the times at which there are significant differences in the connectivity metric. The method divides the scanning period into distinct state phases, so that the connectivity is constant within a state phase but changes across these phases.

2.1.2. Group Fused Lasso for detecting change points in brain connectivity

For a linear regression model (1), let $y \in \mathbb{R}^N$ be the response vector, $X \in \mathbb{R}^{N \times p}$ be the design matrix, $\beta \in \mathbb{R}^p$ be the regression coefficient vector, $\varepsilon \in \mathbb{R}^N$ be the residual vector that conforms to an independently and identically distribution $\varepsilon_n \sim \mathcal{N}(0, \sigma^2)$, $n \in \{1, 2, \dots, N\}$.

$$y = X\beta + \varepsilon \quad (1)$$

Considering the order between variables, Tibshirani et al. [29] proposed a fused lasso model, which not only sparsely limited the regression coefficient, but also the continuous difference of the regression coefficient of adjacent variables. Therefore, it could not only obtain the sparse solution of the regression coefficient, but also make the adjacent regression coefficient change smoothly. The fused lasso solution of the linear regression model in Equation (1) can be expressed as

$$\hat{\beta} = \underset{\beta \in \mathbb{R}^p}{\text{argmin}} \frac{1}{2} \|y - X\beta\|_2^2 + \lambda_1 \|\beta\|_1 + \lambda_2 \sum_{p=2}^p |\beta_p - \beta_{p-1}| \quad (2)$$

Where $\lambda_1 \geq 0, \lambda_2 \geq 0$ are the regularization parameters, β_p is the p -th element of vector β . In Formula (2), the first term is the residual term, ensuring that the error between the response vector represented by design matrix and the original response vector is as small as possible. The second term with $L1$ norm penalty function is the sum of the absolute values of the coefficients, so that the coefficient with a small absolute value is automatically compressed to zero, achieving the purpose of variable selection and coefficient sparsity. The third term $\sum_{p=2}^P |\beta_p - \beta_{p-1}|$ is called fusion penalty, which smoothly changes the regression coefficients of adjacent variables, that is, penalizes the absolute value of the difference between the regression coefficients corresponding to adjacent variables, so that the solution of the fused lasso has the characteristics of piece-wise constant. Obviously, its effect is to make the adjacent regression coefficients almost equal, so the fused lasso can be supposed to have an automatic group effect [30].

In the real data, $p > 2$ ROIs were selected to extract fMRI data for each subject. This results in m ($m = \frac{p(p-1)}{2}$) pairwise correlated connectivity time series, which are interdependent and synergistic to form a dynamic brain network. The connectivity metric Y was modeled as a piecewise-constant signal with each column denoting the time-varying connectivity of two brain regions and locations of change points were assumed to be shared across connections. This makes sense because we focus on changes in brain networks, where partial or all connections change regularly and consistently. The (non-convex) jump numbers were replaced by the (convex) total variation, penalizing the connectivity metric to approximate a segmented constant signal with multiple change points, i.e., solving

$$\min_{U \in \mathbb{R}^{T \times m}} \|Y - U\|^2 \text{ subject to } \sum_{i=0}^{T-1} \delta(U_{i+1,\cdot} - U_{i,\cdot}) \leq k \quad (3)$$

Using the Lagrange multiplier method, the combinatorial optimization problem can be relaxed to a convex optimization problem that contains a quadratic error criterion whose penalty is the sum of the euclidean norm of multidimensional increments, i.e., to consider:

$$\min_{U \in \mathbb{N}^{T \times m}} \frac{1}{2} \|Y - U\|^2 + \lambda \sum_{i=0}^{T-1} \|U_{i+1,\cdot} - U_{i,\cdot}\| \quad (4)$$

The first term in (4) measures the error between observed connectivity metric and solution $U \in \mathbb{N}^{T \times m}$ which is a piece-wise constant matrix for a given $\lambda > 0$. The second term controls the cumulative increment of all m correlations profiles at continuous times, where the element $U_{i+1,j} - U_{i,j}$ refers to difference of the j -th correlation between time point i and $i + 1$. When the change of multivariate connection time series is not significant, the increment is zero, otherwise the increment takes non-zero values. When i is regarded as a change point of network, it indicates that one or more connections in the connectivity change significantly at this time. Obviously, penalty parameter λ influences the number of change points. Many increment vectors $U_{i+1,\cdot} - U_{i,\cdot}$ will be enforced to shrink to zero as λ increases. In conclusion, the solution of (4) offers an approximation of Y by a piecewise-constant matrix $U \in \mathbb{N}^{T \times m}$ which share change points. The addition of regularization norm to U in (4) does not affect the detection location of the change points. For a fixed value of λ , the resulting optimization problem is reformulated as a fused lasso regression problem, which will facilitate the implementation by block coordinate descent or group fused least angle regression (LARS) algorithms, either exactly or approximately [31]. After that, dynamic programming strategy was used to calculate the minimum sum of the squared errors (SSE) from the subset of all possible change points. The curvature of the SSE curve is examined to determine the optimal number whose second-order derivatives do not rise to a certain threshold when more change points are added. Typical threshold values are chosen to range from 0 to 0.5. The smaller the threshold, the higher the tendency to retain more change points.

2.2. Dynamic brain network estimation based on GC

The strength and direction of the connections between ROIs were calculated using multivariate conditional GCA [32] to obtain EC matrices that form a directed dynamic brain EC network for each segment. Node is considered as each brain region associated with a time series while edge represents a direct effect between two different regions. Granger causality model is an analytical method based on vector autoregressive models that predict the value of the current time series by a linearly weighted combination of the values of the past time series. Therefore, it can only be applied to the causality test of time series data, and cannot be used between variables with only cross-section data. No prior knowledge is required, and the emphasis is on the temporal order of the data as they interact with each other. The directionality of information transfer between brain regions or neurons can be reflected in the analysis of brain networks. In general, for two stationary time series x_t and y_t ($t = 1, 2, \dots, T$), if the combination of the historical information of x and y is more conducive to the prediction of the current value of x than only using the historical information of x , that is, the variance of residuals is reduced, then y is considered to be the Granger cause of x . $VAR(q)$ and joint $VAR(q)$ model are defined as

$$x_t = \sum_{i=1}^q a_{1i} x_{t-i} + \varepsilon_{1t} \quad (5)$$

$$x_t = \sum_{i=1}^q a_{2i} x_{t-i} + \sum_{i=1}^q b_{2i} y_{t-i} + \varepsilon_{2t} \quad (6)$$

The order q , i.e., the maximum number of lagged samples, is selected with Akaike information criterion (AIC), Bayesian Information Criterion (BIC), or according to requirements in the empirical considerations of the specific application [13]. a_{1i}, a_{2i}, b_{2i} are the model coefficients, and $\varepsilon_{it}, i = 1, 2$ are stochastic error term which are not correlated with time. The optimal coefficient set to minimize the model error is estimated by multivariate least squares. The Granger causality effect of y on x is defined by the following formula:

$$F_{y \rightarrow x} = \ln \frac{\text{var}(\varepsilon_{1t})}{\text{var}(\varepsilon_{2t})} \quad (7)$$

Obviously, $F_{y \rightarrow x} \geq 0$. If the past values of the time series y do not contribute to the prediction of x , the regression coefficient $b_{2i} \equiv 0$, $\text{var}(\varepsilon_{1t}) = \text{var}(\varepsilon_{2t})$ and $F_{y \rightarrow x} \equiv 0$. Otherwise, if the past values of y can improve the prediction, $\text{var}(\varepsilon_{1t}) > \text{var}(\varepsilon_{2t})$ and $F_{y \rightarrow x} > 0$. The F-test is performed to determine whether the GC value is significant, that is, whether the difference between prediction error in formula (5) and (6) is meaningful. The prediction error is usually expressed statistically as the residual sum of squares (RSS). Formula (5) is a constrained regression with RSS_R and (6) is an unconstrained regression with RSS_{UR} . The null hypothesis is $H_0: b_{21} = b_{22} = \dots = b_{2q} = 0$, which means the lagging term of y is not part of this regression. To test this hypothesis, use F-test with degrees of freedom q and $(T - 2q - 1)$, i.e.,

$$F = \frac{(RSS_R - RSS_{UR})/q}{RSS_{UR}/(T - 2q - 1)} \quad (8)$$

If the value of $F > F_\alpha$ which is calculated at the selected significance level α , the null hypothesis is rejected. In this case, the lagged term of y belongs to this regression, indicating that y G-causes x .

The above binary Granger causality model obviously cannot distinguish direct causality from indirect causality. Hence, Geweke extended traditional Granger causality to multivariate autoregressive model and added another stationary variable z to calculate the conditional Granger causality (CGC) which is one of the measures of direct effective connectivity. $VAR(q)$ and joint $VAR(q)$ model [33] can be formulated as:

$$x_t = \sum_{i=1}^q a_{3i} x_{t-i} + \sum_{i=1}^q c_{3i} z_{t-i} + \varepsilon_{3t} \quad (9)$$

$$x_t = \sum_{i=1}^q a_{4i} x_{t-i} + \sum_{i=1}^q b_{4i} y_{t-i} + \sum_{i=1}^q c_{4i} z_{t-i} + \varepsilon_{4t} \quad (10)$$

Conditional GC from y to x conditional on z (given z) is:

$$F_{y \rightarrow x|z} = \ln \left(\frac{\text{var}(\varepsilon_{3t})}{\text{var}(\varepsilon_{4t})} \right) \quad (11)$$

Similarly, if the Granger causality from y to x is fully mediated by z , $b_{4t} \equiv 0$ and $\text{var}(\varepsilon_{3t}) = \text{var}(\varepsilon_{4t})$, so $F_{y \rightarrow x|z} \equiv 0$. When the modulation effect of z is considered and y to x still has a directed influence, $\text{var}(\varepsilon_{3t}) > \text{var}(\varepsilon_{4t})$ and $F_{y \rightarrow x|z} > 0$. The statistical significance of causality can also be estimated by F test.

2.3. Feature selection and classification

Feature selection is the process of selecting some of the most effective features from the original features to reduce the sample dimensionality, and is an important means to improve the performance of learning algorithms. Here, tree model, an embedded feature selection algorithm, with little operation cost and suitable for automatic processing was used to screen the features so as to reduce the number of features in the model [34]. The idea of using gradient boosted decision tree (GBDT) to assess the importance of features is relatively simple. Generally, the closer a feature to the top node in the tree structure, the more important the feature is. After comparing the contribution between different features, we get a feature set which is sorted from the most important to the least important.

The SVM algorithm with radial basis function kernel (RBF) is used to verify the dynamic brain effective networks modeling method based on fused lasso with an optimal subset of features. SVM uses kernel method to find a new space that is more favorable for nonlinear classification task. In the current experiment, we conducted 10-fold cross-validation strategy for robust classification and repeated 100 times. Classification performance of different kinds of features was reported by average accuracy.

3. Results

3.1. Simulation studies

A series of simulation experiments were conducted to evaluate the capacity of fused lasso method in capturing network dynamics. We adopted simulation settings inspired by those found in previous papers on dynamic connectivity detection (DCD) algorithm [22,35,36]. Below is a brief description of each simulation study, which is repeated 20 times. p and T represent the number of variables and the length of the time series, respectively. The heat maps in Figs. 1–3 displayed the true relations between the variables (that is, covariance matrices). For

detailed values of the covariance matrices, please refer to [supplementary materials](#). Description of Simulations 1–4: No network change point appears in the simulation time and the data is generated by white noise with $p = 15$ and $T = 150$; One network change point appears at time 150 with $p = 10$ and $T = 300$; Two change points appear at times 200 and 400 with $p = 15$ and $T = 600$; Three change points appear at times 500, 1000, and 1500 with $p = 20$ and $T = 2000$.

Fig. 4 illustrates the result of Simulation 1–4 where the x -axis stands for the length of time series, and the y -axis represents the number of simulations. The blue vertical lines denote the actual change points of the network in each simulation setting, and the solid circle expresses the position of detected change points with fused lasso method. It is demonstrated that the transition points of network structure can be found relatively precisely, verifying the effectiveness of the method. Although the existence of some small differences, change points are detected around the ground truth.

The result of Simulation 1 is shown in Fig. 4(A). When the network structure remains unchanged, the fused lasso method gets the correct result without detecting the false change points. In fact, when the time series is shorter (such as 50), the correct result can still be obtained. The result of Simulation 2 is shown in Fig. 4(B) where there is a change point of network at time 150. The mean value of the change points detected in 20 simulations is 149, which is close to the real value of 150, indicating the effectiveness of the method presented in this paper. The result of Simulation 3 which contains two change points at times 200 and 400 is shown in Fig. 4(C). The mean values of the change points detected in the 20 simulations are 205 and 397. However, it was noted that the second change point was not detected in one simulation, which maybe because the change of connectivity around time point 400 did not reach the threshold and was missed. The result of Simulation 4 is shown in Fig. 4 (D) with change points at times 500, 1000, and 1500. The mean values are 501, 1004 and 1495, respectively. In fact, the longer the time of state phase is, the closer the detection effect of the numbers and positions of change points are to the actual situation. As a whole, fused lasso method can judge out the true transition points of the dynamic network structure quite accurately.

Due to the same purpose and similar simulation data, we compared the proposed method with the DCD algorithm [22]. The results showed that DCD method also achieved relatively good detection effect of change points, as shown in [Supplementary Fig. 3](#). However, it generates many spurious checkpoints over the stationary time periods, especially when the number of variables increases.

3.2. Application to experimental data

3.2.1. fMRI data and preprocessing

The resting-state fMRI data analyzed in current study was selected from public dataset published by Alzheimer's Disease Neuroimaging

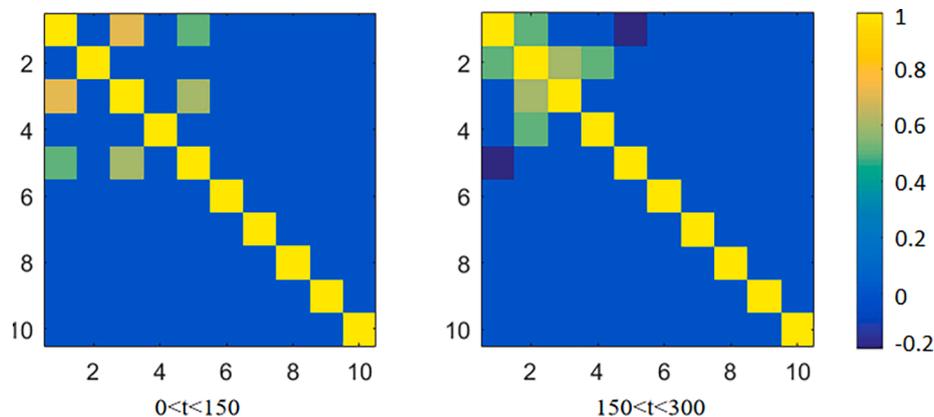


Fig. 1. The covariance matrix applied for the two segments in Simulation 2.

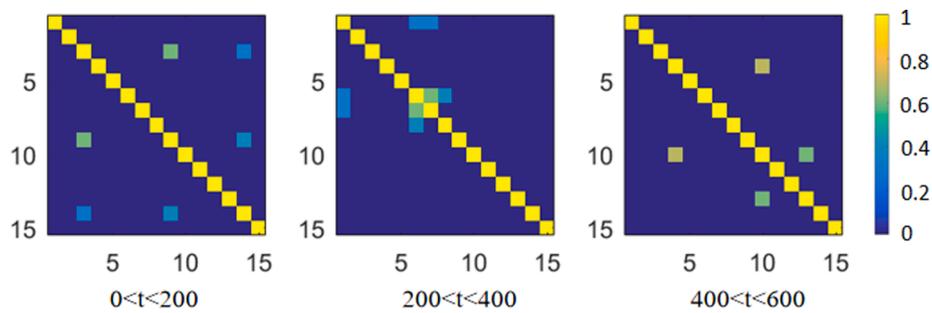


Fig. 2. The covariance matrix applied for the three segments in Simulation 3.

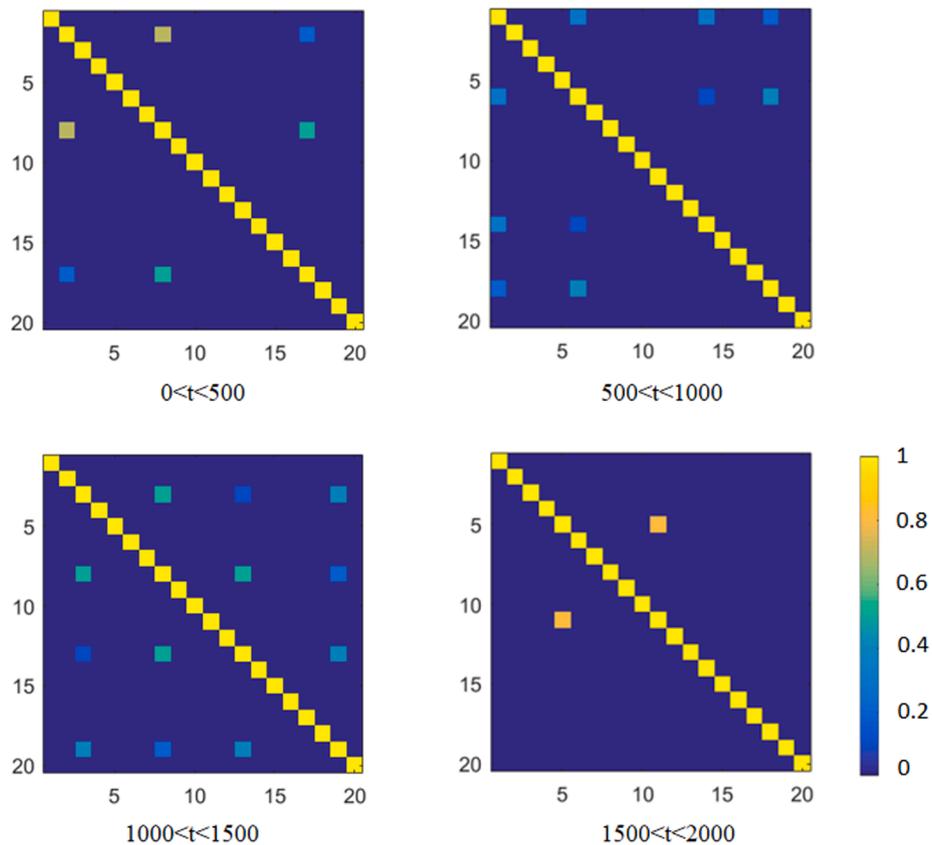


Fig. 3. The covariance matrix applied for the four segments in Simulation 4.

Initiative (ADNI 2) Program [37], including 76 HC subjects and 67 patients with Alzheimer's disease. All subjects were asked to remain relaxed but not to fall asleep during the scan. Images were collected by an echo planar imaging (EPI) sequence on 3.0 Tesla scanner with following parameters: flip-angle (FA) = 80°; repetition time (TR) = 3000 ms; echo time (TE) = 30 ms; imaging resolution/slice thickness = 3.31 mm; slices = 48. A total of 140 volumes/scans (time points) were obtained for each subject. The demographic and clinical characteristics of all subjects studied in this experiment are displayed in Table 1.

Image preprocessing is performed for all fMRI data with a standard pipeline with Data Processing Assistant for Resting-State fMRI (DPARSF) toolbox [38], including removing the first ten volumes of the functional images for subjects' adaptation to the environment; slice timing to the middle slice; realignment for head movements compensation; normalization to the EPI template in standard Montreal Neurological Institute space; resampling to 3-mm isotropic voxels; spatial smoothing with a 4 mm full width half-maximum Gaussian kernel; detrending. During the realignment step, three AD subjects exhibiting excessive head motion

(>3 mm in translational head movement and/or 3° of rotational head movement) were excluded from further analysis. The white matter (WM), cerebrospinal fluid (CSF) signal and six rigid-body parameters about head movement were regressed out as nuisance covariates to reduce non-neuronal BOLD fluctuations and the effects of motion.

3.2.2. Construct dynamic brain network

To construct the causal interaction between the resting state fMRI time series of ROIs for each subject, a 14-node directed graph was constructed. The 14 abnormal functional areas associated with AD are defined as ROI in this paper corresponding to the Anatomical Automatic Labeling (AAL) template which are generally recognized by the academia [39], and are summarized as a set as follows (see Table 2): AAL_ROIs = {Amygdala_R, Calcarine_R, Cuneus_R, Frontal_Inf_Oper_L, Frontal_Inf_Tri_R, Fusiform_L, Lingual_L, Occipital_Sup_L, Olfactory_R, Parietal_Inf_L, Parietal_Sup_R, Postcentral_R, Temporal_Pole_Mid_R, Thalamus_L}. The masks of corresponding ROIs were extracted with wfu_pickatlas_3.0.5b toolbox and the mean time series was attained via

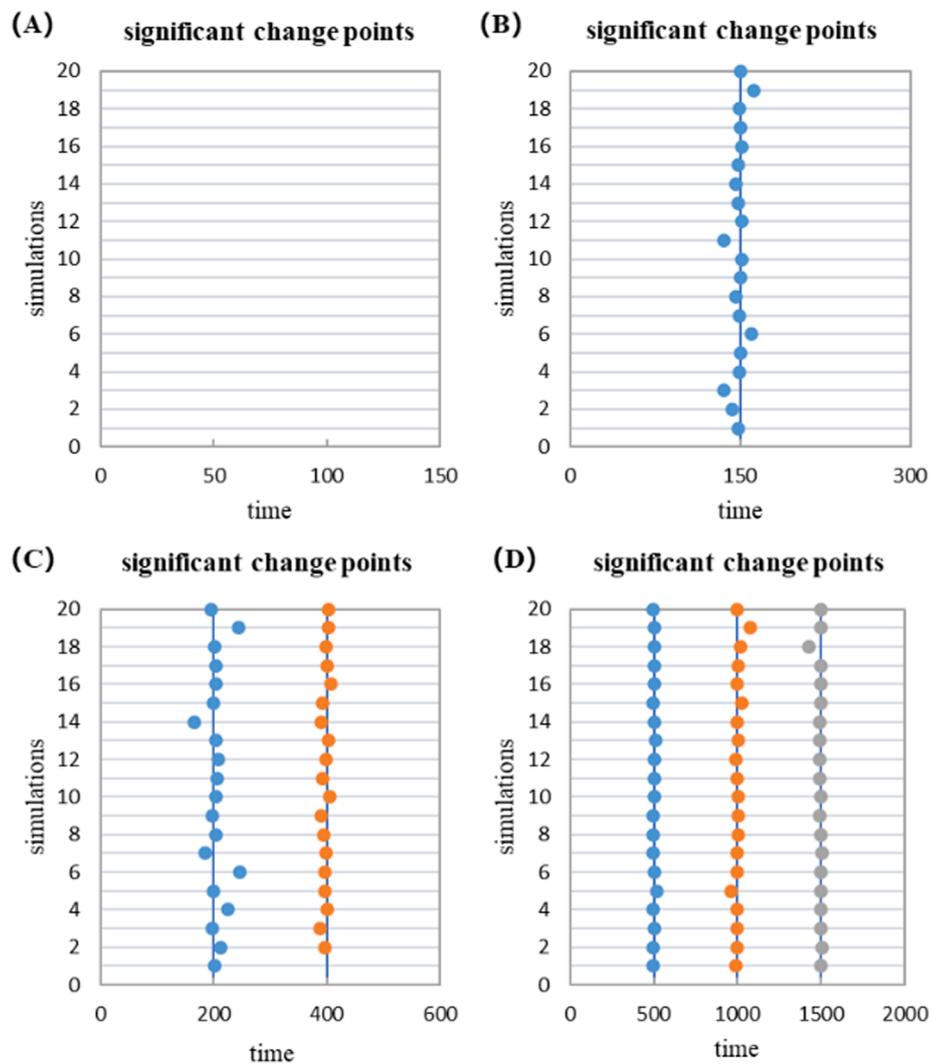


Fig. 4. The results of Simulation 1–4. The circles represent the important change points of dynamic networks found in Simulation 1–4 which is shown in (A) - (D), respectively.

Table 1

Demographic and clinical information of two groups of subjects from ADNI database. MMSE: Mini-Mental State Examination. The values are signified as mean \pm standard deviation.

Information	Healthy Controls (HC)	Patients with AD
Number	76	67
Age (years)	(71–87)	(71–87)
	77.71 \pm 5.7	77.30 \pm 8.53
Sex (male/female)	35/41	32/35
MMSE	28.74 \pm 1.45	21.29 \pm 3.94

averaging signals of all voxels within ROIs, thus generating a representative signal for each ROI. The Unit Root test shows that all time series are stationary, largely depending on the resting state data we use.

There are 14 ROIs resulting in 182 (14 \times 13) directed edges. The mean static effective connectivity matrices were calculated separately for HC and AD group, and significant group differences were marked with asterisks as is shown in Fig. 5. The x-axis and y-axis represent the brain regions, and ij -th element in the matrix represents the quantitative Granger causality from the j -th ROI to the i -th ROI. Compared with HCs, AD patients exhibited significantly increased directed influence from Frontal_Inf_Oper_L to Parietal_Inf_L, from Parietal_Sup_R and Temporal_Pole_Mid_R to Postcentral_R, and decreased effective connectivity from

Table 2

The abnormal brain regions of the Alzheimer’s disease group compared to the healthy control group.

Labels	ROI	Name	Anatomical classification
1	Right	Amygdala	Temporal
2	Right	Calcarine fissure and surrounding cortex	Occipital
3	Right	Cuneus	Occipital
4	Left	Inferior frontal gyrus, opercular part	Prefrontal
5	Right	Inferior frontal gyrus, triangular part	Prefrontal
6	Left	Fusiform gyrus	Temporal
7	Left	Lingual gyrus	Occipital
8	Left	Superior occipital gyrus	Occipital
9	Right	Olfactory cortex	Prefrontal
10	Left	Inferior parietal, but supramarginal and angular gyri	Parietal
11	Right	Superior parietal gyrus	Parietal
12	Right	Postcentral gyrus	Parietal
13	Right	Temporal pole: middle temporal gyrus	Temporal
14	Left	Thalamus	Subcortical

Lingual_L to Cuneus_R, Frontal_Inf_Tri_R and Occipital_Sup_L, from Occipital_Sup_L to Fusiform_L and Lingual_L, from Cuneus_R to Frontal_Inf_Oper_L, from Frontal_Inf_Tri_R and Fusiform_L, from Postcentral_R to Amygdala_R.

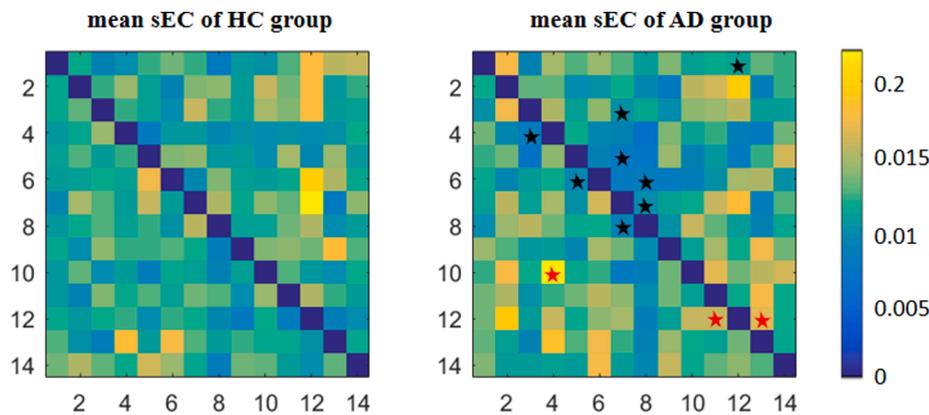


Fig. 5. Mean static effective connectivity (sEC) matrix of HC and AD, respectively. Asterisks denotes significant difference in static effective connectivity as obtained via independent sample *t*-test ($p < 0.05$) where red indicates that the sEC of AD is larger than HC, and black means the opposite.

To further validate the model, dynamic effective connectivity is then determined. The length of overlapping fMRI sub-sequences was set to 5 and the correlation coefficient matrix of brain regions was calculated for each time point. After generating the short-time segmented Connectivity Metric, the fused lasso model is used to find the change points. The segmented brain EC networks were constructed using conditional Granger causality method between adjacent change points, so that a series of dynamic connectivity matrices for each subject could be obtained to extract brain network dynamics. The minimum distance between the candidate change points was set to 30 time points [40], so each subject ended up with 2 to 3 brain network states. One subject was randomly selected in the HC and AD groups respectively, and their static and dynamic effective connectivity networks are displayed in Supplementary Fig. 1.

3.2.3. Classification

Subjects were divided into AD and HC by piece-wise effective connectivity network. Each element in EC matrix is the Granger causality that represents the directed relationship between two brain areas, and it is positive value. Values close to 0 signify that the time series is not the Granger causality of the other. The larger the value is, the more predictive one time series can be to the other time series, that is, one is the Granger causality of the other. We averaged the piece-wise EC matrix of every subject and removed the values in main diagonal since it denotes an area’s predictive influence on itself. Then the remaining part was flattened along the columns and collapsed into a one-dimension vector with dimension $1 \times p(p - 1)$ in which p is the number of related ROIs to retrieve the initial features [32]. Fourteen ROIs were selected, so the procedure resulted in 182 averaged dynamic GC values which consist in the original feature set for the machine learning algorithm.

Gradient boosted decision tree (GBDT) was used for feature selection and feature importance evaluation. The importance of the selected features is sorted. Then, the most discriminative connectivity with high feature importance was input into the RBF_SVM classifier for training under 100 repetitions of the experiment with 10-fold cross-validation. As shown in Table 3, the classification accuracy of static EC, dynamic EC based on fused lasso reached 80.71% and 86.24%, respectively, indicating that these features have certain ability to distinguish between the AD and HC groups. It can be clearly seen that the classification with

Table 3

Classification performance of SVM using features extracted by different methods. FS: feature selection; sFC: static functional connectivity; sEC: static effective connectivity; sw: sliding window.

Method		sFC	sEC	sw-EC	fused lasso-EC
Accuracy(%)	No-FS	72.14	45	50	57.14
	FS	81.67	80.71	82.06	86.24

feature selection is better, especially for the effective connections, which have twice the original number of features than the functional connections. A large number of features may be redundant and contain noise. The window width of the sliding time window was set from 30 to 50, and the EC network in each time window was calculated separately, and then averaged as the classification features for each subject. As shown in Fig. 6, when the window width was 33, the maximum classification accuracy reached to 82.06%. The results of analyses showed that the dynamic EC based on fused lasso improved classification performance, compared with conventional static and window methods.

4. Discussion

There are some difficulties in assessing brain dynamics, especially at rest. Resting state has no explicit tasks designed for state change, so each brain network change point is irregular and dissimilar across runs [41,42]. At this point, we need to explore the relatively stable connectivity state and change points of the brain network for each subject from a data-driven perspective. In this paper, the framework of fused lasso regression was taken advantage to extend the approximation based on total variation to multi-dimensional settings, and detect the change points of time-varying networks constructed by multiple brain regions. The feasibility of the proposed method is validated by simulation experiments. Then, it was applied to fMRI analysis of ADNI dataset and compared with sliding window technique. Inter-group differences between AD and HC were assessed, as well as state transition behaviors, revealing changes in the diseases.

The simulation results proved that the fused lasso was able to

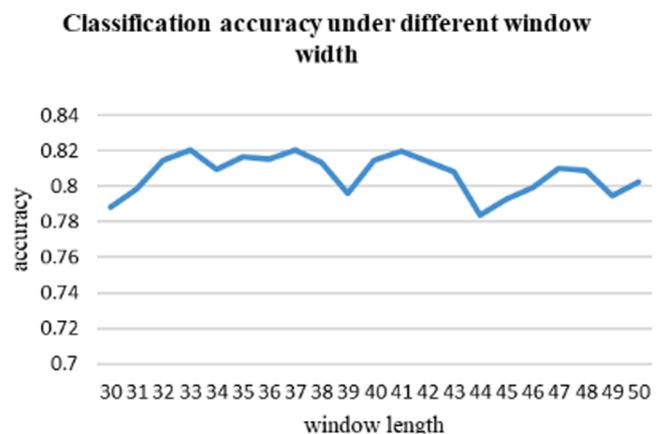


Fig. 6. The classification accuracy of dynamic effective connectivity based on sliding time window method corresponding to different window width.

precisely capture the time-varying connectivity patterns. For state phases of different lengths, the positions found are all near the true transition point. We found that the longer the state phase is, which means the slower the network structure changes, the more accurate the results will be. The difference between the result and the exact value remains within the range of the subsequence length (i.e., 5), indicating the feasibility of the method. But occasionally it will miss some change points as we can see from Fig. 4, this may be because the simulated data are greatly affected by the added noise occasionally and the slope of SSE at that point does not reach the threshold (i.e., 0.5).

Further, the public fMRI data set was used to verify the classification effect of the dynamic EC network constructed by the fused lasso model. The brain network is composed of nodes and edges. The node refers to brain region, which are recorded as one-dimensional fMRI signal, and the edges are represented by the joint interaction of different brain regions. Previous studies of rs-fMRI have discovered that patients with AD dementia have decreased functional connectivity in default network, attention network, executive control network [43,44]. Zhong et al. [45] found that compared with AD patients, the regions in DMN of normal subjects exhibited stronger influential interactions, except that the interactions between medial prefrontal cortex (MPFC) and bilateral inferior parietal cortex (IPC) was weaker than those of AD patients. As demonstrated in Fig. 5, most of the effective connections between abnormal brain regions in AD patients were significantly reduced, indicating that the information flow between brain regions was reduced. There were also three significantly increased connections, all flowing from other regions to the parietal lobe. In patients with AD, there is not only a decrease in information flow in brain regions, but also an increase in information flow related to the parietal lobe, suggesting both impairment (decrease) and compensation (increase) of effective connectivity in patients with AD. However, the properties of dynamic EC are rarely studied.

Finding the time point where brain networks change is an important question in neurophysiological study. Network changing over time actually changes in connectivity which is triggered by dynamic time courses of ROIs. We aimed to discover change points of dynamic network patterns which are constituted by a set of connection sequences which vary over time. That is, we modeled a time-dependent connectivity metric and identified multiple change points shared by multiple connections under a fused lasso approach, in particular, abrupt changes in measurements. The fused lasso is a sparse model that achieves automatic grouping effects by penalizing differences in coefficients. The dynamic network modeling approach based on fused lasso can smooth the connectivity changes of adjacent time points. In current work, we seek to distinguish changes in brain networks in AD patients, and provides an accurate algorithm to interpret time-varying directional interactions among brain regions and classify AD patients with healthy controls by employing a dynamic EC approach based on conditional granger causality and advanced machine learning methods. As shown in Supplementary Figs. 1 and 2, static connections are averaged over the entire scan time, while dynamic connections can capture rich and fine-grained time-varying information. The results show the efficiency of fused lasso model in constructing dynamic EC networks, which can extract more useful features for classification. The change points constructed by the fused lasso method form two or three segmented directed networks since the pre-processed data only has 130 time points, while the sliding time window constructs dozens of continuous networks. Both methods use fMRI data from ADNI to evaluate dynamic EC. However, as shown in Table 3, the accuracy of the former is higher than that of the latter. At the same time, the choice of window size obviously influenced the performance of the sliding window method [46]. Experimental results on 140 subjects in ADNI dataset showed that our proposed method outperformed the sliding window approach, which probably due to better dynamic description. Supplementary Fig. 4 demonstrates the dynamic effective connection matrix of AD, which is averaged by EC of all time points. The change points detection method based on fused lasso

regression revealed 11 significantly different connections between AD and HC. When the window width of the sliding time window was set to 33, there were 14 connections with significant differences between the two groups. Six of the differential connections are the same, indicating that the two methods have certain overlap and commonality, but also have differences. It is noted that the sliding window approach has been utilized in many dynamic studies with significant findings [47]. This also demonstrates the robustness of the change points detection model based on fused lasso method which characterizes the dynamic changes in brain networks more accurately.

Change point model conceptualizes dynamic brain connectivity as a collection of quasi-stable state phases. Corresponding to diverse modulations in the brain, the piecewise constant connectivity states are repeatedly detected to jump at different change points. Although the existing change points model for dynamic connectivity have had some success in explaining temporal changes of brain networks, challenges remained such as only working best in task-based experiments containing multiple subjects. Compared with static EC model and conventional dynamic EC model based on sliding window method, the experimental results show that the fused lasso can characterize the brain network dynamics, and achieve better classification effect. In conclusion, the fused lasso method proposed in this paper builds dynamic effective connectivity network, mining the dynamic changes of information flow in brain regions, effectively preserving the sparsity of the network and the time smoothness of sub-sequences, and improving the classification effect of the algorithm, thus providing help for the diagnosis of brain diseases to a certain extent.

Nevertheless, there are still some aspects that need further study. For example, connectivity of the whole brain network can be studied based on different templates, especially functional partition templates (e.g., 264 Putative functional Areas [48]). In addition, we aim to capture the time-varying properties of brain directed networks. In the future, the conditional Granger causality model can be improved, such as adding nonlinearity and sparsity, so as to obtain more accurate and refined brain networks. Also, fMRI data with task state or longer scan time can be further verified and analyzed.

5. Conclusion

Traditional static FBN studies ignore the rich dynamic time-varying information in the connectivity between brain regions. Recent studies of dynamics have also paid little attention to the dynamic effective connectivity network of the brain. The fused lasso method is a data-driven and computationally efficient approach to address the problem presented in this paper, i.e., temporal change points detection of the multivariate time series of correlations in dynamic brain connectivity research. It divides the brain networks according to the sparsity and smoothness of connectivity changes between ROIs. The conditional Granger causal model estimates a directed graph at each partition/segment, representing the time-varying brain effective network, reflecting the dynamic process of information transmission of neural activity.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grants No.31870979 and 61906117), and Shanghai Sailing Program (Grant No. 19YF1419000).

Data availability statement

Data used in this manuscript is available for download at <http://adni.loni.usc.edu/>.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.bspc.2021.103274>.

References

- [1] K.J. Worsley, K.J. Friston, Analysis of fmri time-series revisited, *Neuroimage* 2 (3) (1995) 173–181.
- [2] K.J. Friston, Functional and effective connectivity: a review, *Brain Connect* 1 (1) (2011) 13–36.
- [3] A. Seth, Granger causality, *Scholarpedia* 2 (7) (2007) 1667, <https://doi.org/10.4249/scholarpedia.1667>.
- [4] K. Friston, R. Moran, A.K. Seth, Analysing connectivity with granger causality and dynamic causal modelling, *Curr. Opin. Neurobiol.* 23 (2) (2013) 172–178.
- [5] L. Barnett, A.B. Barrett, A.K. Seth, Granger causality and transfer entropy are equivalent for gaussian variables, *Phys. Rev. Lett.* 103 (23) (2009), <https://doi.org/10.1103/PhysRevLett.103.238701>.
- [6] A. Khazaee, A. Ebrahimzadeh, A. Babajaniferemi, Classification of patients with mci and ad from healthy controls using directed graph measures of resting-state fmri, *Behav. Brain Res.* 322 (Pt B) (2017) 339–350.
- [7] Y. Li H. Yang B. Lei J. Liu C.Y. Wee Novel effective connectivity inference using ultra-group constrained orthogonal forward regression and elastic multilayer perceptron classifier for mci identification *IEEE Transactions on Medical Imaging* 2018 1 1.
- [8] S. Sulaimany, M. Khansari, P. Zarrineh, M. Daiyanu, N. Jahanshad, P.M. Thompson, A. Masoudi-Nejad, Predicting brain network changes in Alzheimer's disease with link prediction algorithms, *Mol. bioSyst.* 13 (4) (2017) 725–735.
- [9] Y. Du, Z. Fu, V.D. Calhoun, Classification and prediction of brain disorders using functional connectivity: promising but challenging, *Front. Neurosci.* 12 (2018) 525.
- [10] R.M. Hutchison, T. Womelsdorf, E.A. Allen, P.A. Bandettini, V.D. Calhoun, M. Corbetta, S. Della Penna, J.H. Duyn, G.H. Glover, J. Gonzalez-Castillo, D. A. Handwerker, S. Keilholz, V. Kiviniemi, D.A. Leopold, F. de Pasquale, O. Sporns, M. Walter, C. Chang, Dynamic functional connectivity: promise, issues, and interpretations, *Neuroimage* 80 (2013) 360–378.
- [11] T.S. Zarghami, K.J. Friston, Dynamic effective connectivity, *NeuroImage* 116453 (2019).
- [12] C. Jin, H. Jia, P. Lanka, D. Rangaprakash, L. Li, T. Liu, X. Hu, G. Deshpande, Dynamic brain connectivity is a better predictor of ptsd than static connectivity, *Hum. Brain Mapp.* 38 (9) (2017) 4479–4496.
- [13] D. Rangaprakash, M.N. Dreitsch, A. Venkataraman, J.S. Katz, T.S. Denney, G. Deshpande, Identifying disease foci from static and dynamic effective connectivity networks: illustration in soldiers with trauma, *Hum. Brain Mapp.* (2017).
- [14] Z. Fu, A. Caprihan, J. Chen, Y. Du, J.C. Adair, J. Sui, G.A. Rosenberg, V.D. Calhoun, Altered static and dynamic functional network connectivity in alzheimer's disease and subcortical ischemic vascular disease: shared and specific brain connectivity abnormalities, *Hum. Brain Mapp.* 40 (11) (2019) 3203–3221.
- [15] I. Cribben, Y. Yu, Estimating whole brain dynamics using spectral clustering, *J. Roy. Stat. Soc.* (2016).
- [16] D. Xin, Z. Fu, S.C. Chan, Y.S. Hung, Z. Zhang, Task-related functional connectivity dynamics in a block-designed visual experiment, *Front. Hum. Neurosci.* 9 (2015).
- [17] M. Dai, Z. Zhang, A. Srivastava, Discovering common change-point patterns in functional connectivity across subjects, *Med. Image Anal.* 58 (2019) 101532, <https://doi.org/10.1016/j.media.2019.101532>.
- [18] Shakil, S., Billings, J. C., Keilholz, S. D., & Lee, C. H. (2017). Parametric dependencies of sliding window correlation. *IEEE Transactions on Biomedical Engineering*, PP(99), 1-1.
- [19] Y. Murayama M. Adhikari H., & Mantini, et al. Can sliding-window correlations reveal dynamic functional connectivity in resting-state fmri? (vol 127, pg 242, 2016) *NeuroImage* 132 2016 115 115.
- [20] S.B. Samdin, C.M. Ting, S.H. Salleh, M. Hamed, A.M. Noor, Identifying Dynamic Effective Connectivity States in fMRI Based on Time-Varying Vector Autoregressive Models. *International Conference for Innovation in Biomedical Engineering and Life Sciences*, 2015.
- [21] Y. Wang, S. Katwal, B. Rogers, J. Gore, G. Deshpande, Experimental validation of dynamic granger causality for inferring stimulus-evoked Sub-100 ms timing differences from fMRI, *IEEE Trans. Neural Syst. Rehabil. Eng.* 25 (6) (2017) 539–546.
- [22] X. Yuting, M.A. Lindquist, Dynamic connectivity detection: an algorithm for determining functional connectivity change points in fmri data, *Front. Neurosci.* 9 (2015) 285–.
- [23] B. Cai, G. Zhang, A. Zhang, J.M. Stephen, T.W. Wilson, V.D. Calhoun, Y.P. Wang, Capturing dynamic connectivity from resting state fmri using time-varying graphical lasso, *IEEE Trans. Biomed. Eng.* (2018).
- [24] A. Abrol, Z. Fu, Y. Du, V.D. Calhoun, Multimodal Data Fusion of Deep Learning and Dynamic Functional Connectivity Features to Predict Alzheimer's Disease Progression. 2019 41st Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). IEEE, 2019.
- [25] T.-E. Kam, H. Zhang, Z. Jiao, D. Shen, Deep learning of static and dynamic brain functional networks for early mci detection, *IEEE Trans. Med. Imaging* 39 (2) (2020) 478–487.
- [26] Vert, J. P., & Bleakley, K.. (2010). Fast detection of multiple change-points shared by many signals using group LARS. *Advances in Neural Information Processing Systems*.
- [27] L. Barnett, A.K. Seth, The mvgc multivariate granger causality toolbox: a new approach to granger-causal inference, *J. Neurosci. Methods* 223 (2014) 50–68.
- [28] K. Supratak M. Jin P. Jordan M.D. Jennifer G. Ying Estimating dynamic brain functional networks using multi-subject fmri data *NeuroImage* 2018 S1053811918306608-.
- [29] R. Tibshirani, M. Saunders, S. Rosset, J. Zhu, K. Knight, Sparsity and smoothness via the fused lasso. 67 (1) (2005) 91–108.
- [30] J.W. Liu, L.P. Cui, Z.Y. Liu, X.L. Luo, Survey on the regularized sparse models, *Chin. J. Comput.* (2015).
- [31] K. Bleakley, J.P. Vert, The group fused lasso for multiple change-point detection, *Quantitative Biology*, 2011.
- [32] Fundayurdakul, E. Cevher, Determinants of current account deficit in turkey: the conditional and partial granger causality approach, *Procedia Econ. Fin.* 26 (2015) 92–100.
- [33] L. Faes, G. Nollo, S. Stramaglia, D. Marinazzo, Multiscale granger causality, *Phys. Rev. E* 96 (4) (2017), <https://doi.org/10.1103/PhysRevE.96.042150>.
- [34] A.S. Heinsfeld, A.R. Franco, R.C. Craddock, A. Buchweitz, F. Meneguzzi, Identification of autism spectrum disorder using deep learning and the abide dataset, *Neuroimage Clinical* S2213158217302073 (2017).
- [35] M.A. Lindquist, Y. Xu, M.B. Nebel, B.S. Caffo, Evaluating dynamic bivariate correlations in resting-state fmri: a comparison study and a new approach, *NeuroImage* 101 (2014) 531–546.
- [36] I. Cribben, Y. Yu, Estimating whole brain dynamics using spectral clustering, *J. Roy. Stat. Soc.* (2015).
- [37] C.R. Jack, M.A. Bernstein, N.C. Fox, P. Thompson, G. Alexander, D. Harvey, B. Borowski, P.J. Britson, J. L. Whitwell, C. Ward, A.M. Dale, J.P. Felmlee, J. L. Gunter, D.L.G. Hill, R. Killiany, N. Schuff, S. Fox-Bosetti, C. Lin, C. Studholme, C. S. DeCarli, Gunnar Krueger, H.A. Ward, G.J. Metzger, K.T. Scott, R. Mallozzi, D. Blezek, J. Levy, J.P. Debbins, A.S. Fleisher, M. Albert, R. Green, G. Bartzokis, G. Glover, J. Mugler, M.W. Weiner, The alzheimer's disease neuroimaging initiative (adni): mri methods, *J. Magn. Reson. Imaging* 27 (4) (2008) 685–691.
- [38] C.G. Yan, Y.F. Zang, DPARSF: a MatLab toolbox for "pipeline" data analysis of resting-state fMRI, *Front. Syst. Neurosci.* 4 (2010) 13, <https://doi.org/10.3389/fnsys.2010.00013>.
- [39] G. Hao, Z. Fan, C. Junjie, X. Yong, X. Jie, Machine learning classification combining multiple features of a hyper-network of fmri data in alzheimer's disease, *Front. Neurosci.* 11 (2017) 615-.
- [40] Núria M.F., Marc M.F., Laia F.P., David B.F., Lidia V.A., Maribel P.C., & Joan G.O.. Resting-State Functional Connectivity Dynamics in Healthy Aging: An Approach Through Network Change Point Detection. (2020). *Brain Connectivity*, 134-142.
- [41] G. Deshpande, H. Jia, Multi-level clustering of dynamic directional brain network patterns and their behavioral relevance, *Front. Neurosci.* 13 (2020).
- [42] H.J. Park, K. Friston, C. Pae, B. Park, A. Razi, Dynamic effective connectivity in resting state fmri, *Neuroimage* S1053811917309606 (2017).
- [43] M. Ewers, R.A. Sperling, W.E. Klunk, M.W. Weiner, H. Hampel, Neuroimaging markers for the prediction and early diagnosis of alzheimer's disease dementia, *Trends Neurosci.* 34 (8) (2011) 430–442.
- [44] M. Demirta, C. Falcon, A. Tucholka, J.D. Gispert, José.L. Molinuevo, G. Deco, A whole-brain computational modeling approach to explain the alterations in resting-state functional connectivity during progression of alzheimer's disease, *Neuroimage Clinical* 16 (C) (2017) 343–354.
- [45] Y. Zhong, L. Huang, S. Cai, Y. Zhang, K.M. von Deneen, A. Ren, J. Ren, Altered effective connectivity patterns of the default mode network in Alzheimer's disease: an fmri study, *Neurosci. Lett.* 578 (2014) 171–175.
- [46] S. Shakil, C.-H. Lee, S.D. Keilholz, Evaluation of sliding window correlation performance for characterizing dynamic functional connectivity and brain states, *Neuroimage* 133 (2016) 111–128.
- [47] X. Wang, Y.S. Ren, W.S. Zhang, Multi-task fused Lasso method for constructing dynamic functional brain network of resting-state fMRI, *J. Image Graph.* 22 (7) (2017) 0978–0987.
- [48] Jonathan D. Power Alexander L. Cohen Steven M. Nelson Gagan S. Wig Kelly Anne Barnes Jessica A. Church Alecia C. Vogel Timothy O. Laumann Fran M. Miezin Bradley L. Schlaggar Steven E. Petersen Functional Network Organization of the Human Brain *Neuron* 72 4 2011 665 678.