ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 22 (2004) 268–276

Random Field–Union Intersection tests for EEG/MEG imaging

F. Carbonell,[a,*] L. Galán,[b] P. Valdés,[b] K. Worsley,[c] R.J. Biscay,[d] L. Díaz-Comas,[b] M.A. Bobes,[b] and M. Parra[b]

[a] *Institute for Cybernetics, Mathematics and Physics, Havana, Cuba*
[b] *Cuban Neuroscience Center, C.P. 10400 Havana, Cuba*
[c] *McGill University, Montreal, Canada*
[d] *Havana University, Cuba*

Received 14 August 2003; revised 15 January 2004; accepted 15 January 2004

Electrophysiological (EEG/MEG) imaging challenges statistics by providing two views of the same spatiotemporal data: topographic and tomographic. Until now, statistical tests for these two situations have developed separately. This work introduces statistical tests for assessing simultaneously the significance of spatiotemporal event-related potential/event-related field (ERP/ERF) components and that of their sources. The test for detecting a component at a given time instant is provided by a Hotelling's T^2 statistic. This statistic is constructed in such a manner to be invariant to any choice of reference and is based upon a generalized version of the average reference transform of the data. As a consequence, the proposed test is a generalization of the well-known Global Field Power statistic. Consideration of tests at all time instants leads to a multiple comparison problem addressed by the use of Random Field Theory (RFT). The Union-Intersection (UI) principle is the basis for testing hypotheses about the topographic and tomographic distributions of such ERP/ERF components. The performance of the method is illustrated with actual EEG recordings obtained from a visual experiment of pattern reversal stimuli.

Keywords: Event-related potentials; Random Fields; Union Intersection test; Hotelling's T^2; Global Field Power; Average reference; EEG/MEG Source Analysis

Introduction

Recent years have seen the emergence of some methods for electrophysiological (EEG/MEG) neuroimaging. They hold the promise of complementing other imaging modalities, such as PET or fMRI, by providing very high temporal resolution maps of neural activation. These techniques however challenge statistics in several ways. One problem is the need to detect significant components in the event-related potential/event-related field (ERP/ERF) data and their topographic distribution, that is, the *spatiotemporal* detection of the ERP/ERF components in high dimensional, very correlated data. In addition, interest is increasingly focusing on the identification of the current sources that generate such components. As a consequence, tomographic views of the same data should also be analyzed, for example, by the use of linear inverse solutions (Pascual-Marqui et al., 2002), which further increases the dimensionality of variables to be considered.

In practice, temporal components and their spatial location are usually identified by visual inspection of the peaks that appear in the ERP/ERF waveforms. However, this procedure may introduce bias in the analysis due to the large amount of channels recorded and typically excludes less pronounced phenomena. It is difficult to avoid subjectivity, especially when analyzing noisy recordings. A typical example of this type of situation in a clinical setting is the analysis of auditory ERPs to decide the presence of a response to low-intensity signals (Dobie, 1993). As a consequence some statistical decision procedures have been introduced to address the statistical identification of ERP/ERF components.

A common exploratory approach to determine ERP/ERF components in the topographic view is the use of the statistic Global Field Power (GFP). This measure, introduced by Lehmann (1987) is essentially the variance over electrodes of the potential/field of the average reference transformed data. Components are identified as maxima of the GFP as a function of time. It is difficult however to associate statistical significance to this measure. Measures based on statistical procedures are permutation *t* tests or MANCOVA techniques (Blair and Karniski, 1993; Friston et al., 1996; Galán et al., 1997; Raz, 1989). Subsequent testing of hypothesis for the topographic view may then be carried out, for example, by repeated measurement ANOVA, general MANOVA tests (Jenning et al., 1987; Vasey and Thayer, 1987), or permutation *t* tests. These techniques have been limited to a moderate number of sensors and time instants.

The related problem of objectively identifying the presence of the sources of the ERP/ERF component cannot be addressed by the same type of procedures due to the huge number of variables to be considered after the use of inverse solutions. For this reason, the use of novel methods based on Random Field Theory (RFT) and usually known as Statistical Parametric Mappings (SPM) (Friston et al., 1991; Worsley et al., 1992) is a promising alternative (Bosch et al., 2001; Dale et al., 2000; Park et al., 2002).

* Corresponding author. Departamento de Sistemas Adaptativos, Instituto de Cibernetica, Matematica y Fisica, Calle 15, No. 551, e/C y D, Vedado, La Habana 4, C.P. 10400, Cuba.
E-mail address: felix@icmf.inf.cu (F. Carbonell).
Available online on ScienceDirect (www.sciencedirect.com.)

1053-8119/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2004.01.020

The main problem with these approaches is that they use different criteria for topography and for tomography, in fact, employing statistical tests derived from different principles. As a consequence, there is no general method for assessing simultaneously the spatiotemporal detection of ERP/ERF components and the localization of the current density sources that generate them.

This work presents a blend of techniques that provides a global decision threshold for the detection of ERP/ERF components. It is shown that the same threshold can be used for assessing the statistical significance of both the topographic and tomographic distribution of such components. Specifically, the statistical procedure proposed in this paper can be summarized in the following steps:

(1) Definition of the null hypothesis of "no temporal ERP/ERF components".
(2) Calculation of a multivariate T^2 statistic $\phi(t)$ for each t in the time window of analysis to test the hypothesis of Step 1.
(3) Computation, by RFT, of a global decision threshold G_α (with the assumption that the statistics of Step 2 are realizations of a one-dimensional random field). Detection of the ERP/ERF components at those time instants t_0 for which $\phi(t_0)$ exceeds G_α.
(4) Application of the Union-Intersection (UI) tests for examining hypotheses about the topographic and tomographic distribution (at the time instants selected by Step 3) of the ERP/ERF components.

This main idea underlying this approach is that while EEG/MEG data is usually represented by vectors of voltage/field values recorded from a set of sensors, many topographic and tomographic views of the same data are obtained by linear combinations (LCs) of these vectors. For example, different *montages* of the EEG such as bipolar or Laplacian are just LCs of the original data. Additionally, linear inverse solutions are also LCs of the original data. The Union Intersection principle provides tests for any chosen LC.

Thus, the power of this proposal is that one could construct statistical tests for the topographic distribution of the original sensor recordings. On the other hand, it is also possible to provide statistical tests for all LC of generators produced by any conceivable linear EEG/MEG inverse solution. The T^2 statistic developed is constructed to be independent of the reference. This is achieved by developing a novel maximum likelihood estimate of the reference, a generalization of the average reference (Lehmann, 1987) that in turn leads to a generalization of the average reference transformation of the data. The resultant T^2 statistic is an inverse covariance weighted generalization of the GFP statistic, thus providing a testable version of this measure.

The plan of the paper is the following: The basic foundations of the statistical methodology are developed in the first section. Its application to actual EEG data is carried out in the second section. In last section, some assumptions that make possible the use of RFT are briefly discussed.

Method

Let $X = [X]_{sct}$ denote the three-dimensional array that represents the recorded voltage values corresponding to trial s, channel (sensor on the scalp) c, and time instant t ($s = 1 \ldots S$, $c = 1 \ldots C$, $t = 1 \ldots T$). Here, S, C, and T denote, respectively, the numbers of trials, channels, and time instants (within a given time window).

The common model for such ERP data (Aunon et al., 1981) assumes that the recorded voltage values (X_{sct}) are the superposition of the ERP ("signal" μ_{ct}) and a background brain activity "noise" (e_{sct}) noncontingent to the stimuli, that is,

$$X_{sct} = \mu_{ct} + e_{sct},$$

where the vectors $e_{st} = (e_{s1t},\ldots,e_{sCt})$ are assumed statistically independent across trials and normally distributed with mean equal to zero and (typically unknown) covariance matrix Σ_t.

Detection of temporal components

In order to detect ERP/ERF components, one would like to test, in principle, the null hypothesis of a null ERP/ERF average ($\mu = 0$). However, when looking at topographies, especially ERPs, one must consider the possibility of activity in the reference that may shift the entire observation vector by the same constant at each sensor. The null hypothesis may be modified to cover this case by positing a constant mean. This can be stated formally as the general hypothesis:

$$H_0\colon\ \mu_{ct} = r_t \text{ for all } t = 1 \ldots T, c = 1 \ldots C$$

where r_t is an arbitrary constant and μ_{ct} ($c = 1\ldots C$, $t = 1\ldots T$) denote the components of the vector μ. This global null hypothesis can be decomposed in the following T marginal null hypotheses:

$$H_{0t}\colon \mu_{1t} = \mu_{2t} = \ldots = \mu_{Ct} = r_t \text{ for all } t = 1 \ldots T.$$

The likelihood ratio test for H_{0t} is a Hotelling's T^2 statistic (Mardia et al., 1979), say T_t^2, which follows a probability distribution $T^2(C-1, S-1)$ under H_{0t} (see details in the Appendix).

It is to be noted that the statistic proposed in this paper is NOT the usual Hotelling's T^2 statistic used for testing a mean of zero. Rather, it is a modified statistic that does not depend on a given reference. In fact, the proposed statistic is equivalent to calculating a maximum likelihood estimate of the reference, subtracting it from the data, and then performing the usual Hotelling's T^2 test for zero mean.

Statistical tests and RFT

For a fixed time instant t, a test for H_{0t} at the significance level α is given by rejecting this hypothesis when T_t^2 is greater than the $(1-\alpha)100$ percentile of the distribution $T^2(C-1, S-1)$, say c_α. However, a test for H_0 at the level α is not obtained by rejecting H_{0t} if $T_t^2 > c_\alpha$ for some time instant t. This is due to the well-known fact that the application of several tests, each one at the significance level α, yields to the inflation of the type I error, that is, to a test with actual level greater than α. This is the well-known multiple comparisons problem (Hochberg and Tamhane, 1987). One of the classical strategies to solve this problem is the use of the Bonferroni correction. However, this is usually an excessively conservative procedure and leads to an unnecessary loss of statistical power for the case of correlated tests, which is likely to occur in our setting due to the temporal correlation structure that usually appears in multichannel EEG/MEG data (Jimenez et al., 1995 and references therein). A more effective solution to the correlated multiple comparisons problem is the use of RFT (Adler, 1981) for generating distributional approximations. The application of this approach to the current setting

reduces to consider that the statistics T_t^2 are realizations from a (1D) Hotelling's T^2 random field (Cao and Worsley, 1999). Thus, a test for H_0 is based on the statistic $T^2_{\max} = \max(T_1^2 \ldots T_T^2)$, where H_0 is rejected when $T^2_{\max} > G_\alpha$. Here, G_α denotes the $(1 - \alpha)100$ percentile of the distribution of $T^2_{\max}$ under H_0, which can be calculated according to the expressions given in the Appendix. Finally, the ERP/ERF components are detected when the null hypothesis H_0 is rejected, that is, at those time intervals $[t_1,t_2]$ for which $T_t^2 > G_\alpha$ holds for all t in $[t_1,t_2]$.

In the next subsection, a detailed analysis of the scalp topography is carried out to provide important information about the spatial identification of the detected components.

Topographic distribution of the ERP/ERF components

According to the previous subsection, those time instants t for which $T_t^2 > G_\alpha$ should be analyzed. However, to simplify analyses, only a single time instant t_0 is selected, namely, that for which T_t^2 reaches its maximal value (i.e., $T_{t_0}^2 = \max(T_t^2)$).

One of the practical consequences of the UI tests is that if H_0 is rejected, then one can ask which components of the vector μ_{t_0} were responsible. For this purpose, the following hypotheses, which are LC of the vector $\mu_{t_0}=(\mu_{1t_0},\ldots,\mu_{Ct_0})'$, are considered:

$$H_{0t_0}^{\boldsymbol{a}_c}: \boldsymbol{a}_c\mu_{t_0} = r_{t_0},$$

where the vector $\boldsymbol{a}_c = (0,\ldots,1,\ldots,0)$ has a number 1 in the cth component. Note that $H_{0t_0}^{\boldsymbol{a}_c}$ focuses on determining whether the component μ_{ct_0} significantly contributes to the rejection of H_{0t_0}. Each of theses hypotheses yields to the statistic $t_{\boldsymbol{a}_c}$, which distributes t Student with $S - 1$ degrees of freedom (see the expression for $t_{\boldsymbol{a}_c}$ in the Appendix). According to the UI tests, $H_{0t_0}^{\boldsymbol{a}_c}$ is rejected when $(t_{\boldsymbol{a}_c})^2 \geq G_\alpha$. Hence, the identified topography of the ERP/ERF components is determined by those sites c that satisfies this inequality. A practical way to visualize the topographic distribution of such channels is to display the values of the vector $(t^2_{a_1},\ldots,t^2_{\boldsymbol{a}_c})$ as a code-color interpolated topographic map (Galán et al., 1997).

Tomographic distribution of the ERP/ERF components

As is well known, one should be cautious in interpreting the ERP/ERF topography in terms of neural activation regions, which requires the use of specific tomographic techniques. Indeed, the development of physical models and mathematical methods has led to a novel imaging technique, usually called brain electromagnetic tomography (BET), which has become a very useful tool to determine the possible neural generators of ERP/ERF components. It is based on estimates of current density sources inside the brain, by solving the EEG/MEG inverse problem. That is, the problem $\mu_{t_0} = KJ_{t_0}$, where J_{t_0} represents primary current density inside the brain, μ_{t_0} denotes the mean vector of voltage values recorded from sensors at the time instant t_0, and K is known as the lead field matrix.

Although the methods developed in this paper are valid for any linear inverse solution, their use shall be illustrated with the LORETA solution (Pascual-Marqui et al., 1994):

$$\hat{J}_{t_0} = (K'K + \lambda^2 L'L)^{-1}K'\mu_{t_0},$$

where $\hat{J}$ is the estimated current density, λ is a regularizing parameter, and L is a discrete version of the Laplacian operator.

It should be noted at this point that since the EEG/MEG inverse problem is an ill-posed problem, there could exist primary current sources unobservable in $\hat{J}$. In fact, the solution $\hat{J}$ only observes primary current sources in the subspace generated by the rows of the lead field matrix K. As a consequence, one should be careful in the interpretation of the electrophysiological images obtained as solutions of the EEG/MEG inverse problem. In spite of this, various empirical studies have validated the application of the LORETA solution under diverse physiological conditions (Pascual-Marqui et al., 2002 and references therein).

Suppose that the vector $\hat{J} = (\hat{J}_1',\ldots,\hat{J}_{n_g}')'$ has dimension $3n_g$, where n_g denotes the number of generators sampled over a 3-D domain in the brain and $\hat{J}_g$, $g = 1\ldots n_g$, represents the three spatial components of the current density for the gth generator. From the expression given for $\hat{J}$, it is deduced that each vector $\hat{J}_g$ is a certain linear transformation $A_g\mu$ of the vector μ. Therefore, hypotheses about $A_g\mu$ are equivalent to the corresponding hypotheses about $\hat{J}_g$, where the most interesting ones are those of the type

$$H_{0t_0}^{A_g}: J_g = 0.$$

That is, the hypotheses of no primary current density source at the generator g. Each of these hypotheses yields to a Hotelling's T^2 statistic, say $T^2_{A_g}$ (see the expression in the Appendix). Thus, again by the UI tests, the sites of activation correspond to those values of g for which $T^2_{A_g} \geq G_\alpha$.

Application to actual EEG data

Experimental design

This section describes a visual experiment that was designed to show the performance of the statistical methodology proposed in the previous section. The experiment consisted on the application of pattern reversal stimuli for three different experimental conditions:

(A) Pattern reversal with checks of size 27 mm at 70 cm from the computer display.
(B) Pattern reversal with checks of size 7mm at 70 cm from the computer display.
(C) Black display with a fixation point.

It is well known that due to the size of the checks, each of the above experimental conditions should generate ERPs of different amplitudes. Specifically, the amplitude of the ERP generated by the experimental condition A should be, in general, higher than the amplitude of the ERP generated by the condition B and no ERP components are expected from the condition C. For that reason, in the sequel, the conditions A–B shall be called by "High Amplitude", "Low Amplitude" and "No ERP" conditions, respectively.

A sample of 15 EEG recordings was obtained from five healthy subjects. EEG signals were obtained for the three experimental conditions for each subject. The statistical procedure presented in the previous section was applied to each of these 15 recordings. Thus, the main purpose is to compare the statistical results with the actual response obtained by the experimental conditions.

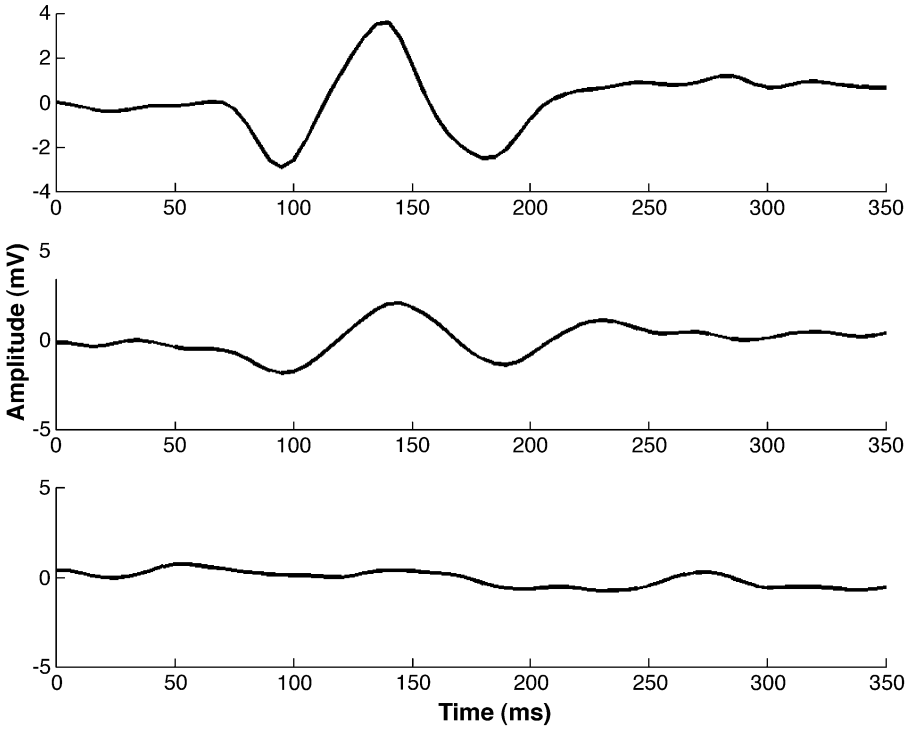

Fig. 1. Waveform of the trials average for one electrode in the occipital region.

Recording technique

The electroencephalographic (EEG) activity was recorded with an electrode cap of 120 channels. The impedance was kept below 10 kΩ. Linked earlobes were used as reference and the forehead was grounded. Two bipolar derivations were used to monitor the horizontal and vertical electrooculogram (EOG). After amplification and filtering from 0.05 to 30 Hz, the EEG was digitized with a 12-bit converter. Digitization was synchronized with the onset of the stimulus, with a sampling period of 5 ms. Epochs of 400 ms

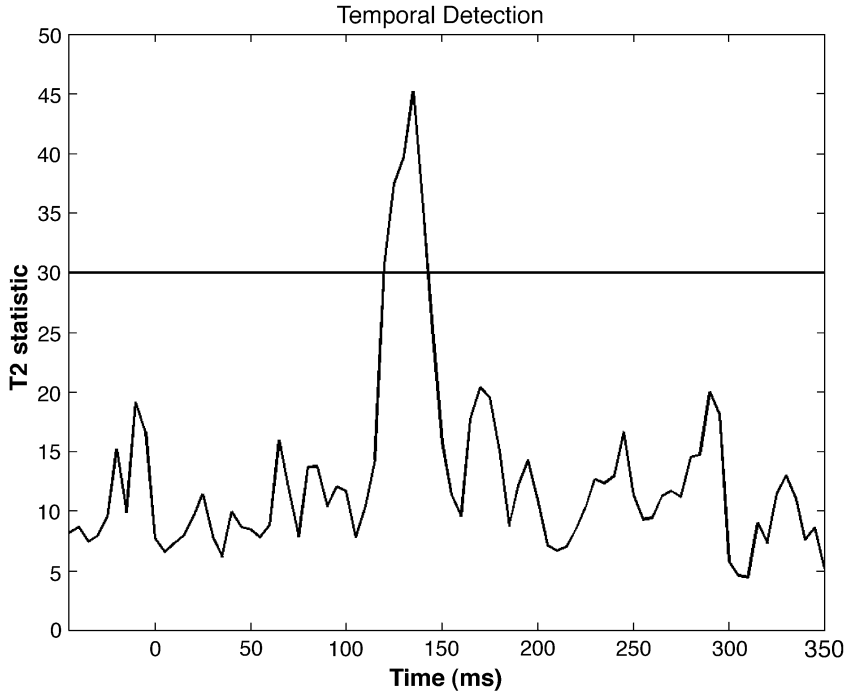

Fig. 2. T^2 statistics for the "High Amplitude" condition. The horizontal line represents the decision threshold at 5% of significance level.

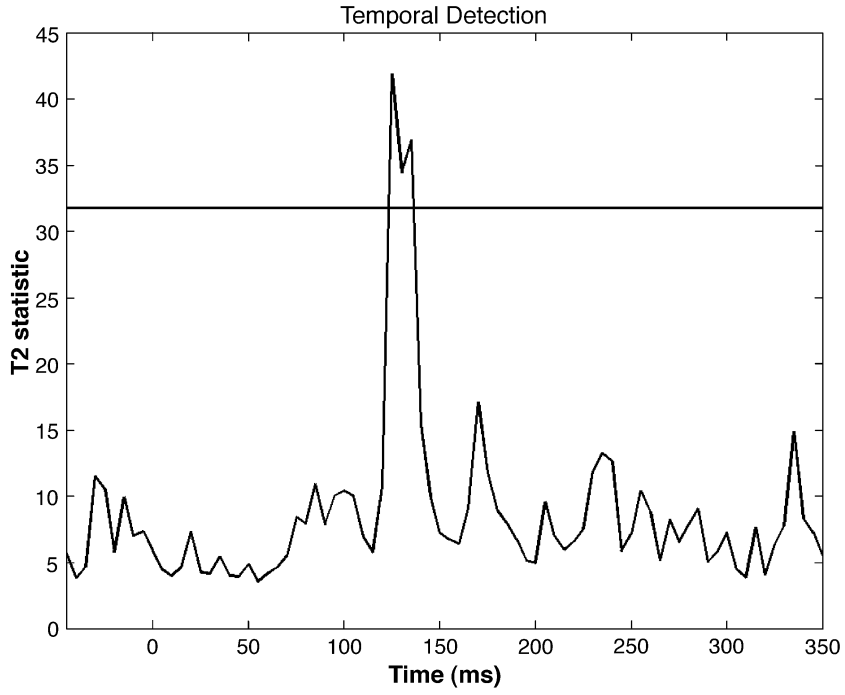

Fig. 3. T^2 statistics for the "Low Amplitude" condition. The horizontal line represents the decision threshold at 5% of significance level.

were selected for each trial, using a 50-ms prestimulus window. Artifact rejection was carried out for the detection of those trials with artifacts or eye movements, which was supervised by trained neurophysiologists. Such trials were discarded for the further analysis.

Results

In all that follows, the probability significance level was fixed to $\alpha = 0.05$ and the time window of analysis to [0 ms, 350 ms], where $t = 0$ corresponds to the presentation of the stimuli.

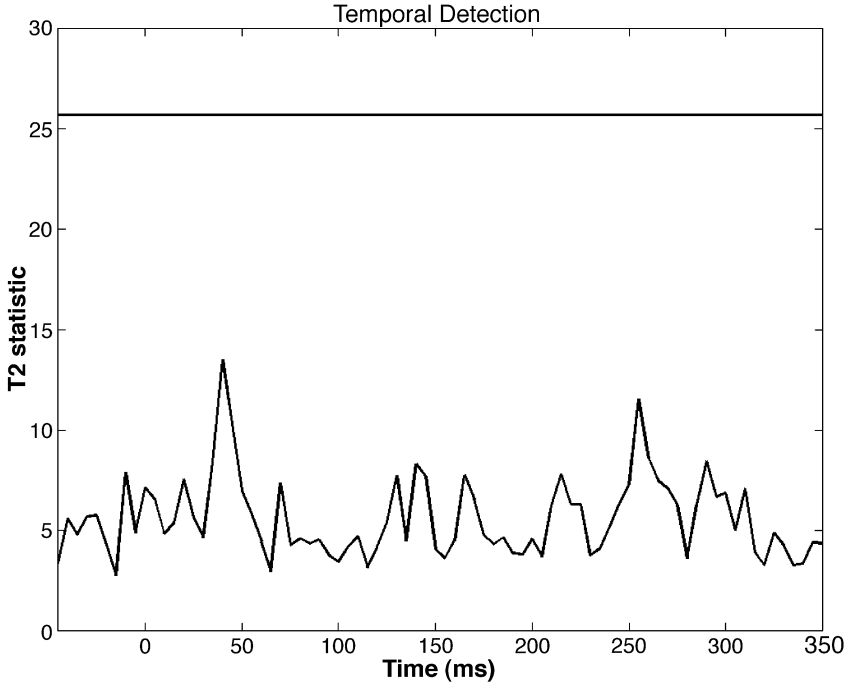

Fig. 4. T^2 statistics for the "No ERP" condition. The horizontal line represents the decision threshold at 5% of significance level.

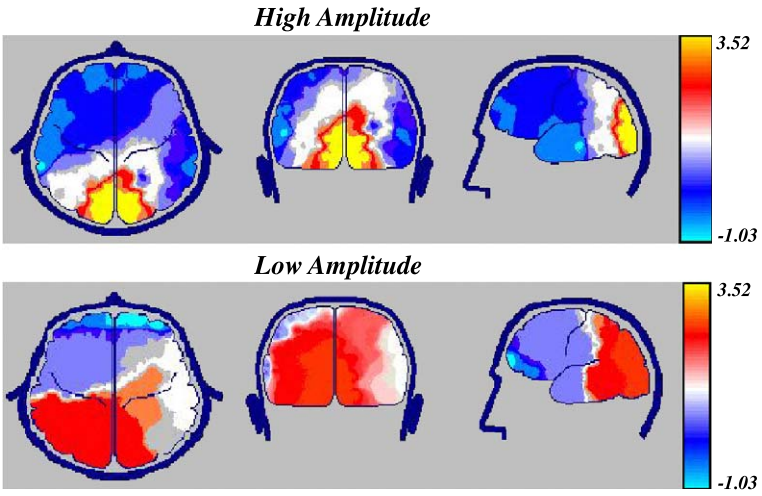

Fig. 5. Trials average topographic distribution of all channels at the time instant of T^2 statistics global maxima.

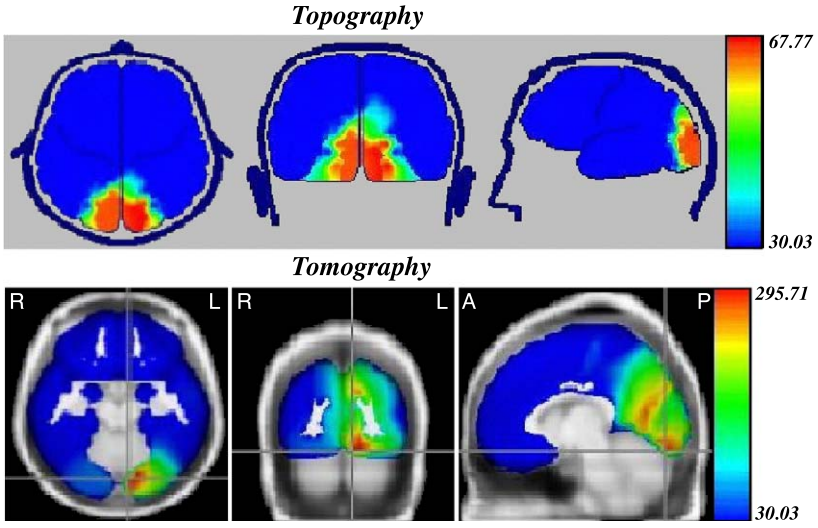

Fig. 6. Topographic and tomographic maps for the "High Amplitude" condition at the time instant of T^2 statistics global maxima.

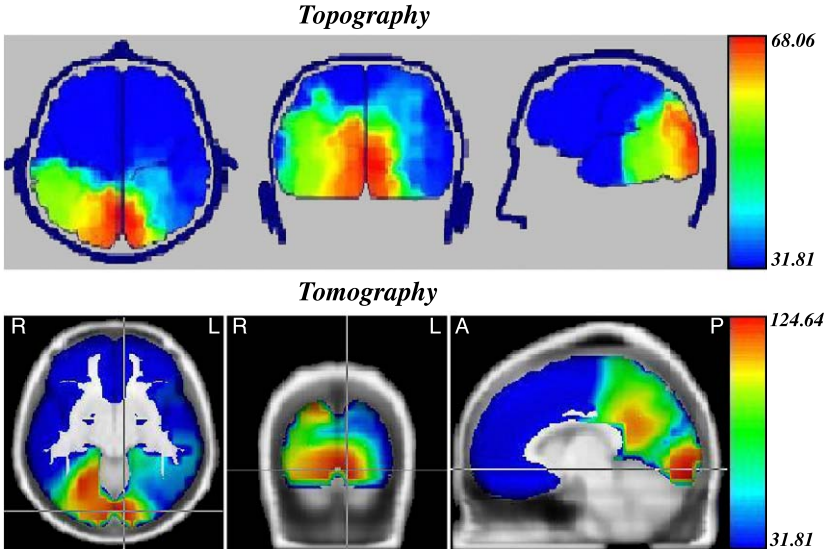

Fig. 7. Topographic and Tomographic maps for the "Low Amplitude" condition at the time instant of T^2 statistics global maxima.

In order to detect the presence of ERP components, the T^2 process (T_t^2) was computed for all subjects and the three experimental conditions. For all these cases, a significant ERP component was detected around 110–150 ms after the presentation of the stimuli. On the other hand, for the second experimental condition, only three of the five cases showed the presence of significant ERP components. In addition, the time window in which these components were significant was narrower than the one corresponding to the first condition. For the three cases, the components appeared around 125–135 ms. Finally, the method did not detect significant ERP components at any of the five cases corresponding to the third experimental condition.

Once the time window in which ERP was detected was determined, the next step was to identify the topographic and tomographic distribution of the significant ERP components (those t from which $T_t^2 > G_\alpha$). Instead of analyzing the whole time window, only that time instant t_0 for which $T_{t_0}^2 = \max_t(T_t^2)$ holds was selected for further analysis.

The statistics $(t_{\mathbf{a}_c})^2$, $c = 1,\ldots,C$ were computed at t_0 for the eight cases (five for first condition and three for the second one) with significant ERP components, and their values were further evaluated by topographic maps. In all the cases, the maps revealed that the set $\{(t_{\mathbf{a}_c})^2, c = 1,\ldots,C\}$ had a topographic distribution with the significant values predominantly in the occipital region. The topographic distribution of these significant values was less concentrated for the three cases corresponding to the "Low Amplitude" condition.

In order to determine the tomographic distribution of the current density sources that might generate the ERP components, the (LORETA) inverse solution was computed for the eight cases abovementioned. For all cases, the inverse solution showed significant activations, located predominantly in the occipital region. However, as in the topographic maps, the significant sources appear more to most spread out for those cases corresponding to the "Low Amplitude" condition.

A single subject was selected to illustrate the results previously described. Fig. 1 shows for this subject (under the three experimental conditions) the average EEG signal (across trials) for one electrode in the occipital region. The top panel corresponds to the "High Amplitude" condition and suggests the existence of an evoked potential around 100–150 ms after the presentation of the stimuli. However, such evoked potential was less evident for the "Low Amplitude" condition and seems to be absent for the "No ERP" condition.

Fig. 2 illustrates the Hotelling's T^2 computed for the "High Amplitude" condition. It can be seen that the time instants for which the T^2 curve exceeds the decision threshold are concentrated around 100–150 ms. This is in correspondence with the result shown in Fig. 1, where a P100 component is evident. Hotelling's T^2 process and the corresponding decision threshold for the "Low Amplitude" hold are shown in Fig. 3. It can be seen that in this case, the amplitude of the component P100 is lower than the amplitude of that component in the "High Amplitude" condition, which is an expected result due to the decrease of the size of the checks in the stimuli. Unlike the results shown in the two previous figures, for the case of "No ERP" condition, the statistical test detected no significant ERP components, which is illustrated in Fig. 4.

Fig. 5 illustrates for both the "High Amplitude" and the "Low Amplitude" conditions the topographic distribution of all channels of the mean ERP at the time instant t_0 (for which T_t^2 has maximal value). Note in both maps the maximal values in the occipital area. Fig. 6 shows, for the "High Amplitude" condition, the topographic and tomographic distribution (at t_0) of the component P100 detected in Fig. 2. The top panel shows the topographic map for the square values of the t statistics corresponding to all channels. It is seen that the highest values of the t statistics are reached for those channels located predominantly in the occipital region. In correspondence with these results, the tomographic view in the bottom panel of the figure shows a significant activation of both hemispheres in the occipital region. Finally, Fig. 7 shows the topographic and tomographic maps corresponding to the case of "Low Amplitude" condition. As in Fig. 6, these maps correspond to the time instant t_0 where T_t^2 is maximal. It can be seen that for both maps, the activated areas are located principally in the occipital region. However, in this case, the activated areas are more widely distributed than in the case of Fig. 6.

Discussion

This paper is based on the use of Hotelling's T^2 to detect ERP/ERF components. If a data offset due to the use of a reference is considered, then the maximum likelihood derivation of the modified test leads to novel estimates of the reference, to a generalized average reference transformation, and to a generalization of the GFP measurement that allows statistical testing. Thus, several intuitively developed measures can be given a firm statistical foundation. Additionally, the fact that the Hotelling's T^2 statistic may be derived from the UI principle allows testing simultaneously the null hypothesis for ANY linear combination of the data. This includes testing any of the individual leads as well as any other that can be obtained from the original leads by a linear combination. Examples are bipolar leads or the Laplacian reference. In particular, linear inverse solutions are particular linear combinations and therefore the methods developed in this paper apply to any of the particular linear tomographic methods as well.

A possible limitation of the methods presented here is the basic model for the ERP/ERF as the linear superposition of the event-related activity and background activity. This model is only a first approximation, useful for exploratory purposes. More realistic, nonlinear and nonstationary models of the ERP/ERF may require more complex procedures that are more computer intensive. It should be mentioned that the procedure proposed in this paper was coded in Matlab 6.5 and took an average of 14 ± 3 s to execute in a Pentium IV personal computer showing that they are quite practical to use.

The extension of the test statistic to a whole time interval hinges on the use of RFT. Some assumptions underlying the application of RFT should be therefore made explicit (see Peterson et al., 1999 for an excellent review on this topic). A general assumption is that the sampled statistical images under analysis should be good approximations of continuous random fields, which is assured for a high spatial sampling rate and a sufficient amount of image smoothing. Other assumptions such as gaussianity and stationarity must also be satisfied, though in general, they are very difficult to ensure in practice. However, RFT has been recently extended for the case of random fields with non-stationarities (Worsley et al., 1999). On the other hand, increasing the number of trials in the electrical recordings and selecting an adequate amount of data smoothing are recommendable solutions in those cases of strong evidences against the gaussian assumption. From the practical point of view, the RFT is only useful

when the data is highly correlated. This is the typical case for EEG/MEG signals, which show a spectral pattern different from a white noise (i.e., flat spectrum) in most experimental settings (Dumermuth and Molinari, 1987 and references therein). However, in those situations with evidence of low temporal correlation, a compromise between the Bonferroni correction and the RFT approach would be a recommendable alternative. A viable choice would be, for instance, the minimum value between the Bonferroni correction and the threshold provided by RFT.

Conclusions

In this paper, a novel application of the RFT and the UI tests was introduced in the framework of EEG/MEG hypothesis-testing problems. RFT was applied to solve the multiple comparisons problem that appears when examining the global null hypothesis of "No ERP/ERF components". On the other hand, the UI tests were used for testing hypothesis about the topographic and tomographic distribution of the detected ERP/ERF components. Therefore, the combination of both techniques provided a general criterion for assessing simultaneously the spatiotemporal detection of ERP/ERF components and the localization of the current density sources that generate them. Finally, physiologically meaningful results were obtained in the application of the statistical methodology to actual EEG data recorded from a visual experiment of pattern reversal stimuli.

Appendix A

The T_t^2 statistics

The hypothesis H_{0t} can be rewritten as H_{0t}: $\mu_t = r_t\mu_0$, where $\mu_0 = (1,\ldots,1)$. The likelihood ratio statistic for this hypothesis was given by (Mardia et al., 1979, pp. 153–154):

$$T_t^2 = d_t' V_t^{-1} d_t,$$

where $d_t = \bar{x}_t - \hat{r}_t\mu_0$, $\overline{x_t} = \frac{1}{S}\sum_{s=1}^{S} x_{st}$, $x_{st} = (X_{s1t},\ldots,X_{sCt})'$, and the prime denotes transposition of vectors.

V_t and $\hat{r}_t$ denote, respectively, the maximum likelihood estimate of the covariance Σ_t and r_t:

$$V_t = \frac{1}{S}\sum_{s=1}^{S}(x_{st}-\overline{x_t})(x_{st}-\overline{x_t})', \hat{r}_t = \mu_0' V_t^{-1}\overline{x_t}/\mu_0' V_t^{-1}\mu_0.$$

It should be noted that for the case $V_t = \sigma^2 I$, it follows that $\hat{r}_t$ is the average reference, d_t is the average reference transform of the data, and the test statistic T_t^2 is a scaled version of the Global Field Power (Lehmann, 1987). Thus, for general V_t, $\hat{r}_t$ can be viewed as a maximum likelihood estimate of the reference, a generalization of the average reference and d_t can be viewed as the generalized version of the average reference transform of the data. Note also that in the expressions above, it has been assumed $S \geq C$ to have V_t^{-1} well defined. However, a similar approach to that in Uhlig (1994) could be followed to obtain T^2 statistics when V_t is a singular matrix, as is the case for $S < C$.

Computation of G_α

For any value z, the explicit approximation for $P({T^2}_{\max} \geq z)$ is given by:

$$P(T^2_{\max} \geq z) \approx \rho_0(z) + Resels_1\rho_1(z)$$

where the EC densities are

$$\rho_0(z) = \int_z^\infty \frac{\Gamma\left(\frac{S}{2}\right)}{\Gamma\left(\frac{C-1}{2}\right)\Gamma\left(\frac{S-C+1}{2}\right)}(1+u)^{-\frac{S}{2}}u^{\frac{C-3}{2}}du,$$

$$\rho_1(z) = (4\log 2)^{\frac{1}{2}}\frac{\pi^{-\frac{1}{2}}\Gamma\left(\frac{S}{2}\right)}{\Gamma\left(\frac{C-1}{2}\right)\Gamma\left(\frac{S-C+2}{2}\right)}(1+z)^{-\frac{S-2}{2}}z^{\frac{C-2}{2}},$$

and Γ is the gamma function. Note that $\rho_0(z)$ is the P-value of ${T_t}^2$ for fixed t, equal to the probability that an F random variable with $C-1$, $S-C+1$ degrees of freedom exceeds $z(S-C+1)/(C-1)$. The number of 1D resels, $Resels_1$, can be estimated unbiasedly as follows (Worsley et al., 1999). Let τ_{ct} be the sample correlation between $X_{s,c,t}$ and $X_{s,c,t-1}$ over s for fixed c, t. Then

$$Resels_1 = (4\log 2)^{-\frac{1}{2}}\frac{1}{C}\sum_{c=1}^{C}\sum_{t=2}^{T}\sqrt{2(1-\tau_{ct})}.$$

Finally the threshold G_α is obtained by solving the equation $P({T^2}_{\max} \geq G_\alpha) = \alpha$.

Union Intersection tests

Suppose that a multivariate hypothesis H_0 can be written as the intersection of the set of univariate hypotheses H_{0a}, that is, $H_0 = \cap H_{0a}$, then the UI test for H_0 is a test whose rejection region can be written as $R = \cup R_a$, where R_a is the rejection region corresponding to the component hypothesis H_{0a}.

If R_a were based on the statistic test z_a and expressed in the form $R_a = \{z_a: z_a^2 > c^2\}$ for some arbitrary critical value c, then the UI test is given by $\max_a z_a^2$ and the hypothesis H_0 is rejected if and only if $\max_a z_a^2 > c^2$.

The statistics t_{a_c} and $T^2_{A_g}$

The t statistic t_{a_c} at the time instant t_0 is given by

$$t_{a_c} = a_c'(\overline{x}_{t_0} - r_{t_0}^{\wedge})/[a_c'\Sigma_{t_0}^{\wedge}a_c/(S-1)]^{1/2}, a_c = (0,\ldots,1,\ldots,0)$$

Let J_g^s be the inverse solution at the generator g obtained from

$$J^s = (J_1^s,\ldots,J_{n_g}^s) = (K'K + \lambda^2 L'L)^{-1}K'x_{st_0}$$

Then, the Hotelling's T^2 statistic $T^2_{A_g}$ is computed by

$$T^2_{A_g} = \overline{J}_g'(\Sigma_{J_g})^{-1}\overline{J}_g, \text{where}$$

$$\overline{J}_g = \frac{1}{S}\sum_{s=1}^{S}J_g^s \text{ and } \Sigma_{J_g} = \frac{1}{S}\sum_{s=1}^{S}(J_g^s - \overline{J}_g)(J_g^s - \overline{J}_g)'.$$

References

Adler, R.J., 1981. The Geometry of Random Fields. Wiley, New York.

Aunon, J.I., McGillen, C.D., Childers, D.C., 1981. Signal processing in evoked potentials research: averaging and modeling. CRC Critical Reviews in Bioengineering 5, 323–367.

Blair, R.C., Karniski, W., 1993. An alternative method for significance testing of waveform difference potential. Psychophysiology 30, 518–524.

Bosch, J., Valdés, P., Virues, T., Aubert, E., John, R., Harmony, T., Riera, J., Trujillo, N., 2001. 3D statistical parametric mapping of EEG source spectra by means of variable resolution electromagnetic tomography (VARETA). Clinical Electroencephalography 32, 47–61.

Cao, J., Worsley, K.J., 1999. The detection of local shape changes via the geometry of Hotelling's T2 fields. Annals of Statistics 27, 925–942.

Dale, A.M., Liu, A.K., Fischl, B.R., Buckner, R.L., Belliveau, J.W., Lewine, J.D., Halgren, E., 2000. Dynamical statistical parametric mapping combining fMRI and MEG for high-resolution imaging of cortical activity. Neuron 26, 55–67.

Dobie, R.A., 1993. Objective response detection. Ear and Hearing 14, 31–35.

Dumermuth, G., Molinari, L., 1987. Spectral analysis of the EEG. Some fundamentals revisited and some open problems. Neuropsychobiology 17, 85–99.

Friston, K.J., Frith, C.D., Liddle, P.F., Frackowiak, R.S., 1991. Comparing functional (PET) images: the assessment of significant change. Journal of Cerebral Blood Flow and Metabolism 11, 690–699.

Friston, K.J., Stephan, K.M., Heather, J.D., Frith, C.D., Ionnides, A.A., Liu, L.C., Rugg, M.D., Vieth, J., Keber, H., Hunter, K., Frackowiak, R.S.J., 1996. A multivariate analysis of evoked responses in EEG and MEG data. NeuroImage 3, 167–174.

Galán, L., Biscay, R., Rodriguez, J.L., Perez-Avalo, M.C., Rodriguez, R., 1997. Testing topographic differences between event related brain potentials by using non-parametric combinations of permutation tests. Electroencephalography and Clinical Neurophysiology 102, 240–247.

Hochberg, Y., Tamhane, A.C., 1987. Multiple Comparisons Procedures. Wiley, New York.

Jenning, J.R., Cohen, M.J., Ruchkin, D.S., Fridlund, A.J., 1987. Editorial policy on analysis of variance with repeated measures. Psychophysiology 24 (4), 474–478.

Jimenez, J.C., Biscay, R., Montoto, O., 1995. Modeling the electroencephalogram by means of spatial spline smoothing and temporal autoregression. Biological Cybernetics 72, 249–259.

Lehmann, D., 1987. Principles of spatial analysis. In: Gevins, A., Gémond, A. (Eds.), Handbook of Electroencephalography and Clinical Neurophysiology, Rev. Ser. Analysis of Electrical and Magnetic Signals, vol. 1. Elsevier, Amsterdam, pp. 309–354.

Mardia, K.V., Kent, J.T., Bibby, J.M., 1979. Multivariate Analysis Academic Press, London.

Park, H.J., Kwon, J.S., Youn, T., Pae, J.S., Kim, J.J., Kim, M.S., Ha, K.S., 2002. Statistical parametric mapping of LORETA using high density EEG and individual MRI: application to mismatch negativities in schizophrenia. Human Brain Mapping 17, 178–186.

Pascual-Marqui, R.D., Michel, C.M., Lehmann, D., 1994. Low-resolution electromagnetic tomography: a new method for localizing electrical activity of the brain. International Journal of Psychophysiology 18, 49–65.

Pascual-Marqui, R.D., Esslen, M., Kochi, K., Lehmann, D., 2002. Functional imaging with low-resolution brain electromagnetic tomography (LORETA): a review. Methods and Findings in Experimental and Clinical Pharmacology 24C, 91–95.

Peterson, K.M., Nichols, T.E., Poline, J.B., Holmes, A.P., 1999. Statistical limitations in functional neuroimaging. II. Signal detection and statistical inference. Philosophical Transactions of the Royal Society of London, Series B 354, 1261–1281.

Raz, J., 1989. Analysis of repeated measurements using nonparametric smoothers and randomization tests. Biometrics 45, 851–871.

Uhlig, H., 1994. On singular Wishart and singular multivariate Beta distribution. Annals of Statistics 22 (1), 395–405.

Vasey, M.W., Thayer, J.F., 1987. The continuing problem of false positives in repeated measures ANOVA in psychophysiology: a multivariate solution. Psychophysiology 24 (4), 479–486.

Worsley, K.J., Evans, A.C., Marrett, S., Neelin, P., 1992. A three dimensional statistical analysis of CBF activation studies in human brain. Journal of Cerebral Blood Flow and Metabolism 12, 900–918.

Worsley, K.J., Andermann, M., Koulis, T., MacDonald, D., Evans, A.C., 1999. Detecting changes in non-isotropic images. Human Brain Mapping 8, 98–101.