Evaluation of different cortical source localization methods using simulated and experimental EEG data

Jun Yaoa and Julius P.A. Dewaldb,c,d,*

aDepartment of Physical Therapy and Human Movement Sciences, Northwestern University, Chicago, IL 60611, USA
bDepartment of Physical Med and Rehab., Northwestern University, Chicago, IL 60611, USA
cDepartment of Biomedical Engineering, Northwestern University, Evanston, IL 60201, USA
dSensory Motor Performance Program, Rehabilitation Institute of Chicago, Chicago, IL 60611, USA

Received 8 February 2004; revised 23 July 2004; accepted 29 November 2004

Different cortical source localization methods have been developed to directly link the scalp potentials with the cortical activities. Up to now, these methods are the only possible solution to noninvasively investigate cortical activities with both high spatial and time resolutions. However, the application of these methods is hindered by the fact that they have not been rigorously evaluated nor compared. In this paper, the performances of several source localization methods (moving dipoles, minimum Lp norm, and low resolution tomography (LRT) with Lp norm, \(p\) equal to 1, 1.5, and 2) were evaluated by using simulated scalp EEG data, scalp somatosensory evoked potentials (SEPs), and upper limb motor-related potentials (MRPs) obtained on human subjects (all with 163 scalp electrodes). By using simulated EEG data, we first evaluated the source localization ability of the above methods quantitatively. Subsequently, the performance of the various methods was evaluated qualitatively by using experimental SEPs and MRPs. Our results show that the overall LRT Lp norm method with \(p\) equal to 1 has a better source localization ability than any of the other investigated methods and provides physiologically meaningful reconstruction results. Our evaluation results provide useful information for choosing cortical source localization approaches for future EEG/MEG studies.

© 2004 Elsevier Inc. All rights reserved.

Keywords: Brain imaging; Source localization; Dipole fit; Current density reconstruction; Somatosensory evoked potentials; Event-related potentials

Introduction

Brain activity has both spatial and temporal characteristics. Conventional scalp EEG recordings are noninvasive global measurements of cortical activity with a millisecond temporal resolution. However, the spatial resolution of EEG recordings is limited, which is mainly caused by the blurring effect of the head volume conductor. In addition, the effects of electrode density, i.e., spatial sampling, signal-to-noise ratio (SNR) of the recordings, and the reference electrode placement, all contribute to the level of spatial resolution obtained by inverse calculations with EEG (Nunez, 1981). To improve the spatial resolution of scalp EEG, tremendous efforts have been made to use cortical source localization techniques to de-blur the scalp EEG recordings by deconvolving the measured scalp EEG distribution into an electrical activity (current or potential) over the cortical surface. These techniques are expected to improve the spatial resolution from centimeter scale on the scalp to millimeter scale on the cortex, and thus make it possible to show cortical spatiotemporal activities with a significant improvement in spatial resolution in a noninvasive way.

In order to de-convolve the scalp EEG, a head volume conductor model and a source model are required. As for the head volume conductor model, a homogeneous sphere head volume conductor model, a boundary element method (BEM) model, or a finite element method (FEM) model can be used. Since the BEM model is a compromise between oversimplifying sphere head volume model and the large computational costs of a FEM model, the subject specific BEM model is currently most widely used (Fuchs et al., 1998). As for the source model, two commonly used models are equivalent current dipole models (Scherg, 1990; Scherg and Buchner, 1993; Williamson and Kaufman, 1981) and a finite element method (FEM) model can be used. Since the BEM model is a compromise between oversimplifying sphere head volume model and the large computational costs of a FEM model, the subject specific BEM model is currently most widely used (Fuchs et al., 1998). As for the source model, two commonly used models are equivalent current dipole models (Scherg, 1990; Scherg and Buchner, 1993; Williamson and Kaufman, 1981) and current distributed source models (Hamalainen and Ilmoniemi, 1994; Ilmoniemi, 1993; Nenonen et al., 1994; Pascual-Marqui and Biscay-Lirio, 1993; Pascual-Marqui et al., 1994; Wagner et al., 2000; Wang et al., 1992, 1993).

When using any source models, regularized solutions are commonly used to solve the ill-posed inverse problem. With the combination of different head models, source models, and associated regulation methods, there are currently dozens of different cortical source localization methods available. The
advantages and disadvantages of these methods were analyzed in principle and described qualitatively by Wagner et al. (2000).

Comparisons between the ability of methods to detect source location(s) have been reported in the literature using experimental EEG data (Cincotti et al., 2004; Haan et al., 2000; Komssi et al., 2004; Pascual-Marqui et al., 1994; Phillips et al., 2002; Stenbacka et al., 2002; Waberski et al., 2000) and simulated EEG data (Babiloni et al., 2000b, 2004; Baillet and Garnero, 1997; He et al., 2002a) with different number of EEG electrodes (ranging from 19 to 128), head models, and regulation parameters involved. All these previous studies incorporated comparisons of two or at most three methods with fixed regulation settings. The purpose of this paper is to expand comparisons to a large number of source reconstruction methods, including dipole methods (single and dual moving dipole), minimum Lp norm methods (p equals to 1, 1.5, and 2), and low resolution electromagnetic brain tomography (LRT) methods with various Lp norms (p equals to 1, 1.5, and 2).

In an effort to estimate the accuracy of these source localization methods, we first created simulated EEG data with the obvious advantage that the exact source locations/regions were generated by the investigator. The shortcomings of using simulated EEG data are (1) the data are not realistic and (2) since the forward and backward calculations are based on the same head model, the results obtained based on simulated EEG cannot reflect errors introduced by the head model but only show the accuracy of source model and associated regulation methods. In addition to using the simulated EEG data, we also qualitatively tested the performance of the various source localization methods using experimental EEG data, including somatosensory evoked potentials (SEPs) from electrical stimulation of the wrist, elbow, and shoulder, and event-related potentials (ERPs) resulting from the generation of isometric shoulder abduction torques. The advantage of using SEPs and ERPs is that the location of sensory-motor cortical activity is well described in the literature ((Babiloni et al., 2000a; Druschky et al., 2003; Emerson and Pedley, 1984; Hari and Forss, 1999; Hashimoto, 2000; He et al., 2002b; Recuero, 1999; Sonoo, 2000; Towle et al., 2003; Valeriani et al., 2000) for SEPs and (Penfield and Rasmussen, 1950; Rao et al., 1995) for ERPs), which permitted us to determine whether solutions obtained with the various source localization methods were reasonable or not.

In the following parts of this paper, a BEM realistic head model was used for both EEG data simulation and inverse calculation. All the computations were performed using the Curry V4.5 (Neuroscan, TX) software package, which is a widely used commercial source localization software package. Other source localization methods, such as cortical potential imaging method (He et al., 2002b) and the three-dimensional resultant vector method (Ricamato et al., 2003), are not included in this study. Parts of this work have been published as a conference preceding (Yao and Dewald, 2003).

Materials and methods

Simulation EEG data

By placing equal strength (1 μAmm) dipoles on the cortex, we simulated three EEG data sets, i.e., (I) single dipole area with a fixed center and an increasing radius ranging from 5 mm to 10 mm, (II) two equal-size areas of dipoles with fixed centers and increasing radii ranging from 5 mm to 10 mm, and (III) two dipoles with increasing distance between them ranging from 4 mm to 20 mm. (For the convenience of representation, in the later parts of this paper, we call these three simulated EEG data sets as ‘simulated data sets I/II/III’.) A single dipole area (simulated data set I) was used to simulate EEG data generated from a concentrated source. Two equal-size areas of dipoles with increasing area size and two dipoles with increasing distance (simulated data sets II and III) were used to simulate EEG data generated from multiple sources.

Each EEG data set is comprised of a total of 163 EEG signals, which are the forward calculation results obtained at the electrodes on the scalp with a reference at Oz. The forward calculation was performed using a realistic BEM model, which consists of three shells: skin, skull, and liquor. The BEM model is developed based on the MRI data (with a resolution of 1.0 × 1.0 × 1.0 mm) of a normal subject. The parameters of BEM model are listed in Table 1.

Since real EEG recordings are usually noisy, we further added noise to the simulated EEG data. Noise was obtained from a real EEG measurement in which the subject was not performing any task and with his eyes open and looking at a fixed point. One hundred epochs were averaged to achieve the noise levels comparable to the typical event-related potentials (ERPs). This corresponds to a SNR of approximately 9–10. Subsequently, the data with noise were low-pass filtered by a 9th order zero-lag filter with the cut-off frequency of 50 Hz.

Experimental EEG data

Each of the inverse methods was also applied to two types of experimental EEG data, i.e., somatosensory evoked potentials (SEPs) and event-related potentials (ERPs).

Recording and processing of somatosensory evoked potentials

SEPs resulting from electrical stimulation of right wrist, elbow, or shoulder of one subject were recorded by a BioSemi Active II system (BioSemi, Amsterdam, Netherlands) with 163 active electrodes and a reference on the right ear lobe. Constant current square-wave pulses of 0.3 ms duration were delivered to the median nerve at the wrist, or lateral epicondyle elbow, or the acromion of the shoulder through disposable neurology electrodes (Ambu USA, Linthicum, MD) with 1 cm between anode and cathode using a Compex2 system (Compex, Annecy-le-Vieux, France). An interstimulus interval of 500 ms was used for a total of 1000 stimuli for each stimulus location. The intensity of the pulses was set to the highest amplitude that the subject could tolerate without discomfort. The SEP data were sampled at 2 kHz. The positions of each EEG electrode and three fiducial landmarks (Nasion, left preauricular point and right preauricular point) were marked by a Polhemus (Polhemus, Colchester, VT) three-dimensional magnetic digitizer. The three fiducial landmarks were used to co-register the 163 EEG electrodes with the subject’s MRI data. The co-registration was repeated by adjusting the fiducial landmarks in MRI data until all the 163 electrodes were co-registered well with the subject’s scalp.

Table 1

<table>
<thead>
<tr>
<th>Shell</th>
<th>Surface triangle no.</th>
<th>Side length (mm)</th>
<th>Conductivity (S/m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skin</td>
<td>2692</td>
<td>10.0</td>
<td>0.25</td>
</tr>
<tr>
<td>Skull</td>
<td>2418</td>
<td>9.0</td>
<td>0.017</td>
</tr>
<tr>
<td>Liquor</td>
<td>3306</td>
<td>7.0</td>
<td>1.79</td>
</tr>
</tbody>
</table>
The SEPs were read into Brain Vision Analyzer V1.04 software (Brain Products GmbH, München, Germany), filtered by a 1–500 Hz (48 dB/oct) band-pass filter with a notch at 60 Hz. Then the intervals that contained eye blink signals were eliminated from further analysis. The data over 40 ms pre- and 60 ms post-stimulus period were baseline corrected and averaged.

**Recording and processing of event-related potentials**

Electroencephalographic potentials related to isometric shoulder abduction torque generated by one subject were recorded by the BioSemi Active II system (BioSemi, Amsterdam, Netherlands). During the experiment, the subject was casted at the wrist and secured to a six-degree-of-freedom (DOF) load cell with the shoulder at a 75° abduction angle and a 40° flexion angle. The tip of the hand was located at a distance from the body corresponding to an elbow angle of 90° with 0° representing full extension of the elbow. In order to minimize the activation of trunk muscles, subjects were seated in a Biodex chair with the trunk secured and shoulders strapped to the back of the chair. A monitor was placed in front of the subject to provide visual feedback of torque generation in the SABD direction.

The forces and torques generated at the wrist were measured using a six-degree-of-freedom load cell (JR³ Inc., Woodland, CA) and converted online to torques at the elbow (flexion/extension) and shoulder (flexion/extension, abduction/adduction, and external/internal rotation) based on a free body analysis of the upper limb (Beer et al., 1995). Online visual feedback of the SABD level was provided during the training section. Maximum voluntary torque (MVT) in shoulder abduction direction was first recorded for the subject. The subject was then trained to generate a self-initiated SABD torque at 25% of his MVT and maintain it for a period of 0.3 s. After the subject felt comfortable with the task, EEG data were collected. During the data collection, the subject was asked to repeat the same task 100 times in blocks of 20 trials with 10 s rest periods between individual trials and 10 min rest periods between blocks to avoid fatigue. In order to reduce brain activity at the visual cortex, no visual feedback was provided during the data collection section; instead, the subject was required to look at a fixed point without eye movement during the whole period of a trial (totally about 8 s including 5–6 s before moving and 2 s after moving). Torque generation performance was reported to the subject at the end of each trial. The EEG data were sampled at 2 kHz. Similarly, co-registration of EEG electrode positions with the subject’s scalp was performed off-line.

The event-related potentials (MRPs) were read into Brain Vision Analyzer V1.04 software. The intervals that contained eye blink signals were eliminated from further analysis. The data over 2000 ms pre- and 500 ms post-torque onset were baseline corrected and averaged.

**Source reconstruction**

A realistic BEM head model and two different source models, dipole and distributed current source models, were used to reconstruct the cortical source of both simulated and experimental EEG data. Each inverse method was applied to simulated EEG data with 10 different time points. Among these 10 points, the noise is random. Regularization methods and parameters used in this paper are listed in Fig. 1. The meaning of each of the parameters is described in the Curry reference guide (NeuroScan, 1999). For convenience of representation, in the rest of this paper, we refer to all the source localization methods with different source model, different regularization methods, or parameters as methods 1–8. The exact makeup of each of the methods is provided by the upper indices in the most right column of Fig. 1 with abbreviations listed beside the indices.

Inverse methods using the dipole source model operated on the assumption that only a discrete number of generators (usually dipoles) were active at a given time or over a time period. In this paper, we investigated single and dual moving dipole methods (MDP1 and MDP2). These methods model cortical currents with equivalent current dipoles that are described by their three-dimensional locations, fixed orientation, and variable amplitude.

---

**Fig. 1. Settings for the various source localization methods.**
(Williamson and Kaufman, 1981). The position of the dipoles was found using a least-squares search that minimizes the difference between the estimated and measured data. Dipole models require a priori knowledge of the number of sources, which is usually unknown.

Conversely, when dealing with current distributed-source reconstruction (CDR) methods, no a priori knowledge of the number of sources is required. Since the neurophysiological potentials follow Ohm’s law, using CDR methods to reconstruct the source is a linear underdetermined inverse problem. Solution to this problem uses a regulation operator. More specifically, the regulation used in minimum Lp norm methods minimizes the variance $\Delta^2$ which reads $\Delta^2 = \|Lx - m\|_p + \lambda \|Wx\|_p$, where $x$ is the vector of source currents; $L$ is the lead field matrix; $m$ is the EEG measurements; $W$ is a diagonal location weighting matrix; $\lambda$ is the regulation parameter; and $p$ ($1 \leq p \leq 2$) is the measure in Banach space. When $W$ is equal to identity matrix (as used in this paper), it is referred to as a regular minimum Lp norm method; otherwise, it is a weighted minimum Lp norm method. By changing the weighting matrix $W$ to $BW$, where $B$ is the Laplacian coupling matrix, it becomes a LORETA (LRT) method. The effect of matrix $B$ is to constrain the neighboring sources to similar strengths and thus smooth the source-current distributions. In this paper, minimum Lp norm and LRT Lp norm methods with $p$ equal to 1, 1.5, and 2 were investigated. Furthermore, solutions of CDR methods investigated in this paper were constrained on a cortex overlay with 5 mm between each point and the directions of sources were rotated.

**Evaluation methods**

**Evaluation of the results on simulated EEG data**

The accuracy of the inverse solution using the simulated EEG data was evaluated by three indices: (1) the error distance (ED), which is the distance between the locations of real and estimated sources, (2) the percentage of undetected source number (PUS), and (3) the percentage of falsely-detected source number (PFS).

The error distance (ED) between the real and the estimated source locations were defined as

$$1 \over N_j \sum_{i,j} \min\{\|\vec{d}_i - \vec{s}_j\|\} + 1 \over N_f \sum_{i,j} \min\{\|\vec{s}_j - \vec{d}_i\|\}.$$  

(1)

By using this definition, one can measure the averaged error between the inverse calculation and the actual source location as a function of distance. In this equation, $s_i$ is the real dipole location of the simulated EEG data; $d_j$ is the source location detected by each inverse calculation method; and $i$ and $j$ are the indices of locations of estimated and the actual sources; and $N_j$ and $N_f$ are the total numbers of estimated and the undetected sources, respectively. In the case that CDR methods were used, a source location is defined as the location where the current strength is larger than the threshold. This threshold is set as the value corresponding to 98% confident level of a Weibull distribution function (Weibull, 1939) that fits well to the estimated cortical currents. The first term of Eq. (1) calculates the mean of the distance from each estimated source to its closest real source; and the corresponding real source is then marked as a detected source. All the undetected real sources made up the elements of data set $J$. And the second term of Eq. (1) calculates the mean of the distance from each of the undetected sources to the closest estimated sources.

The percentages of undetected source number (PUS) and of falsely-detected source number (PFS) were defined as

$$\text{PUS} = \frac{N_{\text{un}}}{N_{\text{real}}}, \quad \text{and PFS} = \frac{N_{\text{false}}}{N_{\text{estimated}}},$$

(2)

where $N_{\text{un}}$, $N_{\text{false}}$, $N_{\text{real}}$, and $N_{\text{estimated}}$ are the numbers of undetected, falsely-detected, real, and estimated sources, respectively. When calculating PUS and PFS, an undetected source is defined as a real source whose location to its closest estimated source is larger than 0.6 times the unit distance of the source layer. (Since we constrained all the solutions on a grid of cortex with 5 mm between each point, the unit distance of the source layer is equal to 5 mm.) Similarly, a falsely-detected source is defined as an estimated source whose location to its closest real source is larger than 0.6 times the unit distance of the source layer.

**Evaluation of the results on real EEG data**

Since the exact locations of SEPs and ERPs are unknown, it is hard to quantitatively evaluate the results obtained by different inverse methods. Qualitative evaluation based on current physiologic knowledge of expected active-region was therefore performed.

EEG source localization associated with SEP data is relatively clean and has been well documented in the literature (Babiloni et al., 2000a; Emerson and Pedley, 1984; Hashimoto, 2000; Recuero, 1999; Sonoo, 2000; Valeriani et al., 2000). The most widely researched and clinically applied SEPs are elicited by stimulation of the median nerve at the wrist (Emerson and Pedley, 1984; Sonoo, 2000; Valeriani et al., 2000). More recently, SEP data resulting from the stimulation of fingers or other sites have also been reported (Downman and Schell, 1999; Druschky et al., 2003; Restuccia et al., 2002; Schaefer et al., 2002; Waberski et al., 1999). Different imaging modalities, such as fMRI, PET, and source localization methods based on EEG/MEG, showed that Brodmann areas 2 and 3a, which receive deep joint and muscle receptor inputs, are believed to be the generator sites for the early components of proprioreception-related evoked responses, in contrast to area 3b activation following cutaneous mechanical and electrical stimulation, which with clearly somatotopical representations of different body parts (Druschky et al., 2003; Hari and Forss, 1999). Direct recording of cortical potentials in the human further confirmed the above findings (Towle et al., 2003).

Event-related cortical activity has mostly been studied using cortical stimulation or extracellular recordings. Early stimulation experiments in monkeys and chimpanzees (Ferrier, 1875; Hines, 1940; Sherrington, 1947; Woolsey, 1958) demonstrated that stimulation of the primary motor cortex (M1) resulted in movement in the contralateral side of the body. Penfield and Rasmussen (1950) performed cortical stimulation experiments in humans and reported results similar to those found in animal work. Furthermore, intracortical microstimulation results in the monkey (Asanuma and Rosen, 1972; Schieber, 2001) and from other studies, including single neuron recordings in monkeys (Schieber and Hibbard, 1993), fMRI recordings in humans (Sanes et al., 1995), and MEGs in humans (Cheyne et al., 1991; Salenius et al., 1997), all suggested that M1 is the source generator of movement-related potentials.

The previously established knowledge of the source for SEPs and ERPs was used to evaluate the performance of different source localization methods qualitatively.
Statistical analysis method

The three evaluation indices (i.e., ED, PUS, and PFS) were subjected to separate multivariate analysis of variance (MANOVA). The two dependent variables of the MANOVA were the mean and standard deviation of each tested index, and the three factors were the inverse method, the type of simulated EEG data set, and the radius of the dipole area (for data set I/II) or the distance between the two dipoles (for data set III). Up to 2-way interactions between the factors were considered when performing the analysis. The post hoc analysis was implemented using the Bonferroni test at the \( P = 0.05 \) level of significance.

Results

Source localization results obtained on simulated EEG data

Example results obtained using each of the methods on simulated data sets I/II/III (with a radius that equals 8 mm for data sets I and II, and with a distance between the two dipoles that equals 18 mm for data set III) are shown in Figs. 2A–C, respectively. Results are shown in a brain coordinate system with \( x- \) and \( y- \) axis representing anterior (negative \( x \))/posterior (positive \( x \)) and left (positive \( y \))/right (negative \( y \)) in mm. In this figure, the first and second rows showed the results of minimum \( L_p \) norm (\( P = 1, 1.5, 2 \)) and LRT \( L_p \) norm (\( P = 1, 1.5, 2 \)), respectively. Color bars on the right side of each subplot show the averaged current strength over the 10 time points on the cortex in \( \mu A/mm \). On the last row, the purple dots on the first two subplots are the mean of the estimated dipole location over the 10 time points obtained by using MDP1 and MDP2 methods. The horizontal, vertical, diagonal lines represent the standard deviation of estimated dipole position on \( x, y, \) and \( z \) directions. Finally, the purple dots in the last subplot represent the real source dipole positions. Quantitative evaluation results and statistical analysis results were performed and will be discussed next.

Location accuracy

The location error distances (EDs) of each inverse method using simulated data sets I/II/III are shown in Fig. 3. In this figure, the gray bars on each of the data points represent the standard error of corresponding results. Results clearly show that the LRT1 method provides the best estimation of source location for all three simulated EEG data sets. When applying CDR methods on simulated data sets I and II, the EDs of LRT1 have general decreasing trends when the sizes of dipole-areas increase. However, such a decreasing trend is not so obvious for the other methods on both data sets I and II (see Figs. 3a and b). Results of EDs obtained by applying LRT with \( L_p \) (\( P = 1, 1.5, 2 \)) norm methods on simulated data set III fluctuate when the distance between the two source dipoles is about 12–16 mm (see Fig. 3c). The reason for this fluctuation is believed to be related to constraints imposed on source location by the 5 mm grid on the cortex (see Materials and methods).

Although the moving dipole methods (MDP1 and MDP2) show relatively low EDs, there is an obvious posterior and medial shift shown on the results of all the three simulated EEG data sets, while such a shift is not found in the results obtained by using CRD methods (see Fig. 2). Even worse, locations estimated by the MDP2 method over 10 time points show a large variance (see Fig. 3). It is worth to note that when applying MDP2 to simulated data set III, there is a decreasing trend in the location error when the distance between the two real dipoles increases. This trend is not obvious for any of the other methods. This result suggested that the a priori knowledge of the number of dipoles could improve the inverse results when using dipole source models. Unfortunately, even when the correct number of dipoles is implemented, the MDP2 method still produces posterior and medial shifts, which result in substantial location errors.

The results of statistical analysis using a three-way MANOVA and Bonferroni post hoc tests on the mean and standard deviation of location errors are shown in Table 2a. In this table, values in the upper and lower triangle are \( P \) values on the mean and standard deviation of different indices, respectively. Results in Table 2a show that location estimated by LRT1 method is significantly more accurate than any other methods tested in this paper (see statistical results on ED in the top triangle of Table 2a). Furthermore, the variability (estimated by standard deviation) of the LRT1 method on estimating the location of source is either significantly less than or similar to other methods (see statistical results on the standard deviations in the bottom triangle of Table 2a). The location error obtained by using LRT15 is significantly smaller than results obtained by using any other methods except LRT1 and MDP1.

The percentage of undetected source number

The percentages of undetected source number (PUS) of each source localization method using simulated data sets I/II/III are shown in Fig. 4. Since dipole methods are not able to detect the area of the source, the PUSs for dipole methods are close to 100%. When using simulated data sets I and II, PUSs of all CDR methods have a tendency to increase when the size of dipole area increases. This tendency is slower for the LRT1 method than that for any of the other tested CDR methods. It is worth noting that when using simulated data set II to evaluate the CDR methods, there is an obvious drop of PUS when the radiiuses of the two source dipole-areas increase from 5 mm to 6 mm. When referring to results obtained with simulated data set II with the radiiuses of the two dipole areas equal to 5 mm in Fig. 2d, one finds that all the CDR methods could not detect the laterally located dipole area; instead, they detected a single enlarged and merged dipole area located closer to the medial dipole area (see Fig. 2d). When increasing the radiiuses from 5 mm to 6 mm, all CDR methods show an unseparated bimodal area, which is related to the drop of PUS observed in Fig. 4b. This result shows that a 5 mm source grid does not allow for the detection of two separated sources each with a 5 mm radius and with 29 mm between their centers.

The simulated data set III consists of two dipoles. For this type of EEG data, all CDR methods are able to detect the two dipoles when the distance between the 2 source dipoles is less than 12 mm. However, most of CDR methods except LRT1 are unable to detect one of the two dipoles when the distance between the 2 source dipoles is larger than 12 mm. In all the circumstances, LRT1 has the lowest PUS from all CDR methods.

The results of statistical analysis using three-way MANOVA and Bonferroni post hoc tests on the mean and standard deviation of the percentages of undetected source number (PUSs) are shown in Table 2b. Results show that LRT1 has a significantly smaller and less variable PUS than all the other methods tested in this paper (except that the mean and standard deviation on PUSs for LRT1 and LRT15 are similar).
The percentage of falsely-detected source number

The percentages of falsely-detected source number (PFSs) of each source localization method using simulated data sets I/II/III are shown in Fig. 5. Results show that when using simulated data sets I and II, PFSs of LRT1 and LRT15 methods have a tendency to decrease when the size of dipole area increases, while this trend is
not apparent for the other CDR methods. PFSs of MDP1 for simulated data sets I/II/III are always equal to 100%, which shows that this method is not able to detect the source location within the source area (caused by the posterior and medial shift). The PFSs of MDP2 decrease with the increase of the radius of source area using simulated data set I, which clearly suggests that adding a second dipole is beneficial in reducing the false detection ratio. However, MDP2 has a significant larger standard deviation compared to other methods (see Table 2c).

The results of statistical analysis using three-way MANOVA and Bonferroni post hoc tests on the mean and standard deviation of the percentages of PFS are shown in Table 2c. Results showed that LRT1 has a significantly smaller PFS than any of the other methods.
Source localization results on real EEG data

The results of different inverse methods on SEPs resulting from electrical stimulation of the subject’s wrist during the time window from 20 to 25 ms after the stimulus are represented in Fig. 6. As shown in Fig. 6, the majority of activities obtained by all the methods stay on the contralateral hemisphere. Inverse results on SEPs from shoulder, elbow, and wrist shift from medial to lateral along the central sulcus. (SEP results of simulating elbow and shoulder are not represented here.) Qualitatively, one can observe that the inverse results obtained by using method MDP1 or MDP2 have a large variance even during a short time window (5 ms). And the mean locations obtained by MDP1 and MDP2 are located at pre-central gyrus (i.e., motor cortex). These results are not consistent with results published in the previous literature, which suggests that the sources of SEPs are located on the primary sensory cortex (S1 area). Results obtained by using CDR methods have a similar ‘hot spot’ located on S1 area, while multiple other active sites besides S1 area also show up. Among all the CDR methods, results of methods L1, L15, and LRT1 are more focused on S1 area; furthermore, methods L15 and LRT1 show relatively stronger activity on S1 area.

Results of ERPs related to self-initiated isometric shoulder abduction during a time window from 110 to 90 ms before the onset of torque are shown in Fig. 7. The majority of activities prior to shoulder abduction obtained by all the CDR methods are located on the contralateral hemisphere. Results by MDP1 and MDP2 are located either at prefrontal lobe or parietal lobe. Activities on prefrontal and parietal lobes are also found on the results obtained by CDR methods. However, one observes that the result obtained by LRT1 is the most focused on the medial pre-central gyrus with very limited activity anywhere else. The L1 CDR method performed only slightly worse than LRT1. All other CDR methods show activity in many regions other than M1. The results obtained with LRT1 are consistent with those reported in the previous literature.

Table 2
Results of Bonferroni post hoc tests on the mean and standard deviation of the ED, PUS, and PFS indices

<table>
<thead>
<tr>
<th>Method</th>
<th>ED</th>
<th>PUS</th>
<th>PFS</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>+++</td>
<td>+</td>
<td>++</td>
</tr>
<tr>
<td>L15</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>L2</td>
<td>+++</td>
<td>+</td>
<td>+++</td>
</tr>
<tr>
<td>LRT1</td>
<td>++</td>
<td>+++</td>
<td>+++</td>
</tr>
<tr>
<td>LRT15</td>
<td>+++</td>
<td>---</td>
<td>+++</td>
</tr>
<tr>
<td>LRT2</td>
<td>+++</td>
<td>---</td>
<td>+++</td>
</tr>
<tr>
<td>MDP1</td>
<td>+++</td>
<td>---</td>
<td>+++</td>
</tr>
<tr>
<td>MDP2</td>
<td>+++</td>
<td>---</td>
<td>+++</td>
</tr>
</tbody>
</table>

In this table, values in the upper (white cells) and lower triangle (gray cells) are P values of the mean and standard deviation, respectively. The table should be read from row to column, e.g., the element in first row and third column of (c) shows that the mean of PFS for the L1 method is significantly larger (+) than the mean for the L2 method, where +/−P < 0.05, ++/−−P < 0.01, +++/−−−P < 0.001, ++++/−−−−P < 0.0001.
literature and demonstrated that M1 is the primary source of activity related to arm activity.

In general, inverse results obtained by applying LRT1 to SEPs and ERPs are qualitatively better than other methods, because it provides results that mirror those obtained with direct cortical recordings or with other noninvasive neuroimaging modalities.

Discussions and Conclusions

In order to evaluate different cortical source localization methods, we applied 8 different inverse methods available in CURRY software to both simulated and experimental EEG data. Although the evaluation on the source localization ability of two or three different methods was already reported previously, this is the first paper that compared these many inverse methods using both simulated and experimental data sets. Our evaluation results provide useful information for choosing the appropriate cortical source localization techniques for scientific and clinical purposes.

Evaluation methods

In order to evaluate the localization ability of the different source localization methods, we designed three types of simulated EEG data sets. Results obtained with each of the simulated EEG data sets represented various facets of the accuracy of the inverse methods. Simulated data sets I and II allowed for the

Fig. 4. The percentage of undetected source number for the three simulated EEG data sets using the eight inverse methods. The gray error bars on each of the data points represent the standard error.

Fig. 5. The percentage of false-detected source number for the three simulated EEG data sets using the eight inverse methods. The gray bars on each of the data points represent the standard error.
evaluation of source localization ability; and data sets II and III allowed for the evaluation of the spatial resolution of the various techniques.

Besides simulated EEG data, we also evaluated the different inverse methods using experimental EEG data. Two typical experimental EEG data sets (i.e., SEPs and ERPs) were used to qualitatively evaluate results obtained with each of the inverse methods. Our assessments are based on our current knowledge of regions of cortical sensory motor activity associated with sensory evoked or motor-related potentials.

Fig. 6. Inverse calculation results for SEPs resulting from electrical stimulation of the medial nerve at wrist in a subject. The figure legends are the same as those in Fig. 2. The green line in each of the subplots is the central sulcus.

Fig. 7. Inverse calculation results for ERPs related to isometric shoulder abduction torque generated by a subject. The figure legends are the same as those in Fig. 2. The green line in each of the subplots is the central sulcus.
In case of applying inverse methods to simulated EEG data, where the real source locations are known, three different indices, i.e., the error distance (ED), percentage of undetected source number (PUS), and percentage of falsely-detected source number (PFS), were used to quantify the performance of the different inverse methods. A method with lower ED, PUS, and PFS is able to detect the source location more accurately. Based on the definition of ED, there is a positive correlation between ED and PUS/PFS, i.e., a larger PUS or PFS will cause an increase in ED. The correlation coefficients between ED and PUS/PFS are shown in Table 3. Because of the highly significant positive correlation between ED and PUS/PFS, one can only refer to ED for the evaluation of the source localization ability of an inverse method.

Factors that may influence our evaluation methods

Several factors that may influence the performance of inverse methods were not studied in this paper. First of all, previous reports showed that results of inverse calculation are sensitive to the conductivity of head (Benar and Gotman, 2002; Gencer and Acar, 2004; van Burik and Peters, 2000; van den Broek et al., 1998; Wen and Li, 2001). Among conductivities of brain, skull, and scalp, the inverse result is most sensitive to the conductivity of skull (Ferree et al., 2000). However, the real conductivity of skull in vivo is still unknown. Direct measurements showed a large variance on this value ranging from 15 mS/m to 80 mS/m (Akhtar et al., 2002; Ferree et al., 2000; Hoekema et al., 2003). The conductivity values used in this paper were listed in Table 1 and corresponded to a skull-scapal-ratio equal to 1:15. Inverse results were also obtained using skull-scapal-ratios equal to 1:26 and 1:80. In all of these cases, LRT1 provides the overall best results.

Secondly, the regulation parameter, $\lambda$, is crucial to the performance of inverse methods. As a general rule, the degree of regularization ($\lambda$) should increase with the level of noise in the data. Based on this rule, $\lambda$ is often determined using $\chi^2$ and $L$-curve methods. Unfortunately, these two methods are not available for the minimum Lp norm and LRT Lp norm methods in CURRY. In this paper, a fixed value ($\lambda = 1$) was used for all the methods. In order to limit the potential influence caused by using a non-optimized regulation parameter, we controlled the signal to noise ratio (SNR) of all the EEG data, including both the simulated and experimental data, to a very consistent level, ranging from 9 to 10. Regardless, evaluation of the different source localization methods by using an optimized regulation parameter remains desirable.

The shape of simulated EEG data may be another factor that will affect the evaluation results. All the three simulated EEG data sets are composed of equal strength dipoles (1 $\mu$Amm). No decay of dipole strength was included. The square-wave-shaped sources used here are sub-optimal for LRT Lp norm methods, which are designed on the assumption that the source activities are distributed smoothly on the source layer. We postulate that using simulated EEG data with decays will further improve the performance of LRT Lp norm methods.

When evaluating the CDR methods, a threshold, which was set at a 98% confidence level, was used to define a source. It is obvious that the choice of threshold also affects our evaluation results. Since current estimates by different CDR methods have different distributions, we choose a data-specific threshold for each of the inverse results (i.e., we first fitted the estimated cortical currents using a Weibull distribution; then we set the threshold at the 98% confidence level). Even with such a threshold, the residuals of the fitting procedure are slightly different depending on results obtained by different localization methods. Therefore, there possibly is a small bias for ED, PUS, and PFS depending on the choice of source localization method.

Using a data-specific threshold at 98% confidence level, all CDR methods tend to provide a relatively large PFS; however, different methods have a different distribution of the falsely-detected sources. For example, when using LRT1 method, the falsely-detected sources are distributed around the real source; while when using other CDR methods, more falsely-detected sources are distributed around the edge of the cortex (see Fig. 8, where the purple crosses represent the detected source positions). By increasing the threshold, we can effectively increase the PUS and decrease the PFS of CDR methods. Since a positive correlation existed both between ED and PUS as well as between ED and PFS, the impact of an increased threshold on ED is not straightforward.

Other factors, such as the number of electrodes, the depth of the source, and the SNR of signals, also have impact on the performance of the various inverse methods. The impact of these parameters were not included as part of the current investigation.

Accuracy of the different source localization methods

According to the results obtained using 3 different simulated EEG data sets, the LRT L1 norm method showed the overall highest accuracy. We further summarized the comparison on the mean and the standard deviation (inside the parentheses) of the three evaluation indices between LRT1 to other inverse methods in Table 4. (Note, only significant results were represented.) As shown in Table 4, the LRT1 method has either significant better or similar performance when compared to the other investigated methods for all of the three indices, which showed an overall superior ability to detect the location of the source. When applying the various methods to experimental EEG data, including both SEPs and ERPs, the LRT1 method also provides the best estimates. These findings further confirm that among all the tested methods, LRT1 has the most accurate and least variable source localization abilities.

Our results further indicate that a weighted minimum norm method, such as LRT-$L_p$, has better source localization ability than the regular minimum norm methods. These results are in accordance with those reported in previous literature (Baillet and Garnero, 1997; BabILONI et al., 2000b, 2004). Furthermore, we also demonstrate that CDR methods provide physiologically meaningful results while MDP1 and MDP2 solutions failed in many situations, which is also in agreement with previous reported results (He et al., 2002a; Pascual-Marqui et al., 1994). Within the same CDR...
The regulation method (minimum norm or weighted minimum norm) a lower Lp value will improve the localization results.

**Spatial resolution of different methods**

The purpose of applying inverse methods is to improve the spatial resolution of EEG measurement. However, to date, no quantitative examination has been performed to provide the spatial resolution of an inverse method. For this reason, we designed simulated EEG data sets II and III to quantify the spatial resolution of different inverse methods for 163-channel EEG signals with a SNR around 9–10. Simulated EEG data set II has two dipole-areas with increasing radiuses (ranging from 5 mm to 10 mm) and a fixed distance (29 mm) between their centers. When applying an inverse method to simulated EEG data II, the spatial resolution of the method is defined as the minimum radius of dipole area that a method can distinguish between the two areas. Simulated EEG data set III consists of two moving dipoles with the distance between them increasing from 4 mm to 20 mm. We defined the spatial resolution of an inverse method applied to this type of data as the minimum distance between two moving dipoles that can be distinguished.

Results on simulated EEG data II showed that when the radiuses of the two dipole areas are equal to 5 mm, none of the CDR methods could separate the two areas (i.e., only one mode is detected). When increasing the radiuses from 5 mm to 6 mm, L15, L1, LRT15, and LRT1 methods showed an un-separated area, however, with a bimodal distribution. Furthermore, two separated areas can be found when using a higher threshold. These results suggest that the spatial resolution for these methods is around 6 mm. Further increasing the radiuses of the two dipole areas did not cause the merge of the two dipole areas. This is a very interesting result. Considering that by increasing the radiuses of the two dipole areas while keeping the distance between their centers the same, the two source regions were getting closer, which we thought would require a better spatial resolution to separate them. However, this is not correct. A source image with a smaller radius has higher frequency components. In the extreme case, the frequency component of a source image with only a single dipole is the highest. A source containing higher frequency components requires a higher spatial resolution to detect. Therefore, our results showed a better ability to separate two areas when the radiuses of the two source dipole-areas are increasing.

The effect of the frequency component of the image would also explain the results obtained using simulated EEG data set III which showed that when increasing the distance between the two dipoles, no obvious change in performance of the CDR methods was observed (P values of post hoc analysis on the effect of radius are larger than 0.05). Obviously, changing the distance between the two dipoles does not change the frequency component of source

---

**Table 4**

Comparison of the three evaluation indices between LRT1 and the other methods

<table>
<thead>
<tr>
<th>Methods</th>
<th>LRT1</th>
<th>L1</th>
<th>L15</th>
<th>L2</th>
<th>LRT15</th>
<th>LRT2</th>
<th>MDP1</th>
<th>MDP2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ED</td>
<td>PUS</td>
<td>PFS</td>
<td>ED</td>
<td>PUS</td>
<td>PFS</td>
<td>ED</td>
<td>PUS</td>
</tr>
<tr>
<td>L1</td>
<td>++++ (+++)</td>
<td>++++ (++++)</td>
<td>+++ (++)</td>
<td>L15</td>
<td>++++ (+++)</td>
<td>++++ (++++)</td>
<td>+++ (++)</td>
<td>L2</td>
</tr>
<tr>
<td>LRT15</td>
<td>++++ (++)</td>
<td>++++ (++++)</td>
<td>++++ (++++)</td>
<td>LRT15</td>
<td>++++ (++)</td>
<td>++++ (++++)</td>
<td>++++ (++++)</td>
<td>LRT2</td>
</tr>
<tr>
<td>MDP1</td>
<td>++++ (++++)</td>
<td>++++ (++++)</td>
<td>++++ (++++)</td>
<td>MDP2</td>
<td>++++ (++++)</td>
<td>++++ (++++)</td>
<td>++++ (++++)</td>
<td>Original</td>
</tr>
</tbody>
</table>

The table should be read from row (method) to column (index), e.g., the element in first row and second column shows that the mean of ED for L1 method is significantly larger than the mean of ED for LRT1 method, where +/ P < 0.05, ++/ P < 0.01, +++/ P < 0.001, ++++/ P < 0.0001. The results with and without parentheses represent the P value of the standard deviation and mean, respectively.
image enough to improve the performance of the CDR source localization methods.

Another aspect that may have influenced our results is the use of the 5 mm grid of source layer to constrain the source locations. We believe that by using a finer grid the spatial resolution for L15, L1, LRT15, and LRT1 methods would further improve the performance of these CDR methods.

Practical implications for the application of the various source localization methods

Dipole source models are more widely used in the previous literature to estimate centers of cortical activity. We believe that they are only suitable for the situation where a single source is presented. In order to model multiple sources, multiple dipole methods were developed. However, our results showed that even using the correct number of dipoles, multiple dipole methods could not determine the accurate location of all the sources.

CDR methods have the ability to determine the current distribution of brain activity. Among all the CDR methods investigated in this paper, the LRT1 method showed the highest accuracy and, therefore, is the most promising method for source localization estimation especially when different cortical regions are expected to be simultaneously active during a particular experimental paradigm.

Acknowledgments

This work was supported by a National Institutes of Health Grant 1R03HD39804-01A1 and by a donation from the RIC Women’s Board. The authors want to acknowledge the constructive suggestions made by Drs. Leo Twole, Wim van Drongelen, Arye Nehorai, and the reviews of NeuroImage during the preparation and revision of the manuscript.

References


