Journal of Neuroscience Methods 271 (2016) 43–49

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

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

Unbiased cluster estimation of electrophysiological brain response

Matthew Frehlich[a,b], Luis Garcia Dominguez[a], Sravya Atluri[a,b], Natasha Radhu[a,b], Yinming Sun[a,b], Zafiris J. Daskalakis[a,b], Faranak Farzan[a,b,*]

[a] Temerty Centre for Therapeutic Brain Intervention, Centre for Addiction and Mental Health, Toronto, Ontario, Canada
[b] University of Toronto, Toronto, Ontario, Canada

HIGHLIGHTS

- Demonstrate dependency of cluster analysis results on user threshold statistic.
- Propose Unbiased Cluster Estimation (UCE), a threshold-free non-parametric approach.
- UCE is validated as a threshold-free approach for calculating statistical significance.

ARTICLE INFO

Article history:
Received 4 November 2014
Received in revised form 21 June 2016
Accepted 22 June 2016
Available online 23 June 2016

Keywords:
Cluster analysis
Correction for multiple comparison
Electroencephalography
Magnetoencephalography
Transcranial magnetic stimulation
Brain mapping
Statistical frameworks

ABSTRACT

Background: Recent increase in the size and complexity of electrophysiological data from multidimensional electroencephalography (EEG) and magnetoencephalography (MEG) studies has prompted the development of sophisticated statistical frameworks for data analysis. One of the main challenges for such frameworks is the multiple comparisons problem, where the large number of statistical tests performed within a high-dimensional dataset lead to an increased risk of Type I errors (false positives). A solution to this problem, cluster analysis, applies the biologically-motivated knowledge of correlation between adjacent voxels in one or more dimensions of the dataset to correct for the multiple comparisons problem and detect true neurophysiological effects. Cluster-based methods provide increased sensitivity towards detecting neurophysiological events compared to conservative methods such as Bonferroni correction, but are limited by their dependency on an initial cluster-forming statistical threshold (e.g. t-score, alpha) obstructing precise comparisons of results across studies.
New method: Rather than selecting a single threshold value, unbiased cluster estimation (UCE) computes a significance distribution across all possible threshold values to provide an unbiased overall significance value.
Comparison to existing methods: UCE functions as a novel extension to existing cluster analysis methods.
Results: Using data from EEG combined with brain stimulation study, we showed the impact of statistical threshold on outcome measures and introduction of bias. We showed the application of UCE for different study designs (e.g., within-group, between-group comparisons).
Conclusion: We propose that researchers consider employing UCE for multidimensional EEG/MEG datasets toward an unbiased comparison of results between subjects, groups, and studies.

1. Introduction

Brain activity captured at a global scale using electroencephalography (EEG) or magnetoencephalography (MEG) recordings can be described as a multidimensional entity, where signals are simultaneously recorded from a large number of sensors, typically at hundreds of samples per second. Under the assumption that specific brain processes are frequency-dependent, an extra dimension can also be added by modeling brain waves as a sum of oscillations over time. In recent research, electrophysiological brain response is often characterized through several dimensions (e.g., frequency, time, and phase) across space (i.e., all sensors). Typical analysis of such multidimensional data evaluates the significance of measured brain "activations", modelled as active regions in a random field (Marroquin et al., 2011). Elements (voxels) of these large datasets are considered "active" if they meet a certain threshold of statistical significance (i.e., t-score or z-score) compared to other voxels in the dataset. Since traditional analysis on a voxel-

* Corresponding author at: Centre for Addiction and Mental Health, Assistant Professor of Psychiatry, University of Toronto, 1001 Queen St. Unit 4 – 118A, Toronto, Ontario, Canada.
E-mail address: faranak.farzan@camh.ca (F. Farzan).

http://dx.doi.org/10.1016/j.jneumeth.2016.06.020
0165-0270/

wise basis involves a very large number of comparisons, it runs into the risk of Type I errors (false positives). This is known as the multiple comparisons problem and can be accounted for by using one of the many available statistical solutions.

Classically, voxel-wise correction methods such as the Bonferroni correction and others (Shaffer, 1995) are widely applied across EEG and MEG studies. However, these methods are excessively conservative and have reduced sensitivity towards detecting a true effect, leading to an abundance of Type II (false negative) errors. To address these issues, cluster-based thresholding frameworks have been introduced (Poline and Mazoyer, 1994; Bullmore et al., 1999; Maris and Oostenveld, 2007). Cluster-based frameworks take advantage of smoothness across one or more dimensions in EEG, MEG and related neuroimaging data (Groppe et al., 2011) by grouping active neighbouring voxels into clusters to represent neural activation events. For example, the spatial extent of an activation across neighbouring sensors in EEG results from volume conduction. Cluster-based frameworks typically involve two steps: 1) setting a threshold statistic (TS) and grouping neighbouring active (supra-threshold) voxels into clusters, and 2) calculating a p-value for each cluster, based on a measure of the size of activation. A well-known caveat of this method is in the choice of an initial cluster-forming TS since it can have a large effect on the resulting cluster and its value of significance (e.g., see Fig. 1 in Method). Moreover, instability from a thresholding process can introduce a signal-type bias, where a low TS favors a low intensity (in terms of t-score or z-score), spatially-extended signal and a high TS favors high intensity, focal events. Common cluster measures may include the number of active voxels in a cluster, the maximum intensity within a cluster, and more recently the cluster "mass". Cluster-extent and cluster-maximum measures introduce further bias towards extended weak signals and high intensity focal events, respectively. Cluster-mass methods (Bullmore et al., 1999; Maris and Oostenveld, 2007) however, have proven to be more sensitive since clusters are calculated over an initial TS and the sum of intensity values within each cluster is used to determine a "mass" value for each cluster. Therefore, the cluster-mass method not only accounts for cluster extent but also gives precedence to the intensity of the values contained within each cluster.

While cluster-mass methods attempt to relieve threshold-signal bias, further improvements are needed and have been introduced recently for multidimensional data analysis. One such recent framework is the threshold-free cluster enhancement (TFCE) method (Smith and Nichols, 2009) proposed for application in magnetic resonance imaging (MRI). TFCE is a cluster-maximum method that enhances the original data, boosting weaker signals lying in a larger cluster by a measure of their "support", i.e., the extent of the surrounding cluster. This involves enhancing each pixel p by a measure of its support over a range of thresholds h (from h_0 to h_p), and has the effect of enhancing weak but broadly supported signals to match sharp, focal signals:

$$\mathrm{TFCE(p)} = \sum_{h=h_0}^{h_p} e(h)^E h^H, \quad (1)$$

where h is the current threshold level, e is the extent of the signal at h, and E, H, are scaling parameters. While Smith and Nichols implemented this method for functional MRI (fMRI) data, it was recently extended to EEG analysis (Mensen and Khatami, 2013) and proved to increase the sensitivity of signal detection over traditional clustering methods. A limitation of this threshold-free approach however, is the introduction of additional parameters (E and H) that can be adjusted to vary results by either increasing bias towards the extent of activation or towards the intensity of activations. Another recent approach (Marroquin et al., 2011) suggests a morphology-based approach for detecting activations in random fields. However, this method similarly makes prior assumptions on the shape of the underlying activation event (requiring an input of structuring elements), introducing a user bias to the results. While these recent methods are more effective in controlling Type I error rate (also known as the family-wise error rate, FWER), a recent simulation study (Pernet et al., 2015) found that, on average, all cluster-based methods control FWER effectively. With this in mind, the current issue cluster analysis faces is not sensitivity and detection capabilities, but rather obtaining a bias-free measure of event significance.

To eliminate the recurrent issue of user bias in cluster analysis, we propose an unbiased cluster estimation (UCE) method and aim to extend existing cluster-based statistical frameworks. UCE is implemented in a true threshold-free manner without the need for tuning parameters, and provides a standardized measure across experiments for direct statistical comparisons. This is accomplished by using a number of different thresholds to create a distribution of p-values and then obtain an average result that represents a threshold-free index. As a result, accurate comparisons can be made between neurophysiological studies without introducing user bias. In this study, UCE is applied to the cluster-mass method (Maris and Oostenveld, 2007) to ensure the characteristics of EEG recordings are defined by a framework that is intuitive, statistically sensitive, and applicable to a wide range of scenarios. To demonstrate the use of the method, UCE is used to analyze multidimensional EEG recordings between two conditions within a group, and then between two independent groups. We will demonstrate that UCE properly detects neurophysiological effects, providing an unbiased measure of significance comparable to previous literature.

2. Method

2.1. Multidimensional dataset

Data is provided from previous transcranial magnetic stimulation (TMS) combined with EEG (TMS-EEG) experiments (Farzan et al., 2010; Radhu et al., 2014). This includes 60-channel EEG recordings during single-pulse and paired-pulse TMS paradigms, applied to dorsolateral prefrontal cortex in 84 subjects from two groups: healthy controls (n = 46), and schizophrenia patients (n = 38). We used this data to examine the degree of a neurophysiological process, cortical inhibition (CI), in two independent groups (within-subject design) and to compare CI between two groups (between-group design). Typically, a TMS-evoked response occurs after every TMS pulse. However, in the paired-pulse paradigm, CI causes the attenuation of the resulting TMS-evoked response (Daskalakis et al., 2008; Farzan et al., 2009). We have published on this research question and previously explored the underlying characteristics of the signals, so we focus on validating the UCE statistical methodology. Finally, this dataset permits application of our statistical approaches on both within and between subject designs.

Using this dataset, CI is evaluated as the difference between the EEG power of the two TMS conditions, single-pulse (one TMS pulse per trial: control paradigm) and paired-pulse (two TMS pulses 100 ms apart per trial: test paradigm) (García Domínguez et al., 2014). Power is defined for 106 time windows, from 0 to 500 ms after stimulation and for 50 frequencies, from 1 to 50 Hz. Thus, TMS conditions for each subject are represented by a matrix of 106 (time) × 50 (frequency) × 60 (channel) values.

2.2. Cluster analysis

Analysis on the spatial properties of high-density EEG data in the time and frequency domain requires a 4-dimensional framework, with an adjacency matrix to define neighbouring relationships. That is, two elements (voxels) from the dataset are neighbors if