INVESTIGATIONS OF DIPOLE LOCALIZATION ACCURACY IN MEG USING THE BOOTSTRAP

F. Darvas [*], *M. Rautiainen*[†], *S. Baillet*[‡], *A. Ossadtchi*[§], *J. C. Mosher*[¶], *R. M. Leahy*[‖]

Signal and Image Processing Institute
University of Southern California
Los Angeles, CA

ABSTRACT

[1] We describe the use of the nonparametric bootstrap to investigate the accuracy of current dipole localization from magnetoencephalography (MEG) studies of event related neural activity. The bootstrap is well suited to analysis of event-related MEG data since the experiments are often repeated 100 or more times and averaged to achieve acceptable SNRs. The set of repetitions or "epochs" can be viewed as a set of i.i.d. realizations of the brain's response to the experiment. Bootstrap resamples can be generated by sampling from these epochs and averaging. In this study we applied the bootstrap resampling technique to MEG data from a somatotopic experiment. Four fingers of the right and left hand of a healthy subject were electrically stimulated, and about 400 trials per stimulation were recorded and averaged in order to measure the somatotopic mapping of the fingers in the S1 area of the brain. Based on the single trial recordings for each finger, we performed 5000 bootstrap resamples. We reconstructed dipoles from these resampled averages, using the RAP-MUSIC source localization algorithm. To find the correspondences between multiple sources in each resample dipoles with similar time-series and forward fields were assumed to represent the same source. These dipoles were then clustered using a GMM (Gaussian Mixture Model) clustering algorithm, using their combined normalized time-series and topography as feature vectors. The mean and standard deviation of the dipole position and the dipole time-series in each cluster were computed to provide estimates of the accuracy of the reconstructed source locations and time-series.

[*] University of Southern California, Los Angeles, CA
[†] Tampere University of Technology, Tampere, Finland
[‡] Cogntive Neuroscience & Brain Imaging Laboratory, Paris, France
[§] University of Southern California, Los Angeles, CA
[¶] Los Alamos National Laboratory, NM
[‖] University of Southern California, Los Angeles, CA

[1] Acknowledgement: This work supported by grant R01 EB002010 from the National Institute of Biomedical Imaging and Bioengineering. The authors thank Sabine Meunier for providing the experimental MEG data.

1. INTRODUCTION

The ECD (Equivalent Current Dipole) is a widely used model for event-related neuronal activity. The location, orientation, and time series of ECDs can be estimated from noninvasive surface measurements of the associated magnetic fields and electric potentials generated by the human brain. The ECD model yields a fixed number of source locations and source time-series from each data set. These dipole source positions can be estimated to an arbitrary precision by means of nonlinear optimization [1]. However, because of nonlinearities in the model, the method is sensitive to noise and can depend critically on the number of dipoles and their relative locations. Thus, the arbitrary precision of the localization results can be misleading, especially in studies where differences in source localizations due to different experimental conditions are analyzed. To achieve acceptable estimates, reconstruction is typically performed on data sets averaged over 10's or 100's of repetitions of the same experiment.

Each averaged data set produces a point estimate of the ECD locations and time-series. It is important for interpretation of these results that some measure of uncertainty is also provided. Lower bounds on parameter variances, using the Cramer-Rao inequality, were described for this problem in [2] and [3]. While the bounds were shown to be reasonably tight in simulation studies, they assume stationary Gaussian noise and a deterministic time series for each source. In practice, these assumptions may not hold. The bootstrap method [4] provides an alternative nonparametric method for assessing the reliability of the estimated sources. Since each experiment is typically repeated many times, bootstrap resamples can be generated by sampling with replacement from the set of repeated trials. The advantage of the bootstrap approach is that no specific assumptions are made regarding the distribution of either the noise or of the dipole time series. A similar approach to analysis of event-related data is described in [5], but the analysis was applied only to the scalp data rather than brain sources estimated from this data.

We applied the bootstrap method to MEG data from a somatotopic experiment in order to estimate the accuracy and reliability of estimated ECDs. Each of the four measurements, which consisted of an electric stimulation of the thumb, index, middle and small finger, was resampled 5000 times. We reconstructed sources from these resamples using the RAP-MUSIC algorithm [6]. Since more than one dipolar source was reconstructed per bootstrap resample of the data, the resulting dipoles were clustered using their normalized time-series and topography.

2. METHODS

2.1. MEG data aquisition

The somatosensory measurement was performed on a healthy right-handed male. The stimulation was an electrical square-wave pulse delivered to four fingers of each hand: thumb, index, middle, and pinky. The stimulation was applied between the middle and distal phalanxes of each finger and the stimulation order was randomized. The pulse duration was 0.2 milliseconds and the amplitude was set to twice the perceptual threshold. The interstimulus interval varied randomly between 350 ms and 550 ms. The magnetic fields were recorded with a CTF Systems Inc. Omega 151 MEG system. For each finger a 300 ms interval, including a 50ms pre-stimulus interval, was recorded. The number of single trials per data set ranged from 386 to 415. The DC-offset of the gradiometers was removed from all single trials based on the pre-stimulus interval and the trials were filtered from 3 Hz to 70 Hz.

2.2. Bootstrap resampling

Bootstrap resampling is based on the approximation of the pdf (probability density distribution) F of a random variable X by its empirical pdf $\hat{F}$. With n realizations of X, $\mathbf{x} = \{x_1, x_2, ..., x_n\}$, the empirical pdf is given by

$$\hat{F}(x_i) = \frac{\#x = x_i}{n}, \tag{1}$$

with $\#x = x_i$ being the number of realizations of x that are equal to x_i. If the samples are independent of each other, a bootstrap resample $\mathbf{x}^*$ of $\mathbf{x}$ can be generated by randomly drawing n samples, with replacement, from $\mathbf{x}$. A statistic $\hat{\theta}(\mathbf{x})$, can be resampled by $\hat{\theta}_i = \hat{\theta}(\mathbf{x}_\mathbf{i}^*)$, where $\mathbf{x}_\mathbf{i}^*$ is the ith bootstrap resample of $\mathbf{x}$. If a total of B bootstrap resamples are generated, the mean and standard deviation of $\hat{\theta}$ are simply given by:

$$\bar{\hat{\theta}}_B = \frac{1}{B}\sum_{i=1}^{B}\hat{\theta}_i \tag{2}$$

$$\hat{\sigma}_B = \sqrt{\frac{1}{B-1}\sum_{i=1}^{B}(\hat{\theta}_i - \bar{\hat{\theta}}_B)^2} \tag{3}$$

The same procedure can be applied if the random variables are replaced with random matrices, i.e. $\mathbf{X} = \{\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_n\}$. In the case of single trial MEG data, the matrices $\mathbf{X}_i$, are the spatiotemporal data matrices $\mathbf{D}_i$ where n is the number of single trials. The statistic from which the dipole locations are subsequently estimated is the average,

$$\bar{\mathbf{D}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{D}_i$$

Let $\mathbf{R}(\bar{\mathbf{D}}) = \{\mathbf{r}_1(\bar{\mathbf{D}}), \mathbf{r}_2(\bar{\mathbf{D}}), ..., \mathbf{r}_s(\bar{\mathbf{D}})\}$ be the locations of the s reconstructed dipoles from the average data with $\mathbf{P}(\bar{\mathbf{D}})$ the corresponding set of dipole forward fields, and $\mathbf{T}(\bar{\mathbf{D}})$ the corresponding set of time-series. Then bootstrap resamples are given by

$$\mathbf{R}_i = \mathbf{R}(\bar{\mathbf{D}}_i) \;\; \mathbf{P}_i = \mathbf{P}(\bar{\mathbf{D}}_i) \;\; \mathbf{T}_i = \mathbf{T}(\bar{\mathbf{D}}_i)$$

where $\bar{\mathbf{D}}_i$ is the ith bootstrap resample of the average. The mean and standard deviation of $\mathbf{R}_i$, $\mathbf{P}_i$ and $\mathbf{T}_i$ cannot be directly calculated using equation (3), because the number s_i of dipoles reconstructed by RAP-MUSIC from the ith resample, $\bar{\mathbf{D}}_i$, is not fixed. Also, the ordering of the dipoles in each resample is arbitrary, so that before we can compute bootstrap statistics for each source, we must first establish the correspondences between sources in each bootstrap resample. We determine the correspondences through clustering of the bootstrap dipoles. To perform clustering we group all estimated source locations together,

$$\mathbf{R}_{all} = \{\mathbf{r}_{11}, \mathbf{r}_{12}, ..., \mathbf{r}_{1s_1}, ..., \mathbf{r}_{Bs_B}\},$$

with identical ordering of their corresponding forward fields and time-series. We then perform a clustering procedure to divide $\mathbf{R}_{all}$ into c subsets,

$$\mathbf{R}_{all} = \{\mathbf{R}_1, \mathbf{R}_2, ..., \mathbf{R}_c\}.$$

where each cluster $\mathbf{R}_j$ is assumed to represent a single ECD source of brain activity at location $\bar{\mathbf{R}}_j$ with topography $\bar{\mathbf{P}}_j$ and time-series $\bar{\mathbf{T}}_j$, $j = 1, \ldots c$, where in each case averages are computed over the bootstrap estimated sources within each cluster. Similarly, we can compute the standard deviation, using eqn. (3), of the location, topography and time series for each cluster.

2.3. Clustering of Bootstrap results

The clustering of the reconstructed sources is performed by applying a GMM algorithm to a set of feature vectors

$$\{\mathbf{f}_1, \mathbf{f}_2, ..., \mathbf{f}_S\} \quad S = \sum_{i=1}^{B} s_i,$$

which are associated with each of the sources ([7], [8]). We used the combined normalized forward field and time-series for each source:

$$\mathbf{f}_i = \begin{pmatrix} \hat{\mathbf{t}} \\ \hat{\mathbf{p}} \end{pmatrix}$$

The maximum number of clusters, which is a parameter for the GMM algorithm used here, was set to the maximum number of dipoles over all bootstrap resamples. Clustering was performed for the data from each finger separately.

3. RESULTS

3.1. Reconstruction on the average data

The mapping of the sources from the average data for the right-hand stimulation and the time-series in the contralateral (left hemisphere) S1 primary somatsensory region of the brain are shown in fig. 1. The spatial locations of these

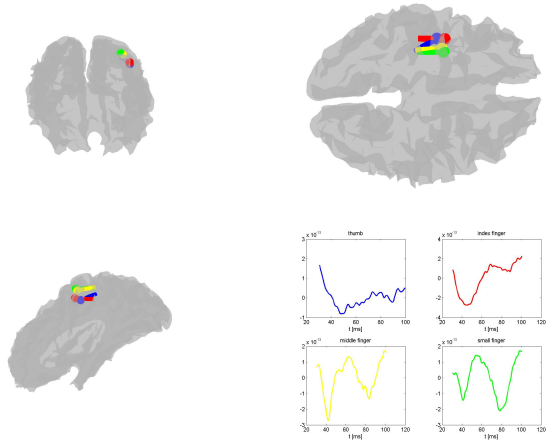

Fig. 1. a) The original somatotopy for the right hand stimulation. The sources are shown as colored points on the cortical white matter surface. The views shown are front, top and left. Blue= thumb, red=index finger, yellow=middle finger and green=small finger. Orientation of the dipoles is indicated by the angle of the rod extending from the sphere representing each source. b) Timeseries of the S1 sources. The same color code is used as in a).

S1 sources are listed in table 1.

stimulation	x	y	z
thumb	37.1	29.1	93.4
index	30.6	34.7	100.0
midlle	28.6	28.2	107.8
small	29.8	26.0	109.3

Table 1. Location of S1 activity in the average data for the 30-100 ms time interval. All values are in mm.

In addition to S1 activity, reconstruction over the interval from 30 - 100 ms also produced sources at other locations. In order to identify S1 activity, which typically peaks around 40 ms, those sources that had the strongest signal power (i.e. squared integrated time-series) in the interval from 30-50 ms were selected as S1. The location of these sources follows the homuncular cortical representation of the fingers as described in [9]. The peak activity for all four fingers takes place between 40 ms and 50 ms. Since areas S1 and S2 are close to each other, it appears that the time series of dipoles accounting for the S1 activity also contain evidence of S2 activity with secondary peaks around 80ms. Activity was also found in the ipsilateral hemisphere (i.e. the right hemisphere for this experiment), in the lower precentral area, rather than the right S1/S2 area as would have been expected from the literature.

3.2. The bootstrapped data

After bootstrapping, a total of 86830 sources were found from 5000 resamples for each of the four fingers. The average number of sources was 4.3 per resample, and all sources were grouped into 29 clusters. Clusters with a larger spatial extent (i.e. a standard deviation $\geq$ 20 mm) were excluded from further analysis, as these clusters most likely represent spurious sources with weak source strength. We show here only the clusters corresponding to S1 activity which were selected in a similar manner to the S1 sources for the original data as described above. The S1 clusters for each finger are shown in fig. 2.

The spatial ordering of the S1 activity clusters reflects the original somatotopy. The standard deviation for this activity was between 2.0 mm and 3.4 mm. Table 2 lists the cluster locations, standard deviation, and bias (relative to the locations estimated from the original data) for the clusters in S1. The bootstrapped data shows similar location uncertainty for each digit, but a substantial bias for the thumb and index finger.

4. DISCUSSION

We applied the bootstrap method to MEG data from a somatotopic expereriment in order to asses the combined sta-

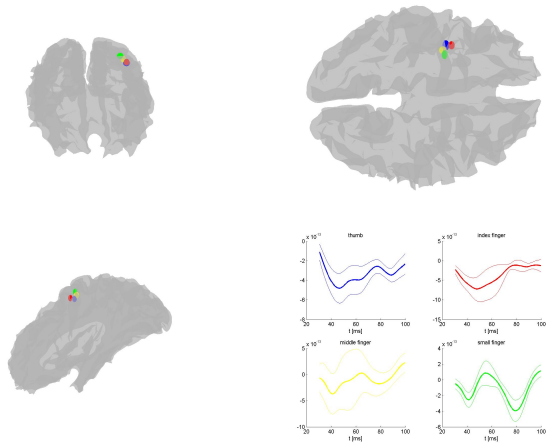

Fig. 2. a) The resampled S1 somatotopy for right hand stimulation. The sources are shown as colored ellipsoids, with semiaxes equal to the standard deviation along the three eigenvectors of their respective cluster covariance. The clusters shown here have the strongest signal power for the interval 40-50 ms and 70-80 ms on the left hemisphere. b) The bootstrapped time-series for these clusters. The thick line is the mean time-series of the cluster, the thin lines are +/- one standard deviation.

tistical variability in the data and the accuracy of the applied dipole reconstruction method. The sources which were reconstructed on the resampled data were clustered using their combined time-series and topography as feature vector. From these clusters, the average location and the standard deviation of the location of the sources in S1 were calculated. For the time intervals in which one would expect S1 activity, the bootstrap method revealed a stable somatotopic mapping on the left hemisphere, with a spatial standard deviation of less than 3.4 mm. However, there was an apparent significant bias to the source locations in the original averaged data, especially for the thumb and the index finger. Since the average data can be seen as just another bootstrap resample,

stimulation	x	dx	y	dy	z	dz	Bias
thumb	31.3	2.3	31.8	3.4	101.2	3.0	10.1
index	35.1	2.1	31.8	2.8	102.2	3.2	5.8
middle	29.2	2.6	28.5	2.8	105.0	3.3	2.9
small	30.5	2.0	25.5	3.2	108.2	3.2	1.4

Table 2. Location in x, y and z, and corresponding standard deviations (dx, dy, dz) for the S1 activity clusters for the 30-100 ms interval. The last column is a bias estimate computed as the distance between the cluster centroid and the location of the source as computed from the original data.

the computed "bias" does not represent the true bias of the estimator but rather is an indication of a potentially large error in the reconstruction from the original averaged data. For the thumb, the center of gravity of the bootstrapped reconstruction lies over 1cm from the position where the original average would have placed the source.

These results demonstrate the potential for bootstrap analysis to provide users of MEG data a nonparametric indication of the reliability of estimated sources, which is very important in the interpretation of brain mapping data for the two primary applications of MEG: presurgical brain mapping and cognitive neuroscience.

5. REFERENCES

[1] M Scherg, "Fundamentals of dipole source potential analysis.," *Advances of Audiology*, vol. 6, pp. 40–69, 1990.

[2] JC Mosher, ME Spencer, RM Leahy, and PS Lewis, "Error bounds for eeg and meg dipole source localization.," *Electronencephalography and Clinical Neurophysiology*, vol. 86(5), pp. 303–321, 1993.

[3] CH Muravchik and A Nehorai, "Eeg/meg error bounds for a static dipole source with a realistic headmodel.," *IEEE Transactions on Signal Processing*, vol. 49(3), pp. 470–484, 2001.

[4] B Efron, "Bootstrap methods: Another look at the jackknive.," *The Annals of Statistics*, vol. 7(1), pp. 1–26, 1979.

[5] F DiNocera and F Ferlazzo, "Resampling approach to statistical inference: Bootstrapping from event-related potentials data," *Behaviour Research Methods, Instruments and Computers*, vol. 32, no. 1, pp. 111–119, 2000.

[6] JC Mosher and RM Leahy, "Source localization using recursively applied and projected (RAP) music.," *IEEE Trans. on Signal Processing*, vol. 47 2, pp. 332–40, Feb. 1999.

[7] IT Nabney, *Netlab*, Springer Verlag, Heidelberg, 2002.

[8] SJ Roberts, DH, I Rezek, and WD Penny, "Bayesian approaches to gaussian mixture modeling," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 11, pp. 1133–1142, 1998.

[9] A Nakamura, T Yamada, A Goto, T Kato, K Ito, Y Ave, T Kachi, and R Kakigi, "Somatosensory homunculus as drawn by meg.," *Neuroimage*, vol. 7, pp. 377–386, 1998.