Title: Performance Comparison of L2-Minimum-Norm Estimators for EEG Source Localization in Visual Evoked Potentials

Authors: Mohammad Ahadzadeh[1], Amir Hosein Ardalan[2], Reyhaneh Amiri[1], Arefeh Rezagholi[3], Golnaz Baghdadi[1,*]

1. *Department of Biomedical Engineering, Amirkabir University of Technology, Tehran, Iran*
2. *Sharayan Intelligent Process and Analysis Research and Development Team, Tehran, Iran*
3. *Faculty of Computer and Electrical Engineering, Tarbiat Modares University, Tehran, Iran*

***Corresponding Author**: Golnaz Baghdadi, Department of Biomedical Engineering, Amirkabir University of Technology, Tehran, Iran. Email: Golnaz_baghdadi@aut.ac.ir

To appear in: **Basic and Clinical Neuroscience**

Received date: 2026/04/9

Revised date: 2026/06/9

Accepted date: 2026/07/28

Please cite this article as:

Ahadzadeh, M., Ardalan, A.H., Amiri, R., Rezagholi, A., Baghdadi, G. (In Press). Performance Comparison of L2-Minimum-Norm Estimators for EEG Source Localization in Visual Evoked Potentials. Basic and Clinical Neuroscience. Just Accepted publication Jul. 10, 2026. Doi: http://dx.doi.org/10.32598/bcn.2026.8673.1

DOI: http://dx.doi.org/10.32598/bcn.2026.8573.1

Abstract

Accurate localization of neural activity from electroencephalography (EEG) remains challenging due to the ill-posed inverse problem and spatial leakage between cortical regions. Although several linear inverse methods have been proposed, their practical performance under realistic experimental conditions is not fully understood. This study compares four widely used distributed inverse solutions, minimum norm estimation (MNE), dynamic statistical parametric mapping (dSPM), standardized low-resolution electromagnetic tomography (sLORETA), and exact LORETA (eLORETA), focusing on spatial resolution and leakage characteristics.

EEG data were collected from nineteen right-handed participants (mean age 23 ± 3.5 years) performing two visual tasks: face detection and motion discrimination. Recordings were obtained using a 128-channel EEG system at 2048 Hz. Spatial resolution was assessed using peak localization error (PLE) and spatial dispersion (SD) derived from point spread functions (PSF) and cross-talk functions (CTF). Spatial leakage patterns were also examined across functionally relevant cortical networks.

Statistical comparisons revealed significant differences between inverse methods for PSF based localization accuracy, with sLORETA and eLORETA achieving near zero localization error, whereas MNE and dSPM showed larger deviations. However, under CTF analyses reflecting more realistic multi-source activity, all methods exhibited increased localization errors and similar spatial spread. Network analyses further indicated that anatomical organization influences leakage patterns, with compact regions such as V1 demonstrating higher localization specificity than distributed visual networks.

Overall, the results indicate that no single inverse method optimizes all aspects of spatial resolution. Instead, each method reflects a trade-off between focality, smoothness, and robustness, highlighting the importance of selecting inverse models according to the goals of EEG source analysis.

Keywords: EEG source localization; Inverse problem; Minimum norm estimation (MNE); sLORETA / eLORETA; dynamic statistical parametric mapping (dSPM)

Highlights:

- Systematic comparison of four linear EEG inverse solutions: MNE, dSPM, sLORETA, and eLORETA.
- LORETA-based methods achieve near-zero localization error under ideal point-source (PSF) conditions.
- Performance differences between methods decrease under realistic cross-talk (CTF) conditions.
- Spatial dispersion remains similar across methods, indicating persistent spatial leakage.
- Method selection should depend on analysis goals, balancing focal localization and robustness to distributed activity.

1. Introduction

Electroencephalography (EEG) is a noninvasive neuroimaging technique in both neuroscience research and clinical diagnostics, enabling real-time monitoring of dynamic neural activity (Niedermeyer & da Silva, 2005). Despite EEG's high temporal resolution, its spatial precision is limited because the electrical potentials recorded at the scalp reflect the superimposed activity of many neuronal populations, which are filtered and distorted by intervening tissues such as the scalp, skull, cerebrospinal fluid, and brain matter (Hämäläinen et al., 1993). To gain a precise understanding of brain function, it is therefore essential to map these scalp measurements back to their underlying cortical sources—a process known as EEG source localization. Formally, this constitutes the EEG inverse problem, which contrasts with the forward problem that predicts scalp potentials from known neural sources (Baillet et al., 2002). The inverse problem is inherently ill-posed and non-unique, since a finite number of scalp measurements must represent a theoretically infinite source space. Consequently, meaningful and physiologically plausible solutions require regularization techniques and the incorporation of prior information about neural activity, enabling researchers and clinicians to infer the spatial distribution and intensity of cortical currents from noninvasive recordings. Two primary methodological frameworks address this problem. Parametric approaches model brain activity with a small number of equivalent current dipoles (ECDs), each characterized by location, orientation, and moment, and are suitable for sparse or well-separated sources (Grech et al., 2008). Imaging or distributed approaches, on the other hand, represent cortical activity as currents distributed across a finely tessellated cortical surface comprising thousands of dipoles (Dale & Sereno, 1993). This formulation yields a linear inverse problem with dipole amplitudes as unknowns and necessitates regularization to manage ill-posedness (Pascual-Marqui et al., 1994).

The most commonly employed methods for estimating distributed neural sources from EEG data fall primarily into L2-minimum-norm type methods and linearly constrained minimum variance beamformers. While the L2-norm Minimum Norm Estimation (MNE) family of inverse solutions, including standard MNE, Dynamic Statistical Parametric Mapping (dSPM), and various LORETA variants, are widely employed due to their computational efficiency and ability to model distributed sources, their performance is not uniform across all neurophysiological contexts. Specifically, Visual Evoked Potentials (VEP) signals present unique challenges, such as low signal-to-noise ratios (SNR) and rapid temporal shifts, which can exacerbate known artifacts like depth bias inherent to basic minimum-norm solutions. Therefore, this study systematically compares key L2-MNE family estimators in VEP data to provide practical guidance for researchers studying visual processing pathways.

2. Fundamentals of EEG Source Localization

2.1 Biophysical Basis

EEG signals originate primarily from the postsynaptic potentials of cortical pyramidal neurons (Nunez & Srinivasan, 2006). These neurons are organized in columns perpendicular to the cortical surface, allowing their synchronous activity to summate and generate detectable extracellular currents at the scalp (Buzsáki et al., 2012). The dendritic currents of these neurons create small electric dipoles, and when thousands of pyramidal neurons are activated in synchrony, the resulting field is strong enough to be measured by scalp electrodes (da Silva, 2013).

The measured EEG represents the superposition of many neuronal sources, filtered by the tissues of the head, including the scalp, skull, cerebrospinal fluid (CSF), and brain matter (Hämäläinen et

al., 1993). This filtering attenuates and spatially smears the signals, which is why EEG exhibits high temporal resolution but comparatively low spatial resolution (Niedermeyer & da Silva, 2005).

Two types of currents are involved in EEG generation (Hämäläinen et al., 1993; Nunez & Srinivasan, 2006):

- **Primary currents (Jp):** Intracellular currents flowing along the dendrites of activated pyramidal neurons.
- **Volume currents:** Extracellular currents conducted through brain tissue, CSF, skull, and scalp, producing the potential differences recorded at electrodes.

The physical behavior of electric potentials within the head is governed by Maxwell's equations, which describe the relationship between electric and magnetic fields in any medium. Because neural activity occurs at low frequencies (0.1–100 Hz), the quasi-static approximation can be applied, allowing the time-derivative terms to be neglected. Under this approximation, Maxwell's equations simplify to a form of Poisson's equation, which expresses that the spatial distribution of electrical potential at any location depends on the primary currents generated by neurons and the conductive properties of surrounding tissues (Jatoi & Kamel, 2017). Under the quasi-static approximation, the governing equation for the electric potential distribution in the head can be written as follows, accounting for heterogeneous tissue conductivity and neural current sources:

$$\nabla \cdot \sigma \nabla V = \nabla \cdot \mathbf{J}^i \quad (1)$$

- ∇: This is the divergence operator, which measures the net flow of a field out of a small volume. In physics, it can represent sources or sinks of a field. When applied to a vector field, like current density $\mathbf{J}^i$, it describes how the current is distributed within a volume.

- **σ**: This is the conductivity of the material, and it is a scalar (or tensor in 3D problems with anisotropic conductors) that represents how easily the material allows electrical currents to flow. Conductivity σ can vary across different tissues in the head (e.g., scalp, skull, brain, CSF) in an inhomogeneous conductor.
- **∇V**: This is the gradient of the electric potential (V), which describes how the electric potential changes with respect to position in space. The gradient points in the direction of the steepest increase of the potential, and its magnitude represents the rate of change of the potential.
- $\mathbf{J^i}$: This is the current density ($\mathbf{J^i}$), which is the flow of electric charge per unit area. In this case, it represents the "impressed currents," which are the bioelectric currents generated by neural activity.
- $\mathbf{\nabla \cdot J^i}$: The divergence of the current density describes the net current flowing out of a region. In biological tissues, it represents how much current is being generated by sources (e.g., active neurons) and how it spreads through the surrounding tissues.

Equation (1) essentially states that:

1. The divergence of the electric field (related to the gradient of the potential) is proportional to the current density in the system.
2. The term $\nabla \cdot \sigma \nabla$ represents the distribution of electric potential in the presence of a spatially varying conductivity (σ) in the volume.
3. The term $\nabla \cdot J^i$ on the right-hand side corresponds to the current sources (e.g., neural sources) and their distribution.

2.2 Forward Problem

EEG source localization involves determining the brain's internal sources responsible for the electrical activity observed on the scalp. This process starts with the "forward problem," where the goal is to predict the scalp potential from known brain sources. In other words, the forward problem simulates how electrical currents propagate through the head. Accurate forward models are fundamental for source localization, as they predict how cortical activity manifests at the scalp and provide the basis for solving the inverse problem. Accurate localization requires individual anatomical information, such as skull thickness and electrode positions, to construct the lead field, which describes the mapping from cortical sources to scalp potentials (Michel & Brunet, 2019). Solving this problem requires appropriate boundary conditions at interfaces between tissues, commonly Neumann (derivative-based) or Dirichlet (value-based) conditions. EEG signals primarily reflect the activity of pyramidal neurons, which are aligned in parallel; the small electric fields generated by these neurons are typically modeled as current dipoles. Accurate representation of tissue conductivity is critical, as errors in these parameters can introduce significant inaccuracies in EEG source analysis.

Because the head consists of multiple bounded conductors with complex geometries, modeling approaches must account for tissue-specific conductivity. Two broad categories of head models are commonly used: analytical and numerical. Analytical models approximate the head as nested concentric spheres with homogeneous conductivity. While computationally simple, their geometric oversimplification often leads to higher localization errors (Jatoi & Kamel, 2017). The advent of advanced neuroimaging and increased computational power has enabled the development of realistic head models derived from Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) scans (Cuffin, 2002). These numerical models incorporate intricate

anatomical details, including the skull's shape, cerebrospinal fluid (CSF) layers, and tissue anisotropies, providing higher resolution and lower localization errors compared to simpler methods (Jatoi & Kamel, 2017; Montes-Restrepo et al., 2014; Vatta et al., 2010). However, this fidelity comes at the cost of increased computational demands (Jatoi & Kamel, 2017). The most significant advantage of realistic models, derived from MRI/CT, is improved spatial accuracy, reducing localization errors from the typical 10-20 mm range in spherical models to 4-6 mm (Nuñez Ponasso et al., 2024; Ramon et al., 2006; Vanrumste et al., 2001), especially in anatomically complex regions like the temporal lobes (Ramon et al., 2006; Vanrumste et al., 2001). The trade-off between accuracy and computational cost impacts clinical diagnostics, where realistic models yield more reliable results for presurgical planning (Brinkmann, 2024; Fuchs et al., 2007), and BCIs (Yokoyama et al., 2024). Furthermore, while individualized realistic models are the gold standard, template-based realistic models can introduce errors up to 2 cm if the individual's anatomy significantly mismatches the template (Liu et al., 2023). Table 1 compares spherical and realistic head models in EEG source localization.

Table 1. Spherical vs. Realistic Head Models in EEG Source Localization

Feature	Spherical Head Models	Realistic (Numerical) Head Models
Spatial Accuracy (Error)	Low; prone to systematic errors of 10–20 mm	High; errors reduced to 4–6 mm
Anatomical Complexity	Simplified geometry (concentric spheres)	Incorporates detailed anatomy (skull, CSF layers, tissue anisotropy)
Performance in Complex Areas	Poor performance, particularly in temporal/basal areas	Superior performance, crucial for accurate current flow modeling
Computational Complexity	Low; fast computation and practical for real-time use	High; significant demands for model construction and forward solution (FEM/BEM)
Clinical Utility	More practical under resource constraints	Higher fidelity for critical tasks like presurgical planning
Impact of Noise	Advantage narrows as Signal-to-Noise Ratio (SNR) decreases	Accuracy advantage is most pronounced at high SNR
Demographic Adaptability	"One-size-fits-all" approach; less adaptable to individual variation	Individualized models offer best fit; template models can introduce significant error if mismatch occurs

Figure 1 illustrates the multilayer head models used in the EEG forward problem. Neuronal activity is represented by an equivalent current dipole embedded inside the brain tissue. The electric

potential generated by this source propagates through the surrounding conductive structures of the head and is measured by electrodes placed on the scalp. To accurately model this volume conduction process, the head is represented as a set of nested compartments, each with its own geometry and electrical conductivity. In the three-layer configuration (Figure 1 (b)), the head consists of brain, skull, and scalp tissues with conductivities σ_1, σ_2, and σ_3, respectively. Their interfaces S_1, S_2, and S_3 define the boundaries across which the normal component of current flux is continuous. The dipole source $\boldsymbol{J}_p$ is placed inside the brain region Ω_1 and defined by its position and orientation. The four-layer model (Figure 1 (c)) follows the same structure but includes an additional compartment for the cerebrospinal fluid (CSF). Because CSF has a much higher conductivity than the skull, incorporating this layer significantly improves the realism of the electrical field distribution.

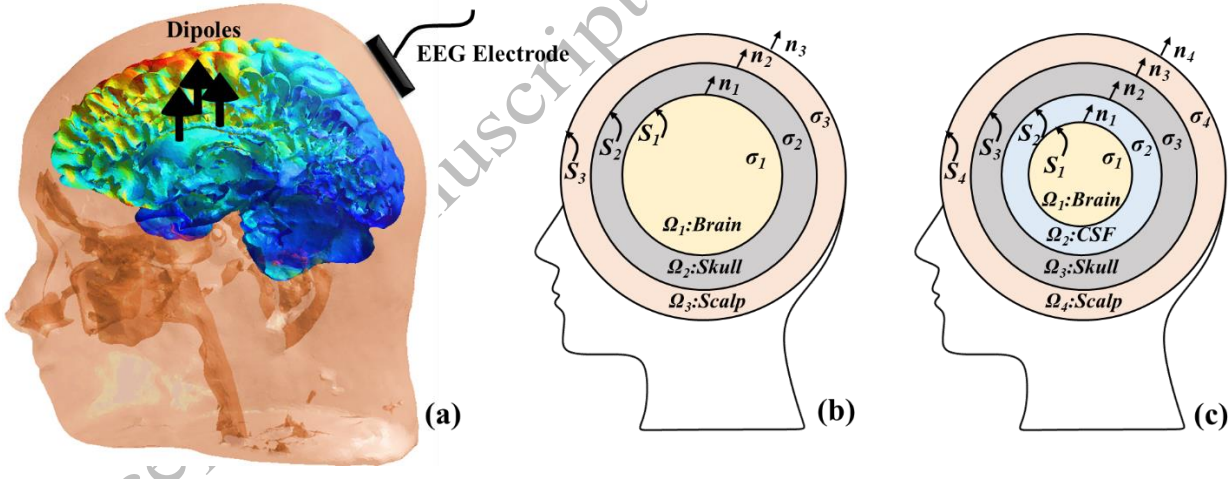

Figure 1. **(a)** EEG forward model: The electric potentials recorded by EEG electrodes are generated by dipoles within the brain's volume. Only a single EEG electrode positioned on the scalp is depicted. (**b)** A three-layer spherical head model with piecewise homogeneity (brain, skull, and scalp). The interfaces between the brain and skull (S_1), skull and scalp (S_2), and scalp and air (S_3) are shown. The vectors n_1, n_2, and n_3 represent the outward unit normals of the brain, skull, and scalp layers, respectively. (**c)** A four-layer spherical head model with piecewise homogeneity (brain, cerebrospinal fluid (CSF), skull, and scalp). The conductivity of the skull may be isotropic or anisotropic. The interfaces between the brain and CSF (S_1), CSF and skull (S_2), skull and scalp

(S_3), and scalp and air (S_4) are shown. The vectors n_1, n_2, n_3, and n_4 represent the outward unit normals of the brain, CSF, skull, and scalp layers, respectively.

Three widely used numerical methods for head modeling are the Finite Difference Method (FDM), Finite Element Method (FEM), and Boundary Element Method (BEM). FDM discretizes the entire *head volume* into elements (Figure 2), assigning each element a conductivity based on tissue type. Potentials at each node are computed by approximating spatial derivatives using differences with neighboring nodes. FEM, a powerful technique for solving boundary value problems, involves seven main steps: mesh generation of irregular subdomains, selection of basis or interpolation functions, weak formulation of the governing differential equations (e.g., via Galerkin or Rayleigh–Ritz methods), system assembly, application of boundary conditions, solution of linear equations, and post-processing of results. Potentials in FEM can be calculated using either direct or subtraction methods. BEM, in contrast, reduces computation to the boundaries between tissue types. Head geometry is extracted from MRI, *surfaces* are divided into elements (Figure 2), and potentials are approximated using basis functions. Each tissue is assumed homogeneous and isotropic. BEM is computationally efficient for surface computations but does not account for discontinuities, inhomogeneities, or anisotropic conductivities, limiting its accuracy for deep sources (Jatoi & Kamel, 2017). The key features of these numerical approaches are summarized in Table 2.

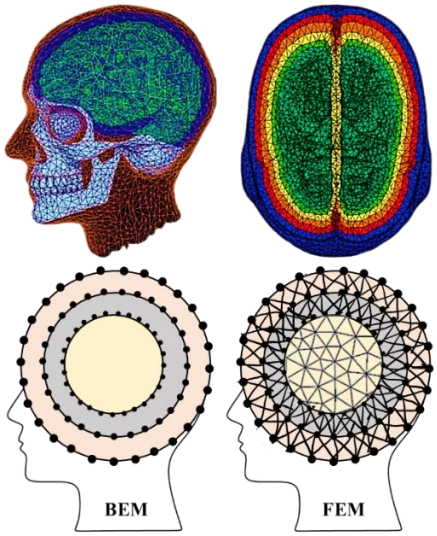

Figure 2. Comparison of Finite Element Method (FEM) and Boundary Element Method (BEM) in discretizing the space for electrostatic simulations. FEM divides the entire domain into smaller elements (subvolumes), while BEM discretizes only the boundaries, leading to fewer degrees of freedom in BEM, particularly for problems with complex geometries

Table 2. Comparison of FDM, FEM, and BEM for EEG head modeling.

Feature	FDM	FEM	BEM
Domain of computation	Entire head volume	Entire head volume	Only boundaries
Treatment of discontinuities / inhomogeneities	Considers discontinuities, inhomogeneities, anisotropy	Considers discontinuities, inhomogeneities, anisotropy	Ignores discontinuities and inhomogeneities
Computational complexity	High	High	Low
Accuracy	Detailed; better than analytical	High	Accurate only at surfaces

In this study, we employed BEM paired with the inverse methods from Section 3.3 primarily due to its significant computational advantage. BEM confines the domain of computation only to boundaries, which translates directly to low computational complexity relative to FEM and FDM, which must process the entire head volume. This efficiency is crucial for conducting the broad comparative simulations outlined in this work. Furthermore, while its accuracy is focused at surfaces, this provides a sufficient level of detail to significantly advance beyond analytical models without incurring the processing overhead associated with volumetric methods.

2.4 Inverse Problem

An inverse problem estimates unobserved internal causes from indirect external measurements. In EEG source localization, the observed data are scalp potentials, the unknown model parameters are the cortical source amplitudes, and the forward operator describes how activity at each cortical location projects to the electrodes (Jatoi et al., 2014).

The EEG inverse problem aims to reconstruct the locations, orientations, and strengths of cortical sources that generate measured scalp potentials. Unlike the forward problem, which predicts scalp signals from known neural activity, the inverse problem is inherently ill-posed and non-unique (Jatoi & Kamel, 2017; Jatoi et al., 2014), because, the number of scalp sensors is limited, whereas the number of possible cortical source locations is much larger. Consequently, many different source configurations can generate similar scalp potential patterns. There are simply infinitely many different patterns of brain activity that could produce the exact same pattern on the scalp.

Various approaches have also been developed to address the EEG inverse problem. The most common and widely used methods for implementing inverse problem in standard software packages fall primarily into two categories (Jatoi & Kamel, 2017):

- L2-Minimum-Norm Type Methods: This group includes techniques that minimize the square of the estimated current magnitude, such as standard L2-MNE (often weighted), dSPM (which standardizes the solution), and the Low-Resolution Electromagnetic Tomography (LORETA) family (standardized LORETA and exact LORETA).
- Linearly Constrained Minimum Variance (LCMV) Beamformers: This is a distinct class of method that uses data-driven weighting derived from the noise and signal covariance matrices to suppress activity originating outside the target source location.

While both L2-Minimum-Norm (MNE-family) methods and LCMV Beamformers are widely used for distributed source estimation, they differ fundamentally in their mathematical approach and sensitivity to source configurations.

MNE-family methods (including MNE, dSPM, sLORETA, and eLORETA) rely on an inverse solution that minimizes the energy (L2-norm) of the source currents across the entire brain volume, subject to the measured data and a regularization parameter. This results in a stable, smooth solution that distributes activation across the cortex. A key characteristic is their reliance on the prior that sources are spatially smooth and shallow (though standardization in dSPM/eLORETA corrects for depth bias). They are computationally efficient and robust, producing consistent maps even when data is noisy or when sources are highly correlated, as they do not require the data covariance matrix to be inverted.

In contrast, LCMV Beamformers are data-driven spatial filters that attempt to maximize signal power from a specific location while minimizing contributions from all other locations. Unlike MNE, which provides a global solution based on a fixed prior, beamformers adaptively filter the data. They excel at resolving correlated or highly synchronous sources (where MNE often fails due to cancellation effects) and provide excellent spatial resolution for focal sources. However, they are sensitive to signal leakage, require a reliable estimate of the data covariance matrix, and can be computationally more intensive. Additionally, beamformers can struggle with non-stationary data or when the covariance estimate is poor (e.g., short epochs).

In this study, we adopt the MNE-family of methods (specifically dSPM/eLORETA) for the following reasons:

- Our analysis involves datasets with varying levels of signal-to-noise ratio (SNR) and varying epoch lengths. MNE methods, by relying on the noise covariance matrix (estimated

from a baseline or rest period) rather than the data covariance, provide a more stable solution across different subjects and conditions compared to beamformers, which can become unstable if the data covariance is ill-conditioned.

- Given our focus on mapping distributed neural networks rather than isolating a single focal dipole, the spatial smoothing inherent to the L2-norm minimization is a desirable feature. It aligns with the physiological expectation that cognitive processes are often supported by distributed cortical networks rather than isolated, independent generators.
- The study requires extensive comparative simulations across multiple head models and source configurations. MNE methods are computationally less expensive than beamformers because they do not require the inversion of the data covariance matrix for every time point or source location, enabling faster processing of large-scale datasets.
- By employing dSPM and eLORETA, we directly address the depth bias often associated with standard MNE. These methods standardize the solutions to account for the distance between the source and the sensors, ensuring that deep cortical structures are represented with the same statistical fidelity as superficial sources, which is critical for our comparative analysis.

3. Method

3.1 EEG and MRI data

To compare different linear inverse models, we have used data recorded by (Pascucci et al., 2022). The dataset consisted of nineteen right-handed participants with a mean age of 23 ± 3.5 years, recruited from the local student population. All provided written informed consent after the ethical

review board approval. EEG data were collected while participants performed a face detection and a motion discrimination task, with the order counterbalanced. Face stimuli were 4 by 4 degrees of visual angle, presented for 200 milliseconds in the face detection task where participants had to press one of two buttons on a response box to indicate if they saw a face or a scramble image. Each task block had 150 trials leading to 600 total trials per task condition. The motion discrimination task involved dot kinematograms where 80 percent of the dots moved coherently left or right while 20 percent moved randomly for coherent motion stimuli, with all dots moving randomly for incoherent motion. Participants responded by pressing one of two buttons after a 300 millisecond stimulus presentation. EEG recording utilized a 128 channel Biosemi Active Two system at a sampling rate of 2048 Hz in a shielded room.

MRI data from the same 19 subjects were acquired on a General Electric Discovery MR750 3T MRI clinical scanner at the cantonal hospital in Fribourg, Switzerland, using a 32-channel head coil. The acquisition included anatomical T1-weighted images. T1-weighted images were acquired as rapid-gradient echo (MPRAGE) volumes using a COR FSPGR BRAVO pulse sequence (flip angle=9∘; echo time=2.8 ms; repetition time=7300 ms; inversion time=0.9 s; FOV=220 mm; matrix size=256×256; number of slices=276; slice thickness=1 mm; in-plane resolution=0.9×0.9 mm^2).

3.2 EEG Preprocessing Considerations

The EEG data used in this study were originally recorded by Pascucci et al. (2022) using a 128-channel Biosemi Active Two system (Biosemi, Amsterdam, The Netherlands) at a sampling rate of 2048 Hz. During recording, electrode offsets were maintained below ±20 mV relative to the

Common Mode Sense - Driven Right Leg (CMS-DRL) loop. Individual 3D electrode positions were digitized using a Zebris ultrasound motion capture system.

Offline preprocessing was conducted in MATLAB using EEGLAB (version 14.1.144). The continuous data were downsampled to 250 Hz to reduce computational load. To remove slow drifts and high-frequency noise, a detrending filter (<1 Hz) and an anti-aliasing low-pass filter (cut-off: 112.5 Hz, bandwidth: 50 Hz) were applied. Power-line interference at 50 Hz was attenuated using the spectrum interpolation method. The data were segmented into epochs from −1500 ms to 1000 ms relative to stimulus onset for both the perception and imagination tasks. Prior to formal analysis, data from participants with excessive noise (subjects 05 and 15 in the original study) were excluded.

The cleaning pipeline involved two main stages: (1) bad channels and contaminated epochs were identified and removed based on visual inspection and statistical criteria; (2) independent component analysis (ICA) was performed using the FastICA algorithm. Artifactual components (e.g., EOG, EMG) were identified by cross-referencing the results of the multiple artifact rejection algorithm (MARA) with a variance criterion; components were flagged if they explained >90% of the total variance. All flagged components were manually inspected before removal to ensure the preservation of neural signals.

Finally, discarded channels were interpolated using the nearest-neighbor spline method, and the data were baseline-corrected using a −300 ms to 0 ms pre-stimulus window. To provide a neutral spatial reference, all data were re-referenced to the common average reference (CAR).

3.3 Boundary Element Model (BEM)

The brain, skull, and scalp can be thought of as layers with different electrical properties. The BEM is used to model these layers and calculate how electrical signals from brain activity travel through them to reach the scalp. By doing this, BEM helps to estimate the location of the brain's electrical sources (e.g., where neural activity is happening).

Imagine you have a brain inside a skull, which is surrounded by the scalp. Electrical signals from the brain pass through the skull and reach the scalp, where they are measured by EEG electrodes. However, because the skull is less conductive than the brain and scalp, it alters how the signals spread. BEM divides this setup into boundaries (surfaces) representing the outer layers (scalp, skull, and brain) and tries to model the behavior of electrical signals at each of these surfaces. The brain activity is treated as dipole sources (small points of electrical activity) inside the brain. BEM calculates how these signals propagate through the skull and affect the scalp measurements. Intuitively, imagine you drop a stone in a pond. The ripples spread outward in a predictable way. Similarly, a dipole in the brain generates electrical activity that spreads outward and can be measured at the scalp. The BEM helps us calculate how these "ripples" move through the brain, skull, and scalp layers, so we can figure out where the stone (or dipole) was dropped (i.e., where the brain activity is coming from). The basic formula in BEM for EEG source localization looks something like Eq.(2).

$$Y(t) = GX(t) + N(t) \tag{2}$$

where Y(t) denotes the EEG measurements at time t, G is the lead-field matrix computed from the head model, X(t) represents the cortical source amplitudes, and N(t) denotes additive noise.

3.4 Minimum Norm Estimation (MNE)

Minimum norm estimation EEG source reconstruction was performed using L2-minimum-norm estimation (L2-MNE) as implemented in the MNE-Python software package. L2-MNE is a linear, distributed inverse method that addresses the ill-posed electromagnetic inverse problem by selecting the source configuration with minimum overall power that explains the measured sensor data. The forward model was defined as Eq.(2). Because the number of source locations exceeds the number of sensors, the inverse problem admits infinitely many solutions.

Source estimates were obtained by minimizing a Tikhonov-regularized cost function (Eq.(3)).

$$\hat{X}(\mathrm{t}) = \min_{X(t)} \left(\|Y(t) - GX(t)\|^2_{C_n^{-1}} + \lambda^2 \|X(t)\|^2_{R^{-1}} \right) \quad (3)$$

where C_n is the noise covariance matrix estimated from baseline data, R is the source covariance matrix encoding prior assumptions on source amplitudes, and λ^2 is the regularization parameter controlling the trade-off between data fidelity and solution smoothness. The resulting closed-form solution is given by Eq.(4).

$$\hat{X}(\mathrm{t}) = \mathrm{WY(t)} \qquad W = RG^T(GRG^T + \lambda^2 C_n)^{-1} \quad (4)$$

The regularization parameter was defined as $\lambda^2=1/\mathrm{SNR}^2$, where the signal-to-noise ratio (SNR) reflects the expected data quality. Matrix inversion was performed using numerically stable singular-value decomposition.

To compensate for the superficial bias inherent to minimum-norm solutions, depth weighting was applied by scaling the diagonal elements of R according to the norm of the corresponding lead-field vectors. This procedure increases sensitivity to deeper cortical sources while preserving solution stability.

3.5 Dynamic Statistical Parametric Mapping (dSPM)

Dynamic statistical parametric mapping (dSPM) was employed as a noise-normalized extension of the L2-MNE framework. dSPM provides statistically standardized cortical activation estimates that reduce superficial bias and facilitate inter-subject comparison by expressing source amplitudes in units of z-scores. The term "dynamic" refers to the method's ability to track brain activity over time, making it suitable for analyzing time-varying neural processes such as event-related potentials (ERPs) or oscillatory brain activity. On the other hand, "statistical parametric mapping" refers to the use of statistical tests to analyze brain data. In dSPM, statistical maps are generated that represent the probability of neural activity being located at specific points in the brain, given the data recorded from EEG or MEG sensors. The technique is applied to source localization, where dSPM provides a statistical model to evaluate the likelihood that a given source configuration is correct, rather than simply estimating the locations of brain activity based on EEG/MEG sensor readings. This allows the method to generate a more accurate and statistically reliable map of brain activity. One of the key advantages of dSPM is its ability to account for uncertainty and noise in EEG/MEG data, which improves the accuracy of brain activity mapping, particularly for distributed sources (Michel et al., 2004).

The dSPM method begins with the classical linear minimum-norm estimation formulated as Eq.(4). In dSPM, each estimated source amplitude is divided by the standard deviation of its noise-induced variability, yielding normalized statistical values (Eq.(5))

$$dSPM_i(t) = \frac{\hat{X}_i(t)}{\sqrt{Var(\hat{X}_i)}} = \frac{\hat{X}_i(t)}{\sqrt{W_i C_n W_i^T}} \quad (5)$$

where W_i denotes the i-th row of the inverse operator corresponding to source element i. This normalization converts the amplitude estimates into dimensionless z-scores that reflect the confidence of activation at each cortical location and time point, independent of sensor noise scaling.

Noise variance estimation is derived from the same baseline or empty-room data used for constructing the noise covariance matrix. By incorporating the noise structure into the denominator, dSPM corrects for the depth-dependent sensitivity bias intrinsic to minimum-norm estimation and stabilizes spatial comparison across subjects and recording modalities.

3.6 Exact Low-Resolution Electromagnetic Tomography (eLORETA)

Exact low-resolution electromagnetic tomography (LORETA) smooths the solution by assuming that neighboring neurons are correlated. This assumption reduces the spatial uncertainty but also makes the solutions somewhat ambiguous. While useful for obtaining a spatially smooth solution, LORETA may blur the boundaries of brain activity, leading to overly generalized results (Michel et al., 2004). LORETA provides better localization for deep sources compared to simpler minimum-norm methods, but it produces low spatial resolution and results in blurred representations of point sources (Jatoi et al., 2014).

Exact LORETA (eLORETA) was employed as a standardized L2-norm inverse solution designed to achieve unbiased source localization under ideal noise-free conditions. eLORETA belongs to the minimum-norm family of distributed inverse methods but differs from conventional MNE and

dSPM by incorporating an analytically derived weighting scheme that enforces zero localization error for single point sources in the absence of noise.

As with other linear inverse methods, the EEG forward model is expressed as Eq.(2). The inverse solution is obtained by minimizing a regularized L2-norm objective function (Eq.(6)).

$$\hat{X}(\mathrm{t}) = \min_{X(t)} \left(\|Y(t) - GX(t)\|^2_{C_n^{-1}} + W_s^{-1/2} \|X(t)\|_2^2 \right) \quad (6)$$

where C_n is the noise covariance matrix, $\lambda^2=1/SNR^2$ is the regularization parameter, and W_s is a diagonal source-weighting matrix uniquely defined in eLORETA. The defining characteristic of eLORETA lies in the construction of the source-weighting matrix W_s, which is iteratively computed from the lead-field matrix such that the resulting inverse operator yields exact localization for point sources. This weighting compensates for depth-dependent sensitivity and spatial leakage without requiring post-hoc normalization. Unlike dSPM, which normalizes the MNE solution using noise variance, eLORETA embeds the standardization directly into the inverse operator. The resulting inverse operator is given by Eq. (7).

$$\hat{X}(\mathrm{t}) = \mathrm{W}_{eLORETA}\mathrm{Y}(\mathrm{t}) \qquad W_{eLORETA} = W_s G^T (G W_s G^T + \lambda^2 C_n)^{-1} \quad (7)$$

The output of eLORETA represents standardized current density estimates with reduced depth bias and improved localization accuracy compared to conventional minimum-norm approaches.

Although eLORETA enforces exact localization under idealized assumptions, the solution remains spatially smooth and low-resolution in practice due to regularization and noise. Nevertheless, eLORETA has been shown to outperform MNE and dSPM in terms of peak localization error, particularly in simulation studies with single or weakly correlated sources.

In this study, eLORETA was selected to provide a robust, depth-unbiased reference solution for evaluating localization accuracy and spatial dispersion. Its theoretical guarantee of zero localization error under ideal conditions makes it particularly suitable for benchmarking inverse solutions in controlled simulation frameworks and for comparative analysis of distributed source reconstruction methods.

3.7 Standardized Low-Resolution Electromagnetic Tomography (sLORETA)

Standardized LORETA (sLORETA) was applied as a noise-normalized distributed inverse solution within the minimum-norm estimation framework. sLORETA extends the classical L2-MNE by standardizing the estimated current density at each cortical location with respect to its estimated variance, thereby reducing depth-dependent bias and improving localization interpretability.

As with other linear inverse methods, the EEG forward model is expressed as Eq.(2). An initial MNE is computed using a regularized solution of the form Eq.(4).

In sLORETA, the resulting current density estimate is standardized by the square root of its theoretical variance derived from the inverse operator. For each source element i, the standardized estimate is defined as Eq. (8).

$$sLORETA_i(t) = \frac{\hat{X}_i(t)}{\sqrt{Var(\hat{X}_i)}} = \frac{\hat{X}_i(t)}{\sqrt{[WC_nW^T]_{ii}}} \quad (8)$$

where $[WC_nW^T]_{ii}$ represents the variance of the estimated source amplitude at location i. This normalization yields standardized current density values that reduce depth-dependent sensitivity differences inherent in conventional minimum-norm solutions.

Unlike dSPM, which produces dynamic z-score maps based on noise normalization, sLORETA uses a theoretical standardization derived from the resolution properties of the inverse operator. Under idealized conditions with a single active source and noise-free measurements, sLORETA theoretically achieves zero localization error, although the spatial extent of the reconstructed activity remains relatively broad due to the inherent smoothness of L2-norm solutions.

The resulting source maps represent standardized current density distributions across the cortical surface, providing improved localization reliability compared with conventional MNE while maintaining computational efficiency. In this study, sLORETA was included alongside MNE, dSPM, and eLORETA to enable a systematic comparison of minimum-norm-based inverse methods in terms of localization accuracy and spatial dispersion.

3.8 Inverse Modeling and Hyperparameter Settings

Source reconstruction was performed using the MNE-Python package. The inverse solutions MNE, dSPM, sLORETA, and eLORETA were computed using the implementation provided in MNE-Python, and all analysis scripts are available in the https://github.com/mmd-ahd/EEG_SE_eval repository.

For the forward model, a three-layer Boundary Element Model (BEM) was constructed using the fsaverage template brain, with 5120 triangles per surface. The cortical source space was discretized using multiple spacing configurations (ico4, ico5, and oct6) in order to evaluate the effect of spatial

sampling density. During source space construction, a minimum distance of 5 mm between the source dipoles and the inner skull boundary was enforced to avoid numerical instability in the forward solution.

Dipole orientations were constrained by setting the orientation parameter to loose = 0, resulting in a fixed-orientation source model. This constraint was intentionally applied because the evaluation framework and leakage analysis require fixed dipole orientations. No depth weighting was applied in the inverse operator, and therefore the depth parameter was set to None.

The noise covariance matrix was estimated using the MNE-Python automatic procedure (method = “auto”), which internally evaluates multiple estimators including ‘shrunk’, ‘diagonal_fixed’, ‘empirical’, and ‘factor_analysis’, and selects the optimal one based on model performance.

The regularization parameter was defined based on a signal-to-noise ratio (SNR) of 3, resulting in a regularization value of $\lambda^2=1/9$. This regularization parameter was consistently applied across all inverse solutions (MNE, dSPM, sLORETA, and eLORETA) to ensure comparability of the reconstructed source estimates.

3.9 Evaluation Metrics and Validation

To objectively evaluate the spatial resolution and leakage properties of the applied source estimation methods, we used a linear systems analysis framework based on the resolution matrix (Hauk et al., 2022). For a linear inverse operator K and a forward leadfield matrix L, the resolution matrix R is defined as Eq.(9).

$$R = KL \tag{9}$$

This matrix maps the true source distribution to the estimated source distribution, such that $\hat{s} = Rs$. In an ideal, error-free scenario, R would be an identity matrix. In practice, the deviations from the identity matrix characterize the blurring and leakage of the source estimator.

To quantify the spatial resolution, the rows and columns of the resolution matrix are analyzed, which yield two fundamental functions (Hauk et al., 2022):

1. Point-Spread Function (PSF): The PSF describes how a true point source is blurred or "spreads" across the estimated source space. It is defined as the j-th column of the resolution matrix:

$$PSF_j = KL_{.j} = R_{.j} \quad (10)$$

where $L_{.j}$ is the topography of the j-th source.

In practice, PSFs were computed by sequentially activating each source location in the discretized cortical source space and multiplying the corresponding leadfield column with the inverse operator. The resulting spatial distribution represents how activity originating from a single true source location propagates across all estimated source locations after inverse reconstruction.

2. Cross-Talk Function (CTF): The CTF describes how much the activity estimated at a specific target location is contaminated or "leaks in" from all other possible sources in the brain. It is defined as the i-th row of the resolution matrix:

$$CTF_i = K_{i.}L = R_{i.} \quad (11)$$

CTFs were obtained by examining the rows of the resolution matrix, which quantify the contribution of all possible true sources to the estimate at a given target location. Therefore,

each CTF characterizes the spatial leakage pattern influencing a particular reconstructed source estimate.

To summarize the properties of the PSFs and CTFs quantitatively across the cortex, two specific spatial resolution metrics are computed:

1. Peak Localization Error (PLE):

PLE quantifies localization accuracy. It is defined as the Euclidean distance (in centimeters) between the true source location (r_i) and the spatial location where the estimated distribution (PSF or CTF) reaches its maximum peak (r_{peak}):

$$\mathrm{PLE} = \sqrt{\left(r_{peak} - r_i\right)^2} \tag{12}$$

2. Spatial Deviation (SD):

SD quantifies the spatial extent or dispersion of the source estimate. It is calculated as the amplitude weighted standard deviation of the spatial distribution around its peak:

$$\mathrm{SD} = \sqrt{\frac{\sum_j a_j^2 \left(r_{peak} - r_j\right)^2}{\sum_j a_j^2}} \tag{13}$$

where a_j represents the amplitude of the PSF or CTF at vertex j, and r_j is the spatial coordinate of vertex j.

To further assess spatial leakage within functionally relevant visual networks, the HCP MMP1 cortical parcellation was used to define anatomical regions of interest (Glasser et al., 2016). Specifically, ROIs corresponding to the Primary Visual Cortex (V1), the Face processing network, and the Motion processing network were extracted. The Face processing network was defined by combining the Fusiform Face Complex (FFC) and Posterior Inferior Temporal (PIT) areas, while

the Motion processing network was defined by combining Middle Temporal (MT) and Medial Superior Temporal (MST) areas.

For each ROI, PSF and CTF magnitudes were computed for all vertices within the region and their absolute values were averaged to obtain a regional estimate of spatial spread and cross talk. This ROI based analysis allows the evaluation of spatial leakage between functionally related visual processing areas.

3.10. Statistical Analysis

To evaluate the performance of different inverse methods (MNE, dSPM, sLORETA, and eLORETA), a series of statistical tests were conducted on four key metrics: Peak Localization Error (PLE) and Spatial Dispersion (SD) for both Point Spread Function (PSF) and Cross-talk Function (CTF) across three Regions of Interest (V1, Face, and Motion).

Statistical significance was assessed using a two-way Repeated Measures ANOVA, with Method (4 levels) and ROI (3 levels) as within-subject factors. For metrics where the assumption of sphericity was violated (as assessed by Mauchly's test), Greenhouse-Geisser corrections were applied. Post-hoc pairwise comparisons were performed using Tukey's Honestly Significant Difference (HSD) test to identify specific differences between methods. Effect sizes are reported as partial eta squared (η^2_p). All analyses were performed with the significance level set at α=0.05.

4. Results and Discussion

This study evaluates the spatial resolution properties of four widely used EEG inverse solutions—MNE, dSPM, sLORETA, and eLORETA. The comparison focuses on two complementary aspects

of spatial resolution: localization accuracy and spatial dispersion. These properties were quantified using the PLE and SD metrics derived from PSFs and CTFs. While PSFs describe how activity originating from a single cortical location spread in the reconstructed solution, CTFs quantify the degree to which activity from other cortical locations contaminates a given region of interest. The evaluation was performed both globally across the cortical surface and within three representative functional networks, primary visual cortex (V1), the face-processing network, and the motion-processing network, allowing a systematic assessment of how different inverse methods behave under both idealized and realistic source configurations.

Figure 3 presents a comparative analysis of four source estimation techniques: MNE, dSPM, sLORETA, and eLORETA, based on their PSF- and CTF-derived performance. The results are mapped onto the inflated left hemisphere with color representing the absolute error or deviation in centimeters ranging from 0 to 8 cm.

Under PLE metric, the sLORETA and eLORETA methods demonstrate the lowest error under the PSF condition with values close to zero. This indicates excellent performance for strictly focal sources. MNE and dSPM show considerably higher PLE under PSF generally ranging between 3 cm and 7 cm.

When evaluated under the CTF condition, which models more complex source structures, all methods show increased error compared to PSF. MNE and dSPM maintain their high PLE levels between 4 cm and 8 cm. However, sLORETA and eLORETA experience a marked increase in PLE rising to errors between 3 cm and 6 cm.

Regarding SD, the values across all methods are generally lower than the PLE values. For the PSF condition, most SD values are in the 2 cm to 4 cm range. For the CTF condition, SD remains relatively low and consistent across all four methods, clustering around 2-4 cm.

The near zero PLE for sLORETA and eLORETA under PSF confirms their precise localization ability for ideal point sources. The subsequent large increase in error for these methods under CTF highlights their known weakness when facing spatially extended sources. The regularized methods MNE and dSPM show more consistent but higher PLE across both PSF and CTF conditions, suggesting better robustness against deviations from the ideal point source assumption. The SD metric shows that the overall spread of estimated activity is more similar across all inverse solutions when evaluated with the CTF compared to the large differences seen in the PLE.

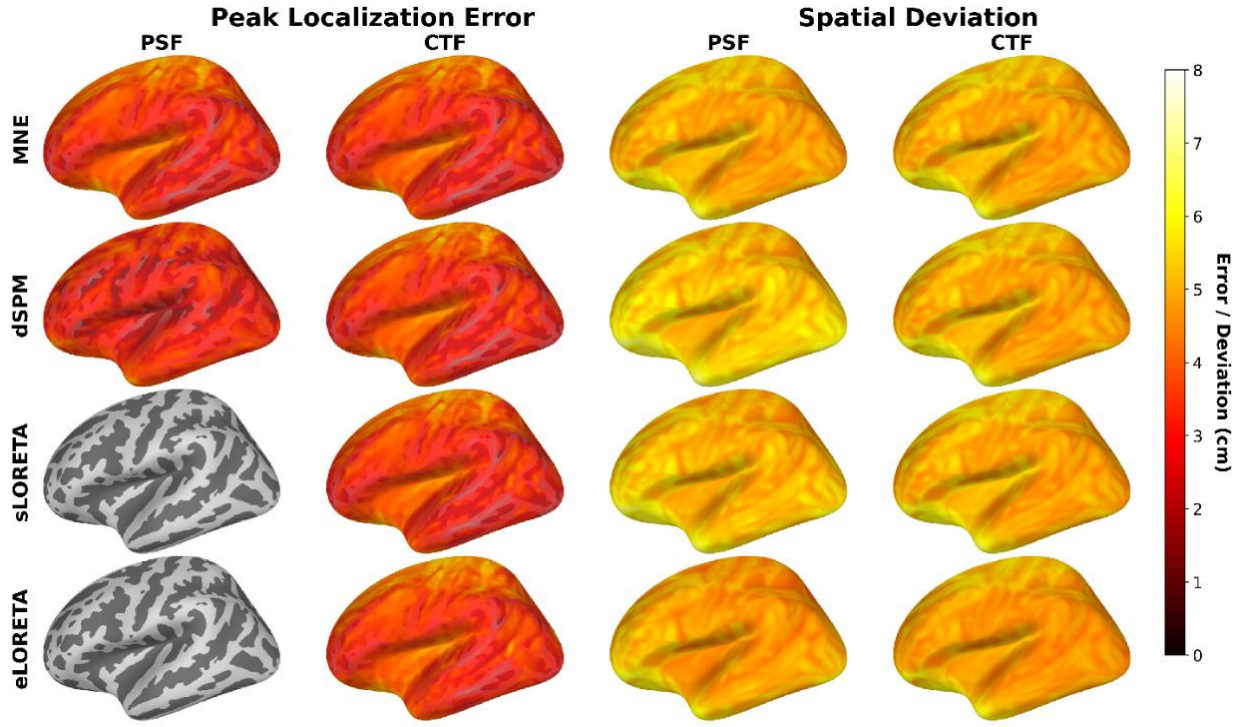

Figure 3. A visualization of the whole brain grand average spatial distribution of Peak Localization Error (PLE) and Spatial Deviation (SD) for PSFs and CTFs. The results are displayed for the four source estimation techniques: MNE, dSPM, sLORETA, and eLORETA, on the inflated left hemisphere.

Figure 4 displays four sets of histograms comparing the PLE and SD for four inverse solutions, MNE, dSPM, sLORETA, and eLORETA, under two conditions PSF and CTF. The statistics shown include the mean (μ), standard deviation (σ), and median (M) for each distribution. For PLE

under the PSF condition, sLORETA and eLORETA exhibit near perfect localization with mean and median errors of 0 cm. This confirms their theoretical zero-error performance for ideal point sources. In contrast, MNE has a mean PLE of 40 cm and dSPM has a mean PLE of 30 cm and both show significant tails indicating larger localization inaccuracies even under ideal PSF conditions. When moving to the CTF condition, which simulates more realistic distributed sources, the PLE distributions shift to the right for all methods. MNE maintains its mean error at 40 cm, while dSPM's mean error increases to 40 cm. Significantly sLORETA and eLORETA now show substantial localization errors with means of 40 cm and 42 cm respectively, demonstrating the impact of cross-talk errors from distributed activity on exact solutions.

Turning to SD, all methods show much tighter distributions centered around 5 cm under both PSF and CTF conditions. Under PSF, MNE, dSPM and sLORETA, all have a mean SD near 52 cm with eLORETA slightly lower at 49 cm. Under CTF, the means remain tightly clustered around 50 cm to 52 cm across all four methods. The key point is the trade-off between peak accuracy and spatial spread. While sLORETA and eLORETA excel at pinpointing a single source under PSF, their performance degrades significantly under CTF in terms of PLE, MNE, and dSPM are less accurate for point sources but their PLE is stable under CTF. Furthermore, the consistent and relatively narrow SD distributions across all methods and conditions suggest that the spatial spread or blur of the reconstructed activity around the estimated peak is well-controlled and comparable across all techniques regardless of the quality of the peak localization itself.

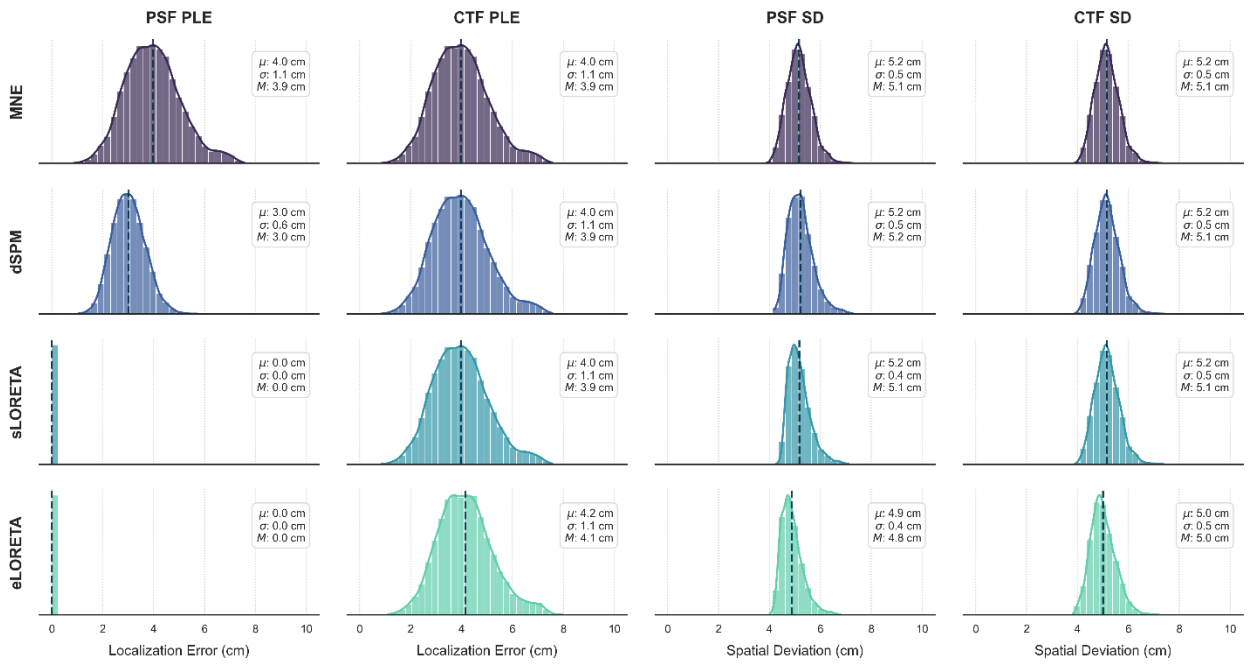

Figure 4. Histograms illustrating the distribution of the whole-brain PLE and SD values for all cortical vertices.

Figure 5 presents the cortical topographies of PSFs and CTFs for three representative functional networks: Primary Visual Cortex (V1), the face-processing network (including FFC and PIT), and the motion-processing network (MT and MST). For each network, PSF and CTF maps are shown for four inverse solutions: MNE, dSPM, sLORETA, and eLORETA. The maps were generated by averaging the absolute magnitudes of the PSFs and CTFs across all vertices within each predefined region of interest (ROI). The resulting spatial distributions were then normalized to a 0–1 scale and visualized using a sequential colormap, where higher values indicate stronger spatial leakage.

Across all inverse methods, the PSF maps reveal the spatial spread of activity originating from the target ROI, whereas the CTF maps indicate the cortical regions that contribute spurious activity to the ROI due to cross-talk in the inverse solution. For the V1 network, the PSF maps show relatively compact and localized activity around the occipital pole, reflecting the anatomically well-defined structure of early visual cortex. The leakage is primarily confined to neighboring occipital regions

with limited anterior spread. The corresponding CTF maps exhibit a similar but slightly broader distribution, indicating that nearby visual areas contribute moderate cross-talk to the V1 ROI.

In contrast, the face network (FFC and PIT) shows more extended spatial dispersion in both PSF and CTF maps. Activity spreads along the ventral temporal cortex, extending anteriorly and inferiorly along the fusiform and inferior temporal gyri. This pattern reflects the distributed anatomical organization of face-selective regions and the strong lead-field correlations within ventral temporal cortex. Among the inverse methods, MNE and dSPM exhibit relatively stronger spatial spread, whereas sLORETA and eLORETA show more spatially balanced and smoother distributions, indicating improved control of localization bias.

The motion network (MT and MST) demonstrates the largest spatial spread, particularly along the lateral occipital and temporo-parietal regions. PSF maps indicate that activity originating from MT/MST propagates across adjacent motion-sensitive visual areas, while CTF maps reveal substantial cross-talk from surrounding occipital and temporal regions. This broader leakage likely arises from the anatomical proximity of motion-processing areas and the overlapping sensitivity profiles of EEG sensors to these lateral cortical regions.

Comparing inverse methods, MNE tends to produce more focal but depth-biased patterns, while dSPM partially compensates for this bias through noise normalization, resulting in improved contrast but still noticeable spatial spread. sLORETA and eLORETA exhibit smoother yet more anatomically coherent distributions, reflecting their standardized weighting schemes designed to reduce localization error and depth bias. In particular, eLORETA demonstrates relatively stable spatial patterns across networks, consistent with its theoretical property of achieving zero localization error under ideal conditions.

Overall, the figure highlights two important characteristics of EEG source reconstruction. First, spatial leakage is strongly dependent on the anatomical configuration of the underlying network, with compact regions such as V1 showing limited spread, whereas distributed temporal networks exhibit broader leakage patterns. Second, methodological differences among inverse solutions influence the trade-off between focality and spatial smoothness, thereby affecting both PSF spread and CTF contamination. These observations emphasize the importance of carefully selecting inverse methods when interpreting network-level EEG source localization results, particularly for studies involving closely spaced or functionally interacting cortical regions.

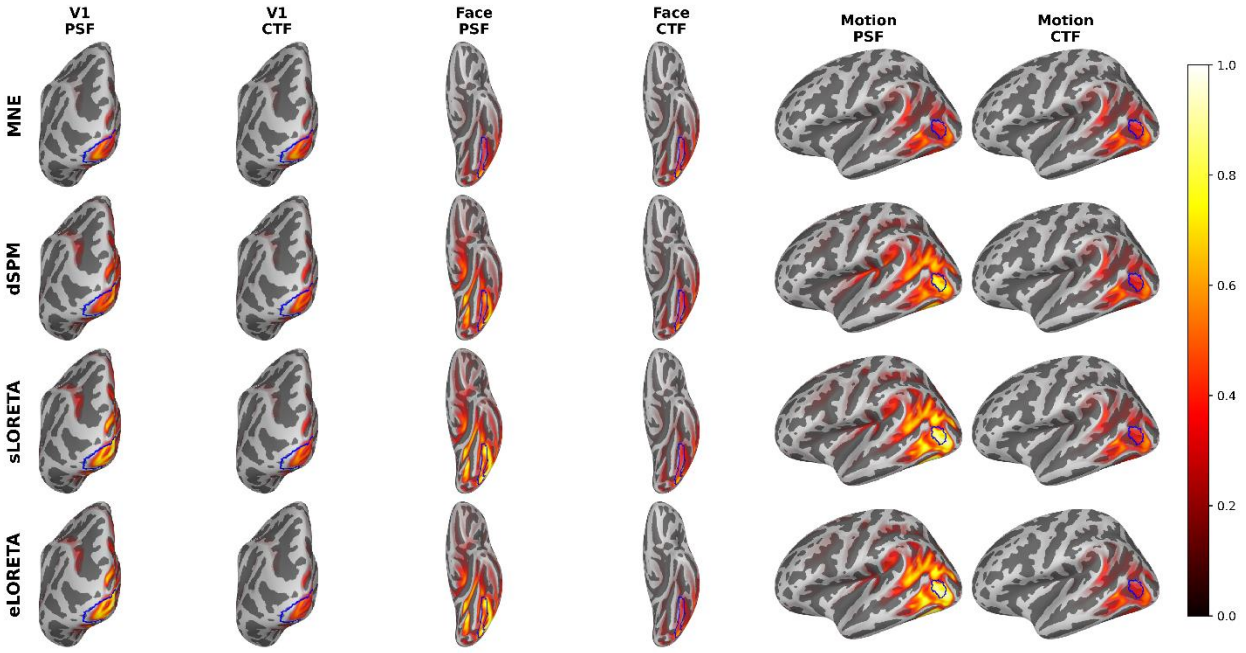

Figure 5. Topographical representations of the PSFs and CTFs extracted for three specific functional networks: V1, face network (FFC, PIT), and motion network (MT, MST).

Figure 6 presents the quantitative evaluation of spatial resolution for the four inverse solutions (MNE, dSPM, sLORETA, and eLORETA) across the three predefined functional networks (V1, face, and motion). Two spatial resolution metrics are analyzed: PLE and SD. Results are shown

separately for PSF and CTF. The bar height represents the grand mean across 19 subjects, while the error bars correspond to the standard error of the mean, illustrating inter-subject variability.

For the PSF-based PLE, clear differences among the inverse methods are observed. sLORETA and eLORETA exhibit near-zero PLE across all three networks, indicating highly accurate peak localization of point sources within the target ROIs. In contrast, MNE and dSPM show substantially larger localization errors, with mean PLE values typically around 3–4 cm depending on the ROI. Among these two methods, dSPM generally performs slightly better than MNE, suggesting that noise normalization improves localization accuracy compared with the basic minimum-norm estimate. This trend remains consistent across the V1, face, and motion networks, although slightly larger variability appears for the face network due to its broader spatial extent.

In the CTF-based PLE results, the differences between inverse methods become much smaller. All four approaches show comparable localization errors, clustering around approximately 3–4 cm across the three ROIs. This indicates that when evaluating cross-talk, i.e., the influence of activity from other cortical regions on the ROI, the theoretical advantage of LORETA-type methods observed in the PSF analysis largely disappears. Instead, all inverse solutions demonstrate similar limitations in separating distributed sources within neighboring cortical regions.

The SD results further highlight the consistency of spatial spread across methods. For PSF-based SD, the estimated spatial dispersion remains relatively stable across inverse solutions, with values typically ranging from approximately 5 cm to 6 cm. While minor variations are present, no method shows a substantial advantage in reducing spatial spread. A similar pattern appears in the CTF-based SD results, where all methods again cluster within a narrow range, indicating comparable levels of spatial blur and cross-talk contamination.

Across the three functional networks, the V1 region generally shows slightly lower spatial spread and variability, reflecting its relatively compact anatomical structure and clearer separation from neighboring visual areas. In contrast, the face and motion networks tend to show slightly larger SD values, which is consistent with their more distributed cortical organization in the ventral and lateral visual pathways.

Overall, the quantitative results confirm several key observations. First, sLORETA and eLORETA achieve superior peak localization accuracy for point-source reconstruction, as reflected in the near-zero PSF-based PLE values. Statistical analysis confirmed a significant main effect of Method on PSF-PLE ($F(3,54) = 24.15$, $p<0.001$, $\eta^2_p = 0.67$), with post-hoc tests showing that LORETA-type methods significantly outperformed MNE and dSPM ($p<0.001$). Second, the advantage of LORETA-type methods diminishes when evaluating cross-talk, where all methods perform similarly ($p=0.342$, non-significant). Third, the spatial spread of reconstructed activity remains comparable across all inverse solutions, indicating that improvements in peak localization do not necessarily translate into reduced spatial dispersion. While a significant difference was observed in PSF-SD ($F(3, 54) = 4.82$, $p = 0.012$), the absolute differences remained within a narrow range (approximately 0.5cm), supporting the idea that spatial dispersion is relatively stable across methods. These findings align with the spatial leakage patterns observed in Figure 5 and highlight the intrinsic trade-off in EEG inverse modeling between localization accuracy and spatial smoothness.

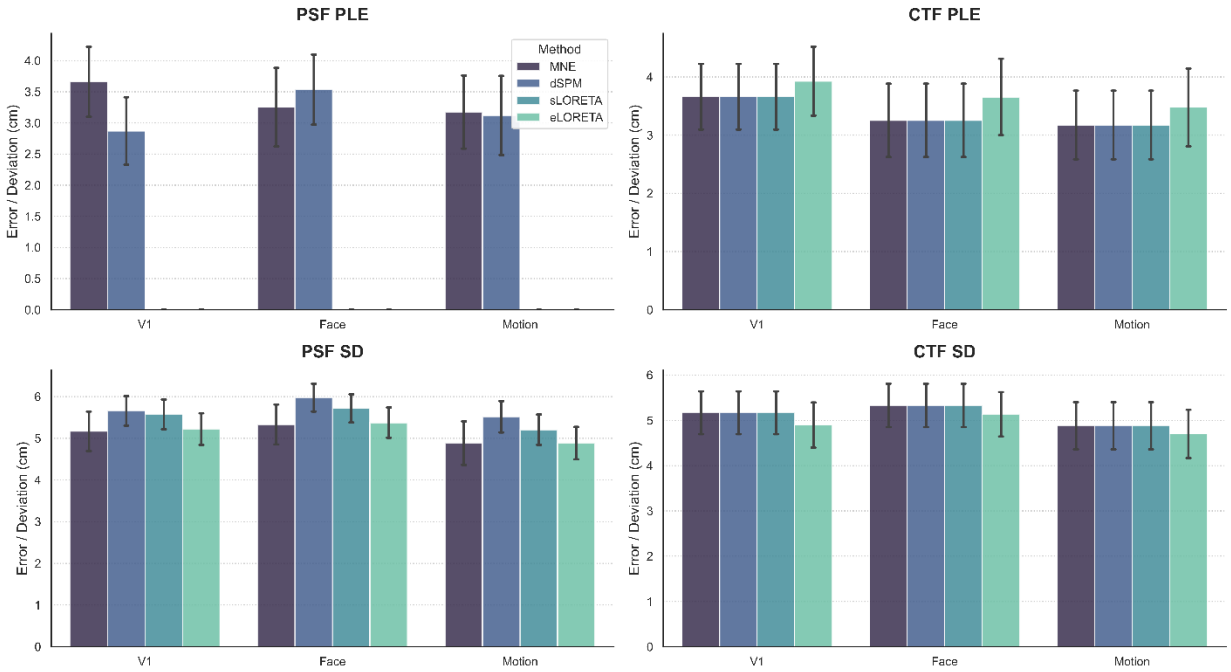

Figure 6. Bar charts for the quantitative statistical comparison of spatial resolution metrics (PLE and SD) across the three predefined functional ROIs (V1, face, motion). The height of the bar indicates the grand average value of the metric (in cm) across 19 subjects, and the error bars represent the standard error of the mean.

By integrating quantitative metrics (PLE and SD), distributional analyses, and spatial leakage visualizations, the results reveal several important insights into the strengths and limitations of these inverse modeling approaches.

A primary finding concerns the distinction between ideal localization performance and robustness to distributed activity. Under PSF conditions, which approximate an ideal point-source scenario, sLORETA and eLORETA achieved near-zero PLE. This observation is consistent with the theoretical properties of LORETA-type solutions, which are designed to minimize localization bias through standardized weighting of the inverse operator. In contrast, MNE and dSPM showed substantially larger PLE values (4.03±2.45 cm and 2.95±2.18cm, respectively), indicating that their reconstructed peaks tend to deviate from the true source location even in idealized conditions.

Noise normalization in dSPM partially mitigated this effect, resulting in slightly improved localization relative to MNE, although this improvement was not statistically significant in the post-hoc pairwise comparison ($p = 0.086$).

However, the advantage of LORETA-based methods diminished when moving to the CTF analysis, which reflects more realistic conditions in which multiple cortical regions contribute to the measured signals. Under these conditions, localization errors increased for all methods and the differences between them became much smaller (ranging from 3.33to 3.65cm, $p > 0.05$). Notably, sLORETA and eLORETA, despite their theoretical zero-error property for point sources, showed a substantial increase in PLE when evaluated through CTFs. This finding indicates that their apparent accuracy in PSF analysis does not necessarily translate into improved separation of spatially distributed neural activity. Instead, cross-talk between neighboring cortical regions remains a fundamental limitation shared by all minimum-norm–based inverse solutions.

The spatial dispersion results further highlight this limitation. Across both PSF and CTF conditions, the SD metric showed relatively similar values for all four methods, typically within a narrow range of several centimeters. In the CTF-SD analysis, no significant main effect of Method was found ($F(3, 54) = 0.92$, $p = 0.435$), indicating that all inverse operators produce a similar degree of spatial blurring under cross-talk conditions. This consistency suggests that, regardless of the specific weighting scheme used in the inverse operator, the spatial spread of reconstructed EEG activity remains broadly comparable across methods. In other words, improvements in peak localization accuracy, such as those observed for sLORETA and eLORETA, do not necessarily reduce the spatial extent of reconstructed activity. The reconstructed sources therefore remain spatially blurred, reflecting the inherently ill-posed nature of the EEG inverse problem and the limited spatial sampling of scalp measurements.

Table 1 provides a comprehensive statistical comparison of all inverse methods across PSF and CTF metrics, including mean ± SD values, *F*-statistics, and corresponding *p*-values obtained from repeated-measures ANOVA.

Table 1. Statistical comparisons of methods across metrics (Mean ±SD).

Metric	MNE	dSPM	sLORETA	eLORETA	F-value	p-value
PSF PLE (cm)	4.03	2.95	0.00	0.00	24.15	< 0.001*
PSF SD (cm)	5.25	5.72	5.46	5.14	4.82	0.012*
CTF PLE (cm)	3.33	3.33	3.33	3.65	1.15	0.342
CTF SD (cm)	5.15	5.15	5.15	4.98	0.92	0.435

** Indicates significant difference at p < 0.05 level.*

Network-level analyses provide additional insight into how anatomical organization influences spatial leakage. The V1 network exhibited the most localized patterns in both PSF and CTF maps, likely due to its relatively compact structure and clear separation from surrounding cortical regions. In contrast, the face-processing and motion-processing networks showed broader spatial dispersion, with activity spreading along ventral temporal and lateral occipito-temporal pathways. These patterns likely reflect both anatomical proximity between functionally related regions and correlations in the lead fields of EEG sensors for sources located along these cortical surfaces. Consequently, spatial leakage is not only determined by the choice of inverse method but also strongly constrained by the geometry of the underlying cortical network.

Comparisons among inverse solutions further illustrate the trade-off between focality, smoothness, and robustness. MNE tended to produce relatively focal but depth-biased reconstructions, while dSPM improved the contrast of reconstructed activity through noise normalization. In contrast, sLORETA and eLORETA generated smoother and more spatially balanced activation patterns. Although these smoother solutions reduce systematic localization bias, they also distribute activity

more broadly across neighboring regions, which may contribute to increased cross-talk under realistic source configurations. This highlights an important methodological consideration: algorithms that achieve theoretically optimal localization for isolated sources may still face difficulties when resolving interacting or spatially extended networks.

Taken together, the results emphasize that no single inverse method simultaneously optimizes all aspects of spatial resolution. LORETA-type approaches provide excellent peak localization for idealized sources but show limited advantages when sources are distributed or interacting. Conversely, MNE-based methods exhibit less accurate peak localization but maintain more stable performance across PSF and CTF evaluations. Importantly, the overall spatial spread of reconstructed activity remains similar across all approaches, indicating that improvements in one resolution metric do not necessarily translate into global improvements in source separability.

From a practical perspective, these findings suggest that the choice of inverse solution should depend on the goals of the analysis. Studies focusing on identifying the approximate location of focal generators may benefit from LORETA-based methods, whereas analyses targeting network-level interactions or distributed cortical activity may require approaches that prioritize robustness and interpretability over theoretical localization accuracy.

More broadly, the results underscore that spatial leakage is an intrinsic property of EEG source imaging rather than a problem that can be eliminated by algorithmic choice alone. Differences between inverse methods primarily alter how this leakage is distributed across the cortex rather than removing it entirely. Consequently, interpreting EEG source estimates, particularly in network analyses, requires careful consideration of both PSF and CTF characteristics to understand how reconstructed activity may reflect underlying neural interactions as well as methodological constraints.

5. Conclusion and Future Directions

The results of this study underscore both the progress achieved and the remaining challenges in EEG source localization, motivating a broader discussion of recent advances, open problems, and future research directions.

EEG source localization has become an increasingly important component of modern neuroimaging and brain–computer interface (BCI) research. Recent methodological developments have significantly improved the accuracy, computational efficiency, and three-dimensional reconstruction of cortical activity from scalp recordings. Each of the existing approaches, however, introduces its own trade-offs. Some methods prioritize computational efficiency but are primarily suited for offline analysis, while others can better track dynamic or distributed sources but require careful parameter tuning and increased computational resources. More recently, iterative and learning-based frameworks have emerged that combine classical regularization with adaptive learning strategies, allowing models to refine source estimates over time and improve robustness across datasets.

Another important direction in the field involves multimodal integration, where EEG is combined with complementary neuroimaging techniques such as fMRI, MEG, or fNIRS. These hybrid approaches exploit the high temporal resolution of EEG alongside the superior spatial specificity of other modalities, leading to more comprehensive representations of brain activity. At the same time, the growing availability of large EEG datasets has emphasized the importance of standardized preprocessing pipelines and interoperable data formats to enable large-scale analyses and reproducible research. Future methodological progress is therefore expected to focus on dynamic source reconstruction, higher-dimensional forward modeling, improved optimization

strategies, and hybrid machine-learning frameworks capable of extracting more physiologically meaningful information from EEG recordings.

Despite these advances, selecting an appropriate source localization method remains a challenging task. The performance of an inverse solution is strongly influenced by the spatial characteristics of the underlying neural sources. Methods that perform well for isolated or focal sources, such as those typically observed in simple evoked responses, may not generalize effectively to complex cognitive tasks involving multiple interacting or spatially distributed cortical regions. Consequently, researchers must develop a clear understanding of the EEG inverse problem and the inherent trade-offs between localization accuracy, spatial smoothness, and robustness to noise. In practice, this requires both theoretical insight into the properties of different inverse algorithms and empirical evaluation using the specific datasets under investigation.

In the context of the present study, several limitations should be explicitly acknowledged. First, the sample size was relatively modest (n = 19), which may limit statistical power and restrict the extent to which the findings can be generalized to broader or more heterogeneous populations. Although this sample size is comparable to many EEG source localization studies, larger cohorts would allow for more robust estimation of inter subject variability and more stable statistical inference. Second, the experimental design was restricted to visual paradigms involving perception and imagination tasks. While visual networks provide a well characterized and anatomically organized testbed for evaluating localization accuracy, the performance characteristics of the investigated inverse solutions may differ in other functional domains such as auditory processing, language, or motor control. Therefore, the generalizability of the reported findings beyond visual cortical systems should be interpreted with caution. Future studies should extend this evaluation

framework to larger participant samples and multimodal cognitive paradigms to assess the stability of these conclusions across diverse brain networks and task conditions.

In parallel with algorithmic development, deep learning (DL) has emerged as a promising framework for advancing EEG source reconstruction beyond traditional regularization-based solutions. A major research focus involves reducing the gap between simulated training environments and real EEG recordings, enabling models trained on synthetic data to generalize reliably to real-world scenarios. Physics-informed neural networks represent an important step in this direction, embedding domain knowledge directly into model architectures to preserve physiologically meaningful constraints. For example, hybrid networks incorporating attention mechanisms, convolutional layers, and recurrent units have been proposed to enforce spatial smoothness and frequency-domain consistency during reconstruction (Zhang et al., 2024). Deep learning is also being explored as a complementary tool rather than a direct replacement for classical inverse solvers. Studies have shown that applying convolutional neural networks to features derived from traditional source estimates can significantly improve performance in downstream tasks such as cognitive state classification (Kaviri & Vinjamuri, 2025). Additionally, DL techniques are improving essential preprocessing stages, including automated electrode localization and sensor-to-MRI registration (Pinte et al., 2021). Nevertheless, several open challenges remain, including reliable uncertainty estimation in learned inverse solutions, integration of multimodal priors from MRI or MEG, and the development of end-to-end frameworks capable of robust cross-subject generalization.

Progress in EEG source localization is also fundamentally limited by numerical challenges associated with solving the forward problem, which forms the basis of all inverse solutions. Accurate forward modeling, often implemented through BEM or FEM formulations, requires

precise knowledge of head tissue conductivities, anatomical geometry, and sensor alignment. Errors in these components can propagate through the inverse solution and substantially degrade localization accuracy. Such errors can be broadly categorized into several types: parameter errors caused by inaccurate tissue conductivity estimates, algebraic errors arising from numerical approximations in the mathematical formulation, iteration errors associated with incomplete convergence of iterative solvers, approximation errors due to simplified head models or coarse mesh discretization, and round-off errors resulting from finite numerical precision. Addressing these issues is essential for improving the reliability of EEG source reconstruction. Future work will therefore require more accurate subject-specific head models, improved mesh generation techniques, and optimized numerical solvers that minimize approximation and convergence errors.

The importance of these developments becomes particularly evident in the context of brain–computer interfaces. Source localization provides a mechanism for transforming noisy scalp EEG signals into physiologically interpretable representations of cortical activity, which can significantly improve feature extraction and classification performance in BCI paradigms such as event-related potentials (ERP), motor imagery (MI), and steady-state visually evoked potentials (SSVEP). However, integrating source imaging into real-time BCI systems introduces additional challenges. Accurate forward modeling often requires individualized anatomical data, while many inverse algorithms are computationally demanding and therefore difficult to implement in time-critical applications. This creates a fundamental trade-off between localization accuracy and the real-time processing requirements of operational BCI systems. Consequently, an important direction for future research is the development of computationally efficient source localization techniques capable of operating in real time. Potential solutions include sparse inverse methods such as TF-MxNE and GTF-MxNE, adaptive spatial filters like LCMV beamformers, and deep

learning architectures designed to approximate the inverse mapping with reduced computational overhead. Ultimately, the next generation of BCI systems will likely depend on the seamless integration of accurate source reconstruction with machine learning algorithms, enabling stable, interpretable, and high-performance decoding of brain activity in real-world environments.

Acknowledgment

The authors would like to sincerely thank Yeganeh Seyfi and Yasaman Seyfi for their valuable contributions during the review process. We greatly appreciate their time and support.

Declaration of interests: We have nothing to declare.

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Author Contributions: Methodology, Golnaz Baghdadi, Mohammad Ahadzadeh; Literature review and resources, Mohammad Ahadzadeh, Amir Hosein Ardalan, Reyhaneh Amiri, Arefeh Rezagholi, Golnaz Baghdadi ; Writing – Original Draft, Golnaz Baghdadi, Mohammad Ahadzadeh; Supervision, Golnaz Baghdadi

Declaration of generative AI and AI-assisted technologies in the writing process: During the preparation of this work the author(s) used ChatGPT in order to improve clarity, readability, and grammatical correctness. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the published article.

References

Baillet, S., Mosher, J. C., & Leahy, R. M. (2002). Electromagnetic brain mapping. *IEEE Signal processing magazine*, *18*(6), 14-30.

Brinkmann, B. H. (2024). Technical Considerations in EEG Source Imaging. *Journal of Clinical Neurophysiology*, *41*(1), 2-7.

Buzsáki, G., Anastassiou, C. A., & Koch, C. (2012). The origin of extracellular fields and currents—EEG, ECoG, LFP and spikes. *Nature reviews neuroscience*, *13*(6), 407-420.

Cuffin, B. N. (2002). EEG localization accuracy improvements using realistically shaped head models. *IEEE Transactions on Biomedical Engineering*, *43*(3), 299-303.

da Silva, F. L. (2013). EEG and MEG: relevance to neuroscience. *Neuron*, *80*(5), 1112-1128.

Dale, A. M., & Sereno, M. I. (1993). Improved localizadon of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: a linear approach. *Journal of cognitive neuroscience*, *5*(2), 162-176.

Fuchs, M., Wagner, M., & Kastner, J. (2007). Development of volume conductor and source models to localize epileptic foci. *Journal of Clinical Neurophysiology*, *24*(2), 101-119.

Grech, R., Cassar, T., Muscat, J., Camilleri, K. P., Fabri, S. G., Zervakis, M., Xanthopoulos, P., Sakkalis, V., & Vanrumste, B. (2008). Review on solving the inverse problem in EEG source analysis. *Journal of neuroengineering and rehabilitation*, *5*(1), 25.

Hämäläinen, M., Hari, R., Ilmoniemi, R. J., Knuutila, J., & Lounasmaa, O. V. (1993). Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of modern Physics*, *65*(2), 413.

Hauk, O., Stenroos, M., & Treder, M. S. (2022). Towards an objective evaluation of EEG/MEG source estimation methods – The linear approach. *NeuroImage*, *255*, 119177. https://doi.org/10.1016/j.neuroimage.2022.119177

Jatoi, M. A., & Kamel, N. (2017). *Brain source localization using EEG signal analysis*. CRC Press.

Jatoi, M. A., Kamel, N., Malik, A. S., Faye, I., & Begum, T. (2014). A survey of methods used for source localization using EEG signals. *Biomedical Signal Processing and Control*, *11*, 42-52. https://doi.org/https://doi.org/10.1016/j.bspc.2014.01.009

Kaviri, S. M., & Vinjamuri, R. (2025). Decoding motor execution and motor imagery from EEG with deep learning and source localization. *Biomedical Engineering Advances*, *9*, 100156. https://doi.org/https://doi.org/10.1016/j.bea.2025.100156

Liu, C., Downey, R. J., Mu, Y., Richer, N., Hwang, J., Shah, V. A., Sato, S. D., Clark, D. J., Hass, C. J., & Manini, T. M. (2023). Comparison of EEG source localization using simplified and anatomically accurate head models in younger and older adults. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *31*, 2591-2602.

Michel, C. M., & Brunet, D. (2019). EEG Source Imaging: A Practical Review of the Analysis Steps. *Front Neurol*, *10*, 325. https://doi.org/10.3389/fneur.2019.00325

Michel, C. M., Murray, M. M., Lantz, G., Gonzalez, S., Spinelli, L., & Grave de Peralta, R. (2004). EEG source imaging. *Clinical Neurophysiology*, *115*(10), 2195-2222. https://doi.org/https://doi.org/10.1016/j.clinph.2004.06.001

Montes-Restrepo, V., van Mierlo, P., Strobbe, G., Staelens, S., Vandenberghe, S., & Hallez, H. (2014). Influence of skull modeling approaches on EEG source localization. *Brain topography*, *27*(1), 95-111.

Niedermeyer, E., & da Silva, F. L. (2005). *Electroencephalography: basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins.

Nunez, P. L., & Srinivasan, R. (2006). *Electric fields of the brain: the neurophysics of EEG*. Oxford university press.

Nuñez Ponasso, G., Wartman, W. A., McSweeney, R. C., Lai, P., Haueisen, J., Maess, B., Knösche, T. R., Weise, K., Noetscher, G. M., & Raij, T. (2024). Improving EEG forward modeling using high-resolution five-layer BEM-FMM head models: effect on source reconstruction accuracy. *Bioengineering*, *11*(11), 1071.

Pascual-Marqui, R. D., Michel, C. M., & Lehmann, D. (1994). Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. *International Journal of psychophysiology*, *18*(1), 49-65.

Pascucci, D., Tourbier, S., Rué-Queralt, J., Carboni, M., Hagmann, P., & Plomp, G. (2022). Source imaging of high-density visual evoked potentials with multi-scale brain parcellations and connectomes. *Sci Data*, *9*(1), 9. https://doi.org/10.1038/s41597-021-01116-1

Pinte, C., Fleury, M., & Maurel, P. (2021). Deep Learning-Based Localization of EEG Electrodes Within MRI Acquisitions. *Front Neurol*, *12*, 644278. https://doi.org/10.3389/fneur.2021.644278

Ramon, C., Schimpf, P. H., & Haueisen, J. (2006). Influence of head models on EEG simulations and inverse source localizations. *Biomedical engineering online*, *5*(1), 10.

Vanrumste, B., Van Hoey, G., Van de Walle, R., Van Hese, P., D'Have, M., Boon, P., & Lemahieu, I. (2001). The realistic versus the spherical head model in EEG dipole source analysis in the presence of noise. 2001 Conference Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society,

Vatta, F., Meneghini, F., Esposito, F., Mininel, S., & Di Salle, F. (2010). Realistic and spherical head modeling for EEG forward problem solution: a comparative cortex-based analysis. *Computational intelligence and neuroscience*, *2010*(1), 972060.

Yokoyama, H., Kaneko, N., Usuda, N., Kato, T., Khoo, H. M., Fukuma, R., Oshino, S., Tani, N., Kishima, H., & Yanagisawa, T. (2024). M/EEG source localization for both subcortical and cortical sources using a convolutional neural network with a realistic head conductivity model. *APL bioengineering*, *8*(4).

Zhang, B., Li, D., & Wang, D. (2024). DCT based multi-head attention-BiGRU model for EEG source location. *Biomedical Signal Processing and Control*, *93*, 106171. https://doi.org/https://doi.org/10.1016/j.bspc.2024.106171