INSTITUTE OF PHYSICS PUBLISHING PHYSICS IN MEDICINE AND BIOLOGY

Phys. Med. Biol. **48** (2003) 2561–2573 PII: S0031-9155(03)59691-9

Assessment criteria for MEG/EEG cortical patch tests

Chang-Hwan Im[1], Kwang-Ok An[1], Hyun-Kyo Jung[1], Hyukchan Kwon[2] and Yong-Ho Lee[2]

[1] ENG420-040, School of Electrical Engineering, Seoul National University, Shillim-dong, Kwanak-gu, Seoul 151-742, Korea

[2] Korea Research Institute of Standards and Science, PO Box 102, Yuseong, Daejeon 305-600, Korea

E-mail: ichism@elecmech.snu.ac.kr

Received 18 February 2003, in final form 17 June 2003
Published 22 July 2003
Online at stacks.iop.org/PMB/48/2561

Abstract

To validate newly developed methods or implemented software for magnetoencephalography/electroencephalography (MEG/EEG) source localization problems, many researchers have used human skull phantom experiments or artificially constructed forward data sets. Between the two methods, the use of an artificial data set constructed with forward calculation attains superiority over the use of a human skull phantom in that it is simple to implement, adjust and control various conditions. Nowadays, for the forward calculation, especially for the cortically distributed source models, generating artificial activation patches on a brain cortical surface has been popularized instead of activating some point dipole sources. However, no well-established assessment criterion to validate the reconstructed results quantitatively has yet been introduced. In this paper, we suggest some assessment criteria to compare and validate the various MEG/EEG source localization techniques or implemented software applied to the cortically distributed source model. Four different criteria can be used to measure accuracy, degrees of focalization, noise-robustness, existence of spurious sources and so on. To verify the usefulness of the proposed criteria, four different results from two different noise conditions and two different reconstruction techniques were compared for several patches. The simulated results show that the new criteria can provide us with a reliable index to validate the MEG/EEG source localization techniques.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Magnetoencephalography (MEG) electroencephalography (EEG) are similar non-invasive techniques for localizing and characterizing the electrical activity of the cerebral cortex using

0031-9155/03/152561+13$30.00

electromagnetic measurements outside the head (Hämäläinen *et al* 1993). To localize the electric sources inside the brain, various source assumptions and reconstruction techniques have been proposed (Baillet *et al* 2001a). Among them, the distributed source model with minimum-norm estimate (MNE) has been widely studied since Hämäläinen and Ilmoniemi (1984, 1994). Their MNE selects the solution where the L2 norm of the current distribution was smallest. They also regularized the solution using truncated singular value decomposition (tSVD). This fundamental study has led to several modifications. For example, the low-resolution electrical tomography method (LORETA) proposed by Pascual-Marqui *et al* (1994) gave deeper sources the same opportunity of being nicely reconstructed by the MNE. Iterative focalization approaches such as FOCUSS (FOCal underdetermined system solution) were proposed to solve the underdetermined inverse problem more effectively and reconstruct more focalized solutions (Gorodnistky *et al* 1995).

Recently, to reduce the dimension of the source space, anatomical constraints have been widely used. Dale and Sereno (1993) first proposed constraining the source space into anatomically known locations (interface between white and grey matter of the cerebral cortex) and orientations (perpendicular to the cortical surface), and weighting the estimate based on *a priori* information. After Dale and Sereno's work, most of the recent publications on the distributed source model have adopted the anatomical constraints (Buchner *et al* 1997, Kincses *et al* 1999, Fuchs *et al* 1999, Gavit *et al* 2001, Baillet *et al* 2001b). The anatomically constrained distributed source model is usually called a cortically distributed source model (Baillet *et al* 2001a, 2001b).

There have been two main approaches to validate a method or an implemented software for the cortically distributed source model. These are experimental verification using a human skull phantom (Baillet *et al* 2001b) and numerical verification by cortical patch tests (Baillet and Garnero 1997, Kincses *et al* 1999, Mosher *et al* 1999a, Wagner *et al* 2002, Chupin *et al* 2002, Jerbi *et al* 2002). In general, the latter is much preferred to the former because it can deal with different kinds of noise conditions, it is easy to simulate, the cortical patches are more realistic than the point source experiments and the size of patch can be easily adjusted. However, in the case of the patch test, no well-established assessment criterion to evaluate the accuracy or noise-robustness of the solutions quantitatively has yet been introduced, whereas, for the point dipole test, well-established assessment criteria were introduced by Baillet *et al* (2001b). They proposed four assessment criteria:

- $\varepsilon_{\mathrm{max}}$: distance between the location of the maximum in absolute intensity of the estimated activity (J_{max}) and the actual source position;
- E_{max}: relative energy contained in J_{max} with regard to the global energy;
- ε_{G}: distance between the actual source location and the position of J_{grav}, the gravity centre of the source estimate;
- E_{s}: relative energy contained in spurious or phantom sources with regard to the original source energy.

Unfortunately, these criteria cannot be applied directly to the cortical patch test. Hence, we need to introduce modified assessment criteria to evaluate the results of the cortical patch tests. In this paper, we suggest four assessment criteria to measure accuracy, degrees of focalization, noise-robustness, existence of spurious sources and so on. It is shown, from four different case studies, that the proposed criteria can quantitatively measure those characteristics in a very effective manner.

In section 2, we explain the simulation set-up such as generation of the cortical patch and construction of forward data. In section 3, the proposed assessment criteria are explained. In section 4, the usage and limitations of the proposed criteria are discussed. In section 5, the

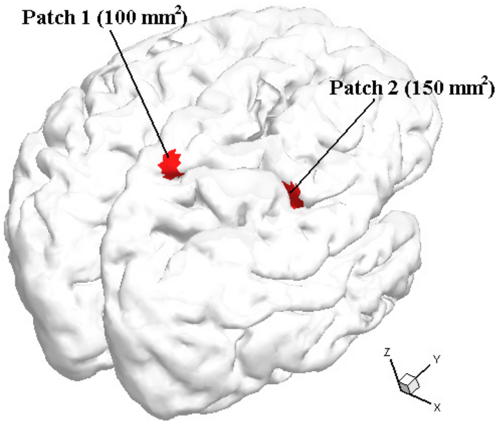

Figure 1. Tessellated cortical surface and an example of patch generation.

usefulness of the proposed criteria is verified by the application to different simulations and conclusions are presented in section 6.

2. Simulation set-up

A tessellated cortex was obtained after segmentation of high resolution MRI data. The tessellated cortical surface consists of 41 472 triangular elements and 20 864 vertices. In this paper, only the MEG simulation was considered for convenience sake[3]. The system configuration used for the simulation was 4D Neuroimaging *Neuromag* 122-channel whole-head system (http://www.4dneuroimaging.com/external-english/html/n122spec.html) with realistic position of the subject's head in the helmet.

To generate cortical patches and construct the forward data set, the concept of virtual area was adopted (Chupin *et al* 2002). The virtual area was assigned to each vertex as one third of the area of all triangles meeting at the vertex. This assumption is valid because the total area remains equal to the actual area of the full tessellation. The cortical patch was generated using the following process:

- a point is selected as a seed of an activation patch area;
- the patch area is extended by including neighbouring vertices around the patch;
- if the total virtual area of the cortical patch exceeds a predetermined surface area, the extension of the activation patch is terminated.

Each patch was made of a set of dipoles with constant current dipole moment density of 1 nAm mm^{-2}. Then, the current dipole moment at each vertex was calculated by the product of the current dipole moment density and the virtual area defined above (Chupin *et al* 2002). Figure 1 shows the tessellated cortical surface and an example of the patch generation. The virtual areas of the two example patches are 100 mm^2 and 150 mm^2.

We used a spherical head model for the forward calculation of magnetic field[4] (Sarvas 1987, Mosher *et al* 1999b). To be more realistic, white Gaussian noise was added to the

[3] In this study, only MEG simulation was performed. Please note that the main concern of the simulation is the distributed source models. We think that the measurement method is not a crucial problem in this study.

[4] Please note that this paper deals only with the assessment criteria to evaluate various inverse techniques. Practically, the accurate forward calculation with accurate head modelling is very important, but not crucial in this study. Hence, we used the spherical head model for convenience sake.

constructed forward data set. The white Gaussian noise was generated by simply drawing a zero-mean Gaussian-distributed random number for each sensor.

3. Suggested assessment criteria

In this study, we suggested four different criteria to measure accuracy, degrees of focalization, noise-robustness, existence of spurious solutions and so on. Each can be selectively used to test specific characteristics of reconstructed results.

3.1. Accuracy criterion I—geodesic distance between original and reconstructed sources: ε_d

To verify the accuracy of the reconstructed source distribution, the distance between original and reconstructed source positions is most widely used (Baillet *et al* 2001b). In the case of the cortically distributed source model, the distance from the peak position of the reconstructed source distribution to the patch's geometric centroid has been used to measure the error of the solution. Jerbi *et al* (2002) used the distance to the centre of mass of a patch's vertices as a measure of localization error, denoted as DTC. The method seems very easy to apply, but it has two crucial problems: the centre of mass of a patch's vertices may not be on the curved cortical surface and the Euclidean distance is simple to calculate but cannot be an accurate measure in the case of the highly curved cortical surface.

The first problem can be solved by adopting the basic concept of the equivalent current dipole (ECD) model. The centroid of a distributed source should be an intuitively appealing location for its equivalent source model. Therefore, we defined the centroid using the following process: (1) evaluating the momental value of an ECD by summing the current dipole moments at all vertices inside a cortical patch area and (2) estimating the best-fitted location by placing the ECD on each vertex and comparing the magnetic field from the ECD with that from the distributed source.

The other problem can be solved by the use of a cortically geodesic distance instead of the conventional Euclidean distance. Bartesaghi and Sapiro (2001) developed a system for the generation of curves on 3D cortical surfaces. The system which can be obtained via a website (http://www.ece.umn.edu/users/guille) enabled us to easily calculate the geodesic distance between two vertices on a cortical surface.

3.2. Accuracy criterion II—L2 norm error estimate: ε_{L2}

In the numerical field calculations with the finite element method (FEM) or boundary element method (BEM), the L2 norm is widely used as an error estimate (Burnett 1987). In this study, the L2 error norm denoted as $\varepsilon_{\mathrm{L2}}$ is defined as

$$\varepsilon_{\mathrm{L2}} = \left(\int_s [J_r - J_a]^2 \, \mathrm{d}S \right)^{1/2} \tag{1}$$

where S represents the whole cortical surface, J_r is the reconstructed dipole moment density and J_a is the actual dipole moment density.

3.3. Criterion to evaluate degrees of focalization: DF

Generally, the distributed source reconstructed by the MNE yields an over-smoothed source distribution, especially when using spatial smoothing algorithms such as LORETA

(Pascual-Marqui *et al* 1994). In the over-smoothed case, it is obvious that the reconstructed energy in a given patch area is smaller than the original one. When special focalizing techniques such as FOCUSS (Gorodnistky *et al* 1995) or multiresolution techniques (Gavit *et al* 2001) are applied, the energy reconstructed in the given area increases to generate an equivalent magnetic field with more focalized activating areas. Therefore, in this paper, to quantify *how well a method can reconstruct a focalized solution*, a criterion denoted as DF is defined as the ratio between the reconstructed and original energies stored in an original patch area. Then, higher DF implies that the method can reconstruct a more focalized source distribution.

3.4. Criterion to check the number of spurious sources: NSS

The current source distribution reconstructed by MNE generally includes several spurious or phantom sources, which usually stem from numerical errors, noisy conditions, lack of regularization techniques and so on. Although most of the spurious sources are small enough in size and magnitude to be ignored, relatively big ones are frequently generated and they are hard to distinguish from actual ones. Therefore, checking the existence of spurious sources is a very crucial work to validate the performance of the implemented software.

In this study, to check the existence of spurious sources, local peaks are detected by scanning all vertices on a cortical surface. The local peak is defined as a vertex that has larger moment density value than all its neighbouring vertices. After the local peak detection, a threshold is set as (maximum moment density) $\times \alpha$ $(0 < \alpha < 1)$. By changing the threshold gradually, the number of local peaks outside the original patch area (number of spurious sources: NSS) is recorded. Using the recorded data, we can estimate and compare the magnitudes of spurious sources in a very effective manner. Moreover, the noise-dependence of a method can also be revealed because the noisy condition generally yields more spurious sources with higher magnitude than the no-noise condition.

4. Usage and limitations of the proposed assessment criteria

Before the application of the proposed criteria, the limitations and usages of the assessment criteria should be discussed. First, note that the proposed criteria are applicable only when exact patch positions are given. In other words, it is obvious that these criteria cannot be applied to measured data, of which exact solutions cannot be estimated *a priori*. In most cases, when reconstructed source distributions are found using some inverse methods, the accuracy and fitness of the results can be recognized intuitively after visually comparing them with exact patch configurations. The main objective of this study is to confirm the accuracy and fitness of the solution in a more quantitative and objective manner. However, it is very difficult to define absolutely satisfactory criteria and even to assess if the defined criteria are reasonable or not. The authors tried to define robust criteria, but they still have some limitations.

First, ε_d should be carefully applied to multi-patch excitations. When several patches are activated simultaneously, we have no strategy to find locations of the multiple maxima automatically. In other words, we should divide the searching regions manually and find maxima in each region. Therefore, the ε_d cannot be applied automatically to a large number of repeated simulations such as the Monte Carlo studies, which should be studied further.

In the case of ε_d, another two questions may be raised. The first one is whether it is really adequate to compare distributed sources by their centres. It may be questioned that the location of the maximum of the MNE may not adequately represent the information contained in the MNE. However, the authors think what it is also important and should be considered is how far the true and estimated sources are apart. To consider the extended information contained in the MNE, another accuracy criterion—$\varepsilon_{\mathrm{L2}}$—was also proposed in this study. To consider both simultaneously, further studies should be performed in the future. The second question is that the ε_d defined in this study may not always represent the accuracy of the reconstructed distributions well. There can be some special cases. Although the geodesic distance is physiologically more meaningful than the Euclidean distance, it sometimes fails to assess the power of an inverse method. The special exception stems from the fact that the spatial resolution of MEG or EEG may be insufficient to distinguish two near walls of sulci. When applying the geodesic distance to sources reconstructed on a wrong wall of a sulcus, the estimator yields a very large error, but, as a mathematical tool, the localizer could be perfect. In that case, the conventional DTC that uses Euclidean distance as a distance measure is more reasonable. Then, a user can select a more proper one considering the positions of assumed patches. It is obvious that the ε_d is more adequate when two walls of sulci are far enough apart to distinguish in the given spatial resolution. The other special case is when a source patch is located on a gyrus, which has a perpendicular normal direction and is hard to measure using MEG. In that case, the equivalent dipole position of the patch that generates the most similar magnetic field may be far from the centre of gravity of the patch. However, it is obvious that the 'blind zone' is an inevitable drawback of the MEG. Then, the use of equivalent position is more reasonable than that of the centre of gravity because the comparison of the equivalent position and the reconstructed maximum better represents the power of the inverse algorithm. Anyway, note that the choices of the activation patches and proper assessment criteria are dependent on the users of the criterion.

Before using the second accuracy criterion $\varepsilon_{\mathrm{L2}}$, the sizes of patches should be considered. When the sizes of the patches are very small, the estimator will always yield a huge error and the comparison may be meaningless. Then, it is proper to use Baillet *et al*'s (2001b) criteria that were introduced in section 1.

In the case of the DF, it can be questioned whether overestimating the source strength always improves the estimate DF. However, note that the criterion checks how much the source is focalized. In most cases, the MNE yields underestimated solutions except when special focalizing algorithms such as FOCUSS are used. This criterion is devised to validate the special focalization techniques and sometimes the overestimation also implies meaningful information. In any of these cases, it is also obvious that the DF gives meaningful information only when patches have sufficient areas and the accuracy measured by ε_d is guaranteed to some extent.

When applying the NSS to multiple-patch simulations, one should recall that some focalizing techniques such as FOCUSS have the possibility of missing some sources that are relatively small in strength. Moreover, if there are several sources that have different currents, some small sources can be underestimated during the post-processing. In those cases, the NSS may not work well. However, those cases are typical and unavoidable problems of general inverse algorithms. The special cases can be identified just by visually observing the reconstructed source distributions. There is no need to apply the NSS for these cases.

In conclusion of the discussions, the proposed assessment criteria are adequate to be applied to relatively large patches with solutions having not much error from exact patch areas. They are very useful when the fitness of the reconstructed source distribution can be estimated roughly and one wants to confirm the result quantitatively. Although they have little limitation in applications to all general problems, note that the locations and sizes of the source patches are dependent upon the users' choice and the criteria are useful enough to

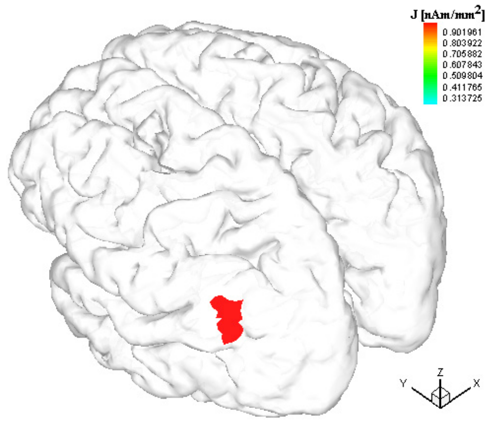

Figure 2. Cortical activating patch used for simulation (virtual area = 200 mm^2).

compare most distinct results because the criteria are devised to yield best results when the original and reconstructed solutions perfectly match each other.

5. Simulation and results

To verify the proposed assessment criteria, reconstructed results from the following four different simulation studies were compared:

- Case I: no-noise, MNE;
- Case II: 15 dB white Gaussian noise, MNE;
- Case III: no-noise, FOCUSS;
- Case IV: 15 dB white Gaussian noise, FOCUSS.

For cases I and II, general MNE with truncated SVD was used. The minimum-norm solver used was SGELSD in LAPACK driver routine, which can be obtained from a website (http://www.netlib.org/lapack/single/sgelsd.f). For the cases III and IV, FOCUSS was applied (Gorodnitsky *et al* 1995, Baillet *et al* 2001b). FOCUSS is a kind of a recursive weighted MNE to concentrate the solution in focal regions. By applying weights to the forward gains of the sources according to their moments at the previous iteration, focal resolution could be obtained. For all the simulations, the number of FOCUSS iterations was set to 5.

Figure 2 shows a cortical activating patch used for the test simulation. The virtual area of the patch was about 200 mm^2 and the current moment density at each vertex was set to 1 nAm mm^{-2}. Figures 3(a) and (b) show the source distribution reconstructed by general MNE for no-noise data (case I) and noisy data (case II), respectively. Figures 4(a) and (b) show the distribution reconstructed by FOCUSS for no-noise data (case III) and noisy data (case IV), respectively. Only sources that exceed 0.25 $\times$ (maximum magnitude) are illustrated. We can see from these figures that noisy data generate more spurious sources and FOCUSS yields more focalized resolution with a reduced number of spurious sources.

Then, we will verify the above facts quantitatively with the suggested assessment criteria. Table 1 presents the three criteria ε_d, $\varepsilon_{\mathrm{L2}}$ and DF for cases I–IV, and table 2 shows the variation

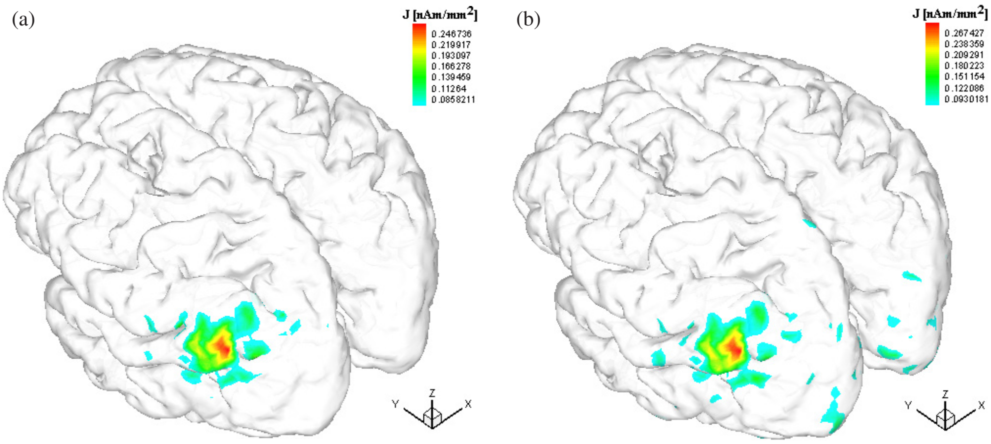

Figure 3. Source distribution reconstructed by general MNE: (a) case I (no-noise), (b) case II (15 dB white Gaussian noise), sources which exceed 0.25 × (maximum magnitude) are presented.

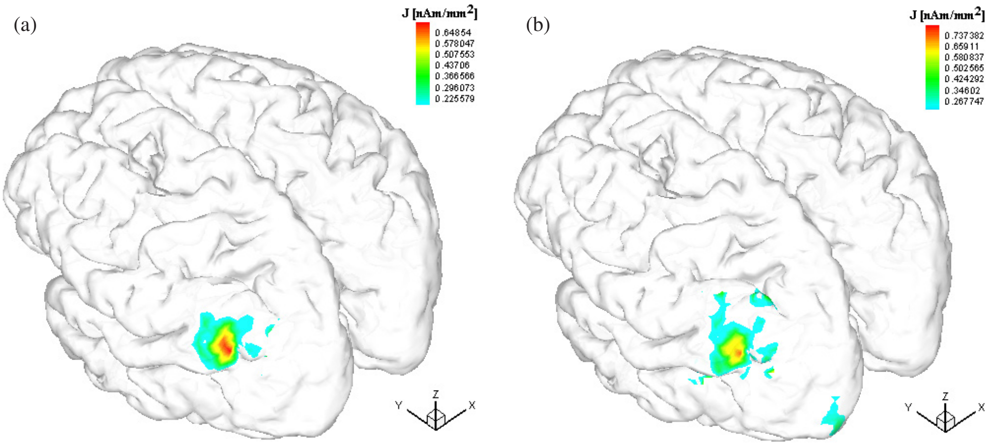

Figure 4. Source distribution reconstructed by FOCUSS: (a) case III (no-noise), (b) case IV (15 dB white Gaussian noise), sources which exceed 0.25 × (maximum magnitude) are presented.

Table 1. Value of ε_d, ε_{L2} and DF for four test cases.

	ε_d (mm)	ε_{L2} (nAm mm^{-1})	DF (%)
Case I	4.71	6.749	19.755
Case II	9.32	8.397	19.520
Case III	3.50	5.391	51.068
Case IV	5.23	7.799	50.409

Case I—no-noise, MNE; case II—noisy, MNE; case III—no-noise, FOCUSS; case IV—noisy, FOCUSS.

of NSS according to α when threshold is defined as (maximum moment density) $\times\ \alpha$ $(0 < \alpha < 1)$. From these results, we can see the following:

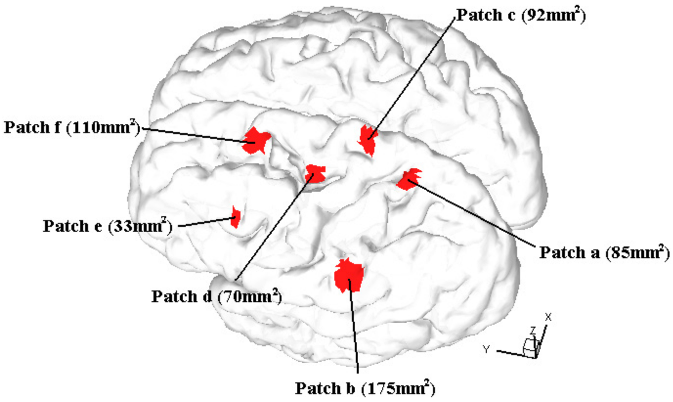

Figure 5. Configurations of several test source patches with various positions and sizes.

Table 2. Variation of NSS according to α when threshold is defined as (maximum moment density) $\times\ \alpha$ $(0 < \alpha < 1)$.

α	0.3	0.4	0.5	0.6	0.7	0.8
Case I	26	14	7	5	1	0
Case II	114	60	33	13	6	3
Case III	10	5	1	0	0	0
Case IV	24	15	6	2	1	0

Case I—no-noise, MNE; case II—noisy, MNE;
case III—no-noise, FOCUSS; case IV—noisy, FOCUSS.

- From ε_d and ε_{L2}, we can see that noisy data highly degrade the solution accuracy. The accuracy can be improved by using the FOCUSS algorithm.
- From DF, we can easily observe that the reconstructed distribution is considerably focalized by adopting the FOCUSS algorithm. Meanwhile, the DF is not affected much by the additional noise.
- From NSS, we can see that noisy data generate much more spurious sources in magnitude or number. The spurious sources can be dramatically reduced by adopting the FOCUSS algorithm. The noise-robustness of the solution is also improved by the FOCUSS as seen from the results.

The above facts extracted from the proposed assessment criteria coincide with those from intuitive understanding of figures 3 and 4. However, the intuitive understanding is inadequate to compare various techniques or conditions because the solution accuracy is hard to estimate intuitively.

The same simulations were performed for several patches with different sizes and locations. Figure 5 shows their configurations. The same conditions and inverse algorithms were tested and the results are shown in tables 3 and 4. We can see from the tables that similar trends could be observed for the several cases and the same conclusions as the previous ones could be made for any of these cases.

Among the several patches used in the previous simulation, two patches b and f were activated simultaneously as shown in figure 6. The current moment density at each vertex was set to 1 nAm mm^{-2}. The reconstructed current density distributions for the given four conditions are shown in figures 7(a)–(d). Tables 5 and 6 show the assessed quantities for

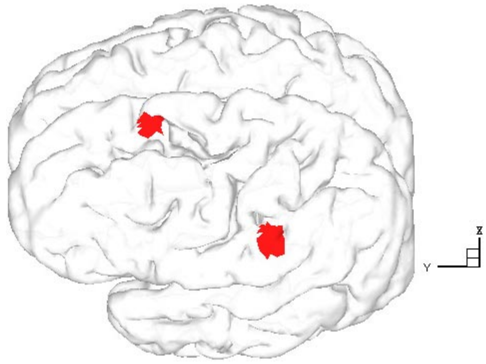

Figure 6. Two simultaneously activated patches (patches *b* and *f* in figure 5).

Table 3. Values of ε_d, ε_{L2} and DF for various patches *a*–*f*.

Patches	Cases	ε_d (mm)	ε_{L2} (nAm mm^{-1})	DF (%)
a	Case I	7.00	4.433	6.894
	Case II	8.72	5.606	6.921
	Case III	4.94	4.079	31.829
	Case IV	6.21	4.859	33.734
b	Case I	2.85	6.364	17.039
	Case II	5.22	8.896	18.494
	Case III	2.84	3.987	73.112
	Case IV	3.13	5.148	72.091
c	Case I	8.11	5.184	10.856
	Case II	8.47	6.574	9.914
	Case III	1.85	4.002	47.470
	Case IV	3.93	4.986	51.526
d	Case I	3.89	4.322	11.907
	Case II	4.27	6.148	10.950
	Case III	1.58	3.182	54.741
	Case IV	2.24	4.154	52.129
e	Case I	5.74	2.893	7.588
	Case II	7.75	4.011	7.514
	Case III	5.14	2.492	33.063
	Case IV	5.90	3.238	32.717
f	Case I	5.55	5.248	8.703
	Case II	9.24	7.035	8.542
	Case III	2.93	3.878	52.043
	Case IV	4.03	4.946	50.311

Case I—no-noise, MNE; case II—noisy, MNE;
case III—no-noise, FOCUSS; case IV—noisy, FOCUSS.

each criterion, where roughly divided regions for searching local maximum points had been assumed before measuring the distance. As seen from the tables, similar conclusions could be deduced in the case of two patch activations.

In conclusion, the proposed assessment criteria provide us with a more detailed and quantitative index to evaluate and compare different reconstructed results.

Table 4. Variation of NSS according to α when the threshold is defined as (maximum moment density) $\times$ α ($0 < \alpha < 1$)—simulations for various patches a–f.

Patches	Cases	0.3	0.4	0.5	0.6	0.7	0.8
a	Case I	13	11	6	2	0	0
	Case II	37	24	13	7	3	1
	Case III	7	3	0	0	0	0
	Case IV	15	9	3	2	0	0
b	Case I	8	2	1	0	0	0
	Case II	33	8	5	3	0	0
	Case III	0	0	0	0	0	0
	Case IV	4	2	1	0	0	0
c	Case I	12	6	3	1	0	0
	Case II	27	15	6	2	2	0
	Case III	1	0	0	0	0	0
	Case IV	6	2	0	0	0	0
d	Case I	25	13	5	3	0	0
	Case II	89	28	14	6	4	1
	Case III	3	0	0	0	0	0
	Case IV	10	6	3	1	0	0
e	Case I	10	6	4	0	0	0
	Case II	23	11	7	2	2	1
	Case III	2	1	0	0	0	0
	Case IV	5	4	4	1	0	0
f	Case I	24	14	11	5	2	0
	Case II	104	59	22	14	7	2
	Case III	7	2	0	0	0	0
	Case IV	16	5	2	2	1	0

Case I—no-noise, MNE; case II—noisy, MNE;
case III—no-noise, FOCUSS; case IV—noisy, FOCUSS.

Table 5. Value of ε_d, $\varepsilon_{\mathrm{L2}}$ and DF for two-patch activation.

	ε_d (mm) Patch b	ε_d (mm) Patch f	$\varepsilon_{\mathrm{L2}}$ (nAm mm^{-1})	DF (%)
Case I	2.84	6.95	8.238	14.064
Case II	3.37	7.10	8.925	13.880
Case III	2.81	6.13	7.172	36.706
Case IV	3.07	6.58	7.754	35.427

Case I—no-noise, MNE; case II—noisy, MNE;
case III—no-noise, FOCUSS; case IV—noisy, FOCUSS.

Table 6. Variation of NSS for two-patch activation according to α when threshold is defined as (maximum moment density) $\times$ α ($0 < \alpha < 1$).

α	0.3	0.4	0.5	0.6	0.7	0.8
Case I	16	6	3	3	1	0
Case II	21	8	5	4	3	1
Case III	5	1	0	0	0	0
Case IV	9	3	1	0	0	0

Case I—no-noise, MNE; case II—noisy, MNE;
case III—no-noise, FOCUSS; case IV—noisy, FOCUSS.

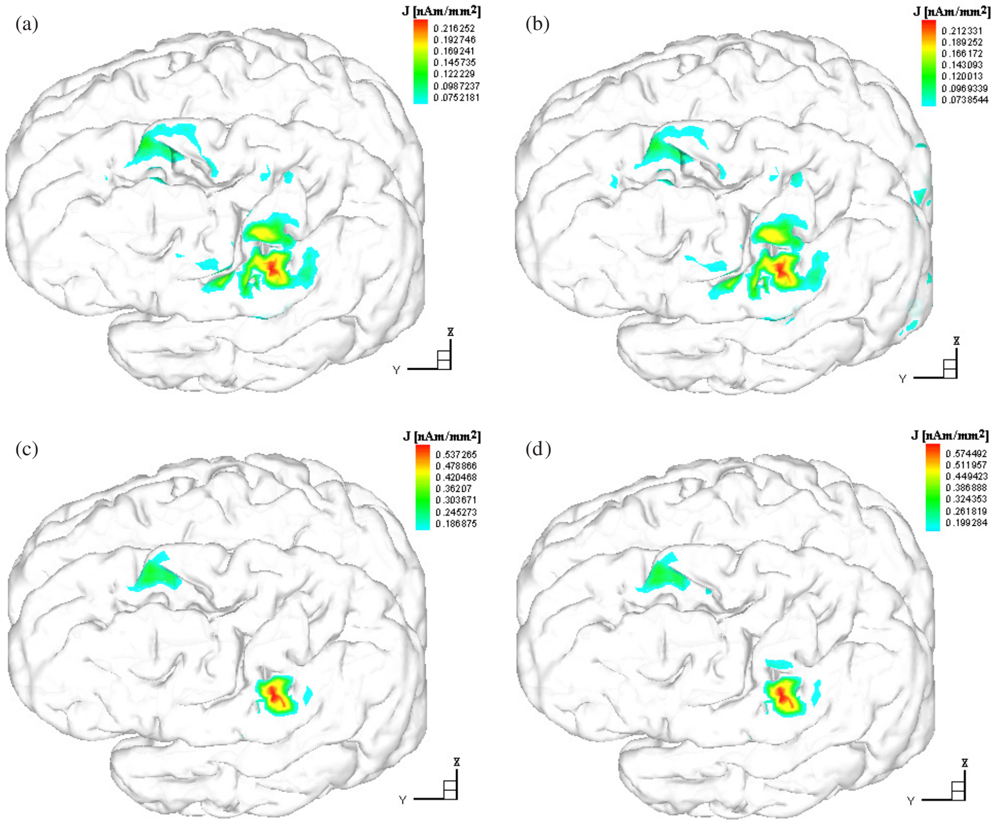

Figure 7. Reconstructed source distribution for four cases: (a) case I (no-noise, MNE), (b) case II (noisy, MNE), (c) case III (no-noise, FOCUSS), (d) case IV (noisy, FOCUSS), sources which exceed 0.25 × (maximum magnitude) are presented.

6. Conclusion

In this paper, we suggested four assessment criteria, which are denoted as ε_d, $\varepsilon_{\mathrm{L2}}$, DF and NSS, to validate MEG/EEG source localization methods for the cortically distributed source model. The first two criteria ε_d and $\varepsilon_{\mathrm{L2}}$ were used to check the accuracy of a solution. The DF was applied to measure the degree of focalization of the solution. The NSS was used to check how many spurious sources are included in the solution. In addition, the noise-dependence of the solution could also be evaluated using some of those criteria.

The validity of the criteria was proved by the application of the proposed criteria to MEG problems with several patch activations. The criteria were applied to four different reconstruction results: (1) no-noise, general MNE, (2) noisy, general MNE, (3) no-noise, FOCUSS and (4) noisy, FOCUSS. The suggested criteria showed superiority of the FOCUSS algorithm very well. From the MEG simulation, we could conclude that the proposed assessment criteria are very promising tools to validate various software or newly developed techniques for the cortically distributed source model.

Note that these criteria are not absolute and may be useless in some special cases as discussed previously. However, we are convinced that the criteria are reasonable enough and will be used as a very reliable assessment measure for the cortically distributed source model. Further study should be performed to improve and strengthen the compatibility of the assessment criteria.

Acknowledgments

This work was supported by a grant of the NRL project from the Ministry of Science and Technology, Republic of Korea.

References

Baillet S and Garnero L 1997 A Bayesian approach to introducing anatomo functional priors in the EEG/MEG inverse problem *IEEE Trans. Biomed. Eng.* **44** 374–85

Baillet S, Mosher J C and Leahy R M 2001a Electromagnetic brain mapping *IEEE Signal Process. Mag.* (Nov) 14–30

Baillet S, Riera J J, Marin G, Mangin J F, Aubert J and Garnero L 2001b Evaluation of inverse methods and head models for EEG source localization using a human skull phantom *Phys. Med. Biol.* **46** 77–96

Bartesaghi A and Sapiro G 2001 A system for the generation of curves on 3D brain images *Hum. Brain Mapp.* **14** 1–15

Buchner H, Knoll G, Fuchs M, Rienäcker A, Beckmann R, Wagner M, Silny J and Pesch Jorg 1997 Inverse localization of electric dipole current sources in finite element models of the human head *Electroenceph. Clin. Neurophysiol.* **102** 267–78

Burnett D S 1987 *Finite Element Analysis: From Concepts to Applications* (Reading, MA: Addison Wesley)

Chupin M, Baillet S, Okada C, Hasboun D and Garnero L 2002 On the detection of hippocampus activity with MEG *Proc. Conf. on Biomagnetism-BIOMAG 2002*

Dale A and Sereno M 1993 Improved localization of cortical activity by combining EEG and MEG with MRI surface reconstruction: a linear approach *J. Cogn. Neurosci.* **5** 162–76

Fuchs M, Wagner M, Kohler T and Wischmann H-A 1999 Linear and nonlinear current density reconstructions *J. Clin. Neurophysiol.* **16** 267–95

Gavit L, Baillet S, Mangin J F, Pescatore J and Garnero L 2001 A multiresolution framework to MEG/EEG source imaging *IEEE Trans. Biomed. Eng.* **48** 1080–7

Gorodnistky I F, George J S and Rao B D 1995 Neuromagnetic imaging with FOCUSS: a recursive weighted minimum norm algorithm *Electroencephalogr. Clin. Neurophysiol.* **95** 231–51

Hämäläinen M S *et al* 1993 Magnetoencephalography. Theory, instrumentation and applications to the noninvasive study of human brain function *Rev. Mod. Phys.* **65** 413–97

Hämäläinen M S and Ilmoniemi R J 1984 Interpreting measured magnetic fields of the brain: estimates of current distributions *Technical Report TKK-F-A559* Helsinki University of Technology

Hämäläinen M S and Ilmoniemi R J 1994 Interpreting magnetic fields of the brain: minimum norm estimates *Med. Biol. Eng. Comput.* **32** 35–42

Jerbi K, Mosher J C, Nolte G, Baillet S, Garnero L and Leahy R M 2002 From dipoles to multipoles: parametric solutions to the inverse problem in MEG *Proc. Conf. on Biomagnetism-BIOMAG 2002*

Kincses W E, Braun C, Kaiser S and Elbert T 1999 Modeling extended sources of event-related potentials using anatomical and physiological constraints *Hum. Brain Mapp.* **8** 182–93

Mosher J C, Baillet S and Leahy R M 1999a EEG source localization and imaging using multiple signal classification approaches *J. Clin. Neurophysiol.* **16** 225–38

Mosher J C, Leahy R M and Lewis P S 1999b EEG and MEG: forward solutions for inverse methods *IEEE Trans. Biomed. Eng.* **46** 245–59

Pascual-Marqui R M, Michel C and Lehman D 1994 Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain *Int. J. Psychophysiol.* **18** 49–65

Sarvas J 1987 Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem *Phys. Med. Biol.* **32** 11–22

Wagner M, Fuchs M and Kastner Jorn 2002 Current density reconstructions and deviation scans using extended sources *Proc. Conf. Biomagnetism-BIOMAG 2002*