DTU Library

Evaluating the influence of anatomical accuracy and electrode positions on EEG forward solutions

Nielsen, Jesper Duemose; Puonti, Oula; Xue, Rong; Thielscher, Axel; Madsen, Kristoffer Hougaard

Link back to DTU Orbit

Evaluating the influence of anatomical accuracy and electrode positions on EEG forward solutions

Jesper Duemose Nielsen [a,b,c,*], Oula Puonti [b], Rong Xue [g,e,f], Axel Thielscher [d,b], Kristoffer Hougaard Madsen [a,b]

[a] Department of Applied Mathematics and Computer Science, Technical University of Denmark, Denmark
[b] Danish Research Centre for Magnetic Resonance, Centre for Functional and Diagnostic Imaging and Research, Copenhagen University Hospital Amager and Hvidovre, Denmark
[c] Sino-Danish Centre for Education and Research, Aarhus, Denmark
[d] Department of Health Technology, Technical University of Denmark, Kgs. Lyngby, Denmark
[e] State Key Laboratory of Brain and Cognitive Science, Beijing MRI Center for Brain Research, Institute of Biophysics, Chinese Academy of Sciences, Beijing, China
[f] Beijing Institute for Brain Disorders, Beijing, China
[g] University of Chinese Academic of Sciences, Beijing, China

ARTICLE INFO

ABSTRACT

Generating realistic volume conductor models for forward calculations in electroencephalography (EEG) is not trivial and several factors contribute to the accuracy of such models, two of which are its anatomical accuracy and the accuracy with which electrode positions are known. Here, we investigate effects of anatomical accuracy by comparing forward solutions from SimNIBS, a tool which allows state-of-the-art anatomical modeling, with well-established pipelines in MNE-Python and FieldTrip. We also compare different ways of specifying electrode locations when digitized positions are not available such as transformation of measured positions from standard space and transformation of a manufacturer layout.

Substantial effects of anatomical accuracy were seen throughout the entire brain both in terms of field topography and magnitude with SimNIBS generally being more accurate than the pipelines in MNE-Python and FieldTrip. Topographic and magnitude effects were particularly pronounced for MNE-Python which uses a three-layer boundary element method (BEM) model. We attribute these mainly to the coarse representation of the anatomy used in this model, in particular differences in skull and cerebrospinal fluid (CSF). Effects of electrode specification method were evident in occipital and posterior areas when using a transformed manufacturer layout whereas transforming measured positions from standard space generally resulted in smaller errors.

We suggest modeling the anatomy of the volume conductor as accurately possible and we hope to facilitate this by making it easy to export simulations from SimNIBS to MNE-Python and FieldTrip for further analysis. Likewise, if digitized electrode positions are not available, a set of measured positions on a standard head template may be preferable to those specified by the manufacturer.

1. Introduction

Electroencephalography (EEG) can be used to analyze brain activity with high temporal resolution. The measurements consist of potential differences observed on the scalp which directly reflect the synchronous activity of a large body of pyramidal cells oriented perpendicular to the cortical surface (Kirschstein and Köhling, 2009). However, the relationship between measurements and brain activity is complicated by the fact that the former is a spatially low-pass filtered representation of the latter where the filter consists of the tissues separating the neural sources from the sensors (Buzsáki et al., 2012). Due to volume conduction, all

sources in the brain affect the potential at all electrodes simultaneously (Baillet et al., 2001).

In some situations, it may be beneficial to analyze the data in the source domain rather than the sensor domain, e.g., when investigating the connectivity between different brain regions (Mahjoory et al., 2017; Nguyen-Danse et al., 2021; Schoffelen and Gross, 2009) or when localizing spike activity in epilepsy (Kaiboriboon et al., 2012; van Mierlo et al., 2020). Reconstructing the neural generators of an observed EEG signal is an inverse problem. In order to solve this, one first needs to solve the corresponding forward problem which consists of estimating the potential distribution on the scalp due to a neural source placed arbitrarily

---

in the brain (Baillet et al., 2001). Forward and inverse solvers are implemented in several open-source software packages for EEG analysis (Delorme and Makeig, 2004; Gramfort et al., 2014; Oostenveld et al., 2011; Tadel et al., 2011).

To solve the forward problem in EEG, one typically starts by constructing a physical model of the head, the volume conductor model, which enables simulation of volume conduction effects due to activity in the brain. Naturally, one would expect the accuracy of this volume conductor model to impact the accuracy with which such effects can be modeled (Vorwerk et al., 2014).

In this work, we investigate some of the elements which affect the accuracy of the forward solution, specifically, the anatomical accuracy of the head model and the accuracy with which electrode positions are known. We compare existing methods with SimNIBS which is able to generate state-of-the-art head models automatically based on magnetic resonance imaging (MRI) scans. Our aim is to explore the extent to which more accurate modeling leads to improvements in the forward solution and perhaps also in source localization accuracy (although the latter is not investigated here).

Numerous methods have been employed to solve the EEG forward problem (Hallez et al., 2007). In simple geometries, e.g., nested sphere models, (quasi-)analytical solutions can be derived, however, for realistically shaped models numerical methods are needed (Ary et al., 1981). In the BEM (Gramfort et al., 2010; Hämäläinen and Sarvas, 1989; de Munck, 1992; Oostendorp and van Oosterom, 1989), it is common to use a three-layer surface model (consisting of inner skull, outer skull, and skin) of relatively low resolution due to the dense nature of the problem. However, efficient ways of solving the system, e.g., the fast multipole method, exist allowing for high-resolution, multi-compartment BEM models as well (Makarov et al., 2021). In the finite element method (FEM), the whole volume is explicitly discretized. This allows anisotropy to be incorporated into the model and FEM is also capable of handling highly complex models (Schrader et al., 2021; Vorwerk et al., 2018; Wolters et al., 2004).

One important aspect of the forward model is its anatomical accuracy. In a systematic evaluation of the effect of distinguishing different tissue compartments, Vorwerk et al. (2014) concluded that modeling cerebrospinal fluid (CSF) as well as distinguishing white matter and gray matter were most important whereas the effects of modeling spongy bone and the anisotropic conductivity of white matter were found to be smaller (albeit still significant). Using a similar strategy, Azizollahi et al. (2016) corroborated these results in neonates. However, we note that other studies have argued for the importance of including anisotropy in the model (Haueisen et al., 2002; Marin et al., 1998; Wolters et al., 2006).

The skull compartment has received special attention in the literature due to its importance in shaping the observed fields in EEG (Hämäläinen et al., 1993). Lanfer et al. (2012) investigated the effect of several geometrical simplifications and errors on the forward solution, concluding that localized modeling errors (e.g., skull holes, erroneous thickness) result in forward errors mostly in the vicinity of such geometrical inaccuracies whereas simplifications of a more general nature (e.g., cutting the model at the base of the skull or approximating the base of the skull with constant thickness) show increased errors for large array of positions. Similar findings were reported by Chauveau et al. (2004) and Li et al. (2007). Forward modeling errors have also been observed when not modeling skull openings (Fiederer et al., 2016). Errors in skull thickness were also investigated by Chauveau et al. (2004), who found sources close to the skull to be most affected, however, their results also demonstrated that errors were seen predominantly on magnitudes and not topography. Several of these studies also report increased dipole localization errors in the vicinity of such forward modeling errors (Chauveau et al., 2004; Fiederer et al., 2016; Lanfer et al., 2012) suggesting that these effects indeed translate to errors in source localization in EEG.

To construct anatomically accurate volume conductor models, images from one or more structural sequences (e.g., T1- and T2-weighted MRI or in rare cases even computed tomography (CT)) are used. However, they may not always be available. To reconstruct sources without such anatomical information, one will have to rely on some kind of average anatomy. This was investigated by Acar and Makeig (2013) who compared individualized head models to models based on template anatomy and found median source localization errors of about 5 mm, particularly in the inferior part of the brain, since the template model was cut at the base of the skull.

Taken together, it seems that the anatomical accuracy of the forward model plays an important role in shaping not only the forward solution but likely also the final source reconstruction. FEM allows us to specify geometrically complex models, however, realistic modeling of the head geometry is difficult and has so far not been easily accessible to the EEG community. SimNIBS (Thielscher et al., 2015) is a tool for simulating electrical fields in the brain due to noninvasive brain stimulation which can construct realistic anatomical volume conductor models of the human head with reasonable accuracy from MRI scans as validated against manual segmentations based on MRI and CT (Nielsen et al., 2018; Puonti et al., 2020b). Given the intimate relationship between the EEG forward problem and the problem of simulating the effect of transcranial electrical stimulation (TES)—related by reciprocity (Ruffini, 2015; Wolters et al., 2004)—SimNIBS can also be used to solve the EEG forward problem and we explore this in further detail in the current work.

Another important aspect of the forward model is how electrode positions are estimated and coregistered to the anatomical model. Different systems to acquire electrode positions exist (e.g., the Polhemus Fastrak system, https://polhemus.com), however, if such systems are not available, one may have to use a template description of the electrode positions. To this end, EEG cap manufacturers typically provide spherical angles of the different electrodes which can then be mapped onto a sphere of a radius corresponding to the size of the subject head.

Dalal et al. (2014) found that localization accuracy and output signal-to-noise ratio (SNR) of a linearly constrained minimum variance beamformer was impacted by electrode digitization technique and coregistration method. Another study by Wang and Gotman (2001) found only minimal effect of electrode position errors while simulating displacements individually for each electrode. This may be a realistic model of pure measurement errors in a scenario where positions are digitized. However, errors due to coregistration between MRI scans and digitized electrode positions or the use of template positions, might be expected to be more spatially correlated, thus introducing a general bias in electrode positions (e.g., due to tilting, stretching etc.) (Acar and Makeig, 2013; Homölle and Oostenveld, 2019). Homölle and Oostenveld (2019) showed that digitizing electrodes using a structured-light 3D scanner resulted in slightly smaller errors compared to transformation of a custom template which again performed better than simply using the manufacturer template positions. Errors were also evident in the forward solution and a subsequent dipole fit. The errors of the latter reached 30 mm over the parietal cortex for the manufacturer template coinciding with the largest spatial bias in electrode positions. Acar and Makeig (2013) simulated coregistration errors by tilting electrode positions and found increased source localization errors (on average 5 mm) affecting predominantly superficial sources. Thus, we expect the accuracy with which electrode positions are determined to affect not only the forward model but also subsequent source estimates.

Here we explicitly consider how these two aspects of the forward model, namely its anatomical accuracy as well as how electrode positions are estimated, affect the forward calculations. The geometrical specification of the model will, to a large extent, be determined by the particular choice of forward modeling pipeline. Consequently, we first compare pipelines from different software packages for generating such models. This will result in aggregated effects from various sources but also provide practically relevant insights into the performance of the tested pipelines. Specification of electrode locations will depend on the

available equipment and resources. Since it may not always be possible to obtain individual estimates during the experimental session, one may have to resort to a set of template positions. Thus, we compare different template layouts to determine their effect on the final solution in the second part of the study.

Finally, although not the focus here, we would like to point out that several other parameters of the forward model are important as well, such as the influence of the conductivity profile of different tissues. The conductivity of some tissues in the human head are fairly well known (e.g., that of CSF) (Baumann et al., 1997) whereas that of others (e.g., compact and spongy bone) (Aydin et al., 2014) generally show large variability across studies. Even though this is known to affect modeling accuracy (Bangera et al., 2010; Dannhauer et al., 2011; Marin et al., 1998; McCann and Beltrachini, 2022; Vallaghe and Clerc, 2009; Vorwerk et al., 2019), one will typically have to resort to standard values from the literature although calibration paradigms have also been suggested with (Aydin et al., 2014) or without (Acar et al., 2016) the use of magnetoencephalography. Finally, model resolution is also important (Kaipio and Somersalo, 2007). In most cases, though, researchers will probably rely on the default settings of a particular pipeline (as determined from validation experiments).

We start by evaluating the numerical accuracy of SimNIBS in a spherical model and a realistic three-layer head model. Next, we compare forward solutions from SimNIBS with standard pipelines in MNE-Python and FieldTrip[1]. We also include a model based on standard anatomy representing a scenario where electrode positions are known but the anatomy is not. As we will see, the pipelines differ substantially in terms of anatomical accuracy. Finally, we compare different ways of specifying electrode positions when digitized positions are not available such as transformation of positions from standard space and transformation of manufacturer layout positions. As in the first sub-study, we also include a model based on standard anatomy.

2. Methods

2.1. Calculating EEG forward solutions in SimNIBS

Starting from the quasi-static approximation of the Maxwell equations, the EEG forward problem takes the form

$$-\nabla \cdot (\sigma \nabla u) = -\nabla \cdot \mathbf{j}_s = f \tag{1}$$

subject to boundary conditions

$$u_i = u_j \quad \text{in } \Omega \tag{2}$$

$$\langle \mathbf{j}_i, \mathbf{n} \rangle = 0 \quad \text{on } \partial\Omega \tag{3}$$

where i and j are two neighboring elements (e.g., tetrahedra) and $\mathbf{n}$ is the vector normal to their interface. $\langle \mathbf{j}, \mathbf{n} \rangle$ is the projection of the current density, $\mathbf{j}$, in the normal direction, and Ω and $\partial\Omega$ refer to the interior of the domain and its surface, respectively. Equation (2) ensures that the solution is continuous across interfaces of neighboring elements whereas equation (3) states that no current can leave the domain. We are interested in finding u, the potential distribution, for a given f (assumed to be known), which can be interpreted as a monopolar source configuration, i.e., the source of the current density, $\mathbf{j}_s$ (Hallez et al., 2007).

Equation (1) states that the divergence of the return current density (left-hand side) is equal to the divergence of the source current density. This is important for EEG source analysis as we measure potential differences on the scalp due to the return currents induced by the primary currents (i.e., the neural sources) along with other effects of non-interest (e.g., respiration, heartbeat).

Due to the principle of reciprocity, the EEG forward problem is related to the estimation of electric fields induced by noninvasive brain stimulation. Specifically, the potential between two points, a and b, due to a dipolar source at a certain position is related to the electric field at that position resulting from running a current between a and b (Weinstein et al., 2000). SimNIBS is a software package for solving the latter problem. To generate so-called leadfields for EEG, we exploit the intimate relationship between the EEG forward problem (estimating scalp potentials from current sources in the brain) and TES (estimating the electric field resulting from applying a potential difference between two electrodes or, equivalently, applying a current at one electrode and removing it at another). SimNIBS returns a smooth estimate of the electric field from which the gain (or leadfield) matrix can be computed. We describe how this is achieved below.

In SimNIBS, the volume conductor model is constructed using tetrahedral elements (potentially with anisotropic conductivities) and linear basis functions are employed to model the potential. The electric field is estimated by differentiating the solution and nodal values are recovered separately for each tissue compartment using superconvergent patch recovery (SPR) (Zienkiewicz and Zhu, 1992). In EEG, the sources are located in the gray matter and we therefore interpolate the field on a surface representing the center of the gray matter sheet.

Having solved the equation system for each electrode wrt. the reference, we start by computing the (negative) gradient of the solution to obtain the electric field, $\mathbf{e}_i$, in each element, i,

$$\mathbf{e}_i = -\mathbf{D}_i \mathbf{u}_i. \tag{4}$$

where $\mathbf{D}_i$ is the gradient operator and $\mathbf{u}_i$ contains the potentials at the nodes of i. As we use linear basis functions, this field estimate is not continuous and therefore very inaccurate for locations other than the element barycenters. To recover nodal field values, we use SPR to construct a smooth, interpolating function for each node by fitting a linear model to the values in an element patch (i.e., the collection of all elements associated with a given node) around each node such that a field component at an arbitrary position, r, can be estimated as

$$e(r) = \mathbf{x}^\top \mathbf{b} = (1 \; x \; y \; z) \, \mathbf{b} \tag{5}$$

where (x, y, z) are the coordinates at position r and $\mathbf{b}$ is a vector of parameters to be estimated. This is solved by minimizing the least squares error between the values of the function $\mathbf{f}$ which we are trying to fit (here the values of the electric field) and the interpolant at the barycenters

$$\arg\min_{\mathbf{b}} \|\mathbf{f} - \mathbf{Xb}\|^2. \tag{6}$$

Here, $\mathbf{X}$ is constructed by row-wise stacking of components $\mathbf{x}_i^\top$ corresponding to the barycenter of each element in the patch (augmented by a column of ones as above). We can compute the projection matrix at r as $\mathbf{x}^\top (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top$ and apply it to each component of the electric field to interpolate the field at this location (which, in this case, is the coordinates of the node defining the element patch). Finally, the electric field at any location can be estimated using barycentric interpolation (see supplementary material of Saturnino et al., 2019, for details).

Following Ruffini (2015)[2], if we consider a particular location in the brain, r, and a pair of electrodes, a and b, on the scalp, we have

$$u_{ab}^{\text{EEG}} I_{ab}^{\text{TES}} = -\mathbf{e}^{\text{TES}}(r)^\top \mathbf{j}^{\text{EEG}}(r) \tag{7}$$

$$u_{ab}^{\text{EEG}} = -\frac{\mathbf{e}^{\text{TES}}(r)^\top}{I_{ab}^{\text{TES}}} \mathbf{j}^{\text{EEG}}(r) \tag{8}$$

$$u_{ab}^{\text{EEG}} = \mathbf{G}_{ab}(r) \mathbf{j}^{\text{EEG}}(r) \tag{9}$$

where $\mathbf{e}(r)$ is the electric field induced by TES at position, r, using current I, and u is the potential difference at a wrt. b due to a dipole source $\mathbf{j}(r)$.

---

[1] Here, we restrict the comparison to MNE-Python and FieldTrip although other popular software packages also exist, e.g., Brainstorm and EEGLab.

[2] Instead of a current density, we simply consider a point dipole with unit A m and no associated (or infinitesimally small) volume and thus we have dropped the δV term.

Table 1
Resolution of surface meshes and number of nodes and elements in the corresponding volume meshes used in the validation experiments. The number and "C" in parenthesis refers to the number of compartments in the model.

Surface	Volume	
nodes/mm^2	Nodes	Tetrahedra
Spherical models (3C)		
0.065	24,880	126,844
0.125	61,154	322,872
0.25	133,952	710,780
0.5	356,742	1,934,983
1.0	903,074	4,977,163
Realistic models (3C)		
low	22,969	120,130
high	697,793	3,861,891
Realistic models (5C)		
0.5	638,247	3,530,439
0.5 (refined)	4,880,956	28,243,512
1.0	1,955,900	11,171,938
1.0 (refined)	15,228,577	89,375,504

$\mathbf{G}_{ab}(r)$ is the potential field value at *a* wrt. *b* for unit dipoles at *r* aligned with each principal axis. Collecting these for all locations and electrode pairs and augmenting it with the reference electrode (effectively, adding zeros corresponding to the reference electrode, and applying an average reference), we arrive at the final matrix of leadfields, i.e., a matrix describing potential fields induced by (ideal) dipolar sources in the brain. Efficient implementations based on reciprocity have also been presented by others (Weinstein et al., 2000; Wolters et al., 2004).

2.2. Validation of SimNIBS for EEG

We validate the leadfields generated using SimNIBS in two ways. First, we compare with an analytical solution in a simple geometry consisting of nested spheres. Second, using a realistically shaped three-compartment model, we compare with existing BEM implementations as BEM generally has high numerical accuracy compared to first order FEM. However, we also wanted to put these numerical differences into perspective by showing differences due to (not) modeling the CSF compartment (as has also been done before, e.g., by Vorwerk et al. (2012)). To this end, we compare a five compartment FEM model with a three-compartment BEM model.

2.2.1. Spherical model

We constructed a model consisting of three concentric spheres representing the inner skull, outer skull, and outer skin surfaces with radii 80, 86, and 92 mm (Saturnino et al., 2019; Vorwerk et al., 2012) and conductivities 0.3, 0.006, 0.3 Sm^{-1}, respectively. We generated such models with different node densities, 0.065, 0.125, 0.25, 0.5, and 1.0 nodes/mm^2 (see Table 1 for information on number of nodes and elements in the volume meshes), to investigate the effect of resolution on the numerical accuracy and used Gmsh (Geuzaine and Remacle, 2009) to generate a volume mesh representation of each of the spherical models. Following Vorwerk et al. (2012), we placed point sources aligned with each axis (*x*, *y*, and *z*) in steps of 1 mm ranging from 2 to 77 mm to the sphere origin. This was done for 1000 random directions. Thus, all sources resided within the innermost compartment of the model. The outermost sources were at a distance of 15mm to the skin surface which is similar to the closest sources in a realistic head model (Lu et al., 2019). Finally, we placed 100 electrodes equidistantly on the outermost surface. In the simulations, these were modeled as rings with 10mm radius and 4mm thickness and were meshed onto the spherical models whereas in the analytical solutions they were treated as point electrodes[3]. The

conductivity of the electrodes was set to 29.4 Sm^{-1}. Analytical potentials were computed at the location of the electrodes using equations 2 and 2a from Ary et al. (1981).

2.2.2. Realistic geometry

To validate the models in a realistically shaped geometry, we extracted surfaces corresponding to inner skull, outer skull, and skin of a head model generated using the *headreco* pipeline in SimNIBS. We downsampled each surface to approximately 5000 nodes and used these in the simulations. Additionally, we upsampled these surfaces by uniform remeshing using MeshFix (Attene, 2010) to a resolution close to what we would normally use in our FEM calculations (approximately 80,000 nodes). Running this model in BEM was not feasible with the implementations used here due to the excessive computational resources needed. Conductivities were the same as in the three-layer sphere model. Electrodes were modeled as described above and were placed by creating a standard montage of the 64 channel BioSemi cap (https://www.biosemi.com) included in MNE-Python. For comparison, we used BEM implementations from FieldTrip (specifically, the *dipoli* implementation) and MNE-Python, the latter of which was used as a reference.

We also show numerical differences between FEM models with different surface densities (specifically, 0.5 and 1.0 nodes/mm^2) and refined (achieved by splitting the tetrahedra) versions of these models to illustrate the effect of geometrical accuracy compared to that of resolution. See Table 1 for information on number of nodes and elements in the volume meshes.

Lastly, we show results of a comparison between a five-compartment FEM model and a three-layer BEM model. This allows us to compare the numerical differences between FEM and BEM (from the previous simulation) with the differences resulting from modeling the anatomy at different levels of geometrical accuracy and detail. Here, the FEM model was used as reference.

2.3. Impact of anatomical accuracy on the forward solution

The dataset used in this study consists of 20 subjects each of which had a T1- and a T2-weighted MRI scan (for details on this dataset, please see Farcito et al., 2019). Importantly, manual segmentations of 16 tissue classes were available for all subjects. However, in this study we only distinguished 9 tissues[4] (10 including air pockets). An overview of this study and the one presented in Section 2.4 is provided in Table 2.

2.3.1. Volume conductor models

To study the effect of anatomical accuracy we generated four head models using different approaches, specifically, (1) SimNIBS with complete head anatomy reconstruction method (CHARM) (a new segmentation and meshing pipeline which is available in SimNIBS 4.0) (Puonti et al., 2020b), (2) MNE-Python (Gramfort, 2013) which is based on FreeSurfer (Dale et al., 1999), (3) FieldTrip (Oostenveld et al., 2011) which uses SPM12 (https://www.fil.ion.ucl.ac.uk/spm/software/spm12/) for segmentation and DUNEuro (Schrader et al., 2021) for FEM modeling, and (4) SimNIBS with CHARM using standard anatomy. As a reference, we used the manual segmentations which we meshed using SimNIBS[5]. Here we briefly describe each of the pipelines.

---

[3] Using point electrodes (as is common in EEG forward calculations) did not have a strong influence and so here we present the results for meshed electrodes which is the default approach in SimNIBS.

[4] White matter, gray matter, CSF, spongy bone, compact bone, skin, eyes, veins, muscle.

[5] Upon visual inspection, many sulci were closed in the manual segmentations, hence these were "corrected" using FreeSurfer by (1) enforcing gray matter within the cortical gray matter mask from FreeSurfer, (2) relabeling any cortical gray matter voxels outside of this mask which overlapped with the white matter mask from FreeSurfer to white matter, and (3) relabeling any unlabeled voxels to CSF. This seemed to work well for opening sulci, however, at the price of slightly reduced accuracy around cerebellum.

Table 2

Overview of the two sub-studies. Letters in the source space column do not carry any meaning *per se* but are only meant to indicate which models use the same (or different) source space definitions *within* each study. (R) denotes that this is the reference model. MNI denotes that the corresponding image is an MNI152 template image. Man. seg. denotes manual segmentation.

Model	Software	Input	Source Space	Electrode Locations	No. tissues	Dataset
Anatomy Study						
Manual (R)	SimNIBS*	Man. seg.	a	Custom	9	20 subjects from
SimNIBS-CHARM	SimNIBS	T1, T2	b	Custom	9	Farcito et al. (2019)
FieldTrip-SPM	FieldTrip	T1	a	Custom	5	
MNE-FS	MNE-Python	T1	a	Custom	3	
MNI-Template	SimNIBS	T1 (MNI)	c	Custom	9	
SimNIBS-CHARM (T1 only)	SimNIBS	T1	d	Custom	9	
Electrode Study						
Digitized (R)	SimNIBS	T1, T2	a	Digitized	9	32 subjects from
Custom-Template	SimNIBS	T1, T2	a	Custom	9	Madsen et al. (2019)
Man-Template	SimNIBS	T1, T2	a	Manufacturer	9	and
MNI-Digitized	SimNIBS	T1 (MNI)	b	Digitized	9	Karabanov et al. (2021)

* Custom processing pipeline.

Table 3

Resolution of models in the anatomy study and the electrode study. Number of nodes and elements are reported as mean and standard deviation over all subjects in each study. The standard deviation for MNI-Template is low because a subject uses one of four different models (from different CV runs). std. denotes standard deviation.

Model	Nodes		Elements	
	Mean	Std.	Mean	Std.
Anatomy Study				
Reference	547,688	41,957	3,218,804	248,273
SimNIBS-CHARM	679,216	40,779	4,026,987	243,927
FieldTrip-SPM	598,710	56,011	571,700	54,818
MNE-FS	15,360	0	30,708	0
MNI-Template	514,931	3517	2,987,265	22,726
SimNIBS-CHARM (T1 only)	666,753	40,642	5,075,927	313,296
Electrode Study				
Individual*	658,013	59,032	3,893,451	349,907
MNI-Digitized	606,976	0	3,530,879	0

* All models based on individual anatomy share the same mesh.

SimNIBS-CHARM The CHARM pipeline uses the SAMSEG framework (Puonti et al., 2016) to segment one or more MRI scans into fifteen different tissue classes. Subsequently, the segmentation is upsampled to 0.5 mm^3 isotropic resolution and post-processed using morphological operations to smooth the segmentation and to ensure that the brain is surrounded by at least one voxel (0.5mm) thick layer of CSF. The central cortical surface is reconstructed using adapted versions of the corresponding CAT12 (Dahnke et al., 2013) functionalities. The pial surface is then estimated from the central cortical surface and used to improve the sulci representations in the original segmentation. Finally, a volume mesh representation is generated using GGAL (https://www.cgal.org) by directly meshing the updated segmentation. The size of the tetrahedral elements are locally adapted using sizing fields such that elements that are far from tissue borders are larger and elements that are close to tissue borders are smaller. The initial tetrahedral mesh is post-processed to improve the quality of the tissue surfaces. This is done by reassigning the tissue type of tetrahedra representing "spikes" in the surfaces. In addition, the surface nodes are smoothed using a Taubin approach (Taubin, 1995) which is adapted to ensure good quality of the tetrahedra connected to the surfaces. These steps are implemented as standard steps of the CHARM segmentation and meshing pipeline.

Please note that the dataset described in Section 2.3 was used to create the tissue priors used in CHARM (Puonti et al., 2020b), however, when running CHARM, we used priors from a four-fold cross-validation (CV) split such that a subject was never included in the prior used when creating its segmentation to circumvent biased accuracy estimates due to overfitting. The FEM simulations were run using standard settings except that the Parallel Direct Sparse Solver (PARDISO)[6] from the Intel Math Kernel Library (version 2022.0.1) was used to calculate multiple FEM solutions efficiently.

MNE-FS MNE-Python integrates with FreeSurfer and relies on its *recon-all* pipeline. Here we created the BEM surfaces (inner skull, outer skull, and skin) based only on a T1-weighted image[7]. This procedure uses the watershed algorithm in FreeSurfer. The final node count in each surface was 5,120. We note that since the inner and outer skull surfaces sometimes overlapped, we had to fix this in order to be able to run the BEM calculations. This was done using MeshFix by pushing the inner skull inwards to ensure a minimum distance of 2 mm between the inner and outer skull since the problem seemed to be that the inner skull surface extended too far outwards. Consequently, we view this as an improvement of the model. We used MNE-Python 0.24 and FreeSurfer 7.1.0.

FieldTrip-SPM FieldTrip[8] uses SPM12 to segment a T1-weighted image into five tissue classes (gray matter, white matter, CSF, bone, and skin). This segmentation is post-processed using morphological operations (e.g., smoothing) to, for example, ensure that the brain is surrounded by CSF which again is surrounded by bone. Based on this segmentation, a hexahedral mesh was created with a downsampling factor of two and a node shift of 0.3. Node shifts were applied to hexahedra at tissue interfaces in order to smooth the boundaries thus allowing us to model smooth interfaces (Camacho et al., 1997; Wolters et al., 2007). We used the DUNEuro solver with default settings.

MNI-Template This model is created simply by running CHARM on the MNI152 T1-weighted image. We padded the image such that the lower part of the face and neck was filled from the prior. Otherwise the procedure is exactly as described for SimNIBS-CHARM.

The purpose of this work is not to provide a thorough evaluation of the anatomical accuracy of each of these models (we refer to Nielsen et al. (2018) and Puonti et al. (2020b)), however, visual inspection revealed that the SimNIBS-CHARM models were generally more anatomically accurate than the other models (see Figs. 6 and 7 for two examples). To summarize, SimNIBS-CHARM is a first order FEM model

---

[6] For details on usage and implementation, see https://www.intel.com/content/www/us/en/develop/documentation/onemkl-developer-reference-c/top/sparse-solver-routines/onemkl-pardiso-parallel-direct-sparse-solver-iface.html

[7] MNE-Python also provides a tool to create surfaces from a T1-weighted image and a sequence of fast low angle shot (FLASH) images, however, since FLASH was not available for this dataset we decided to only use the T1-weighted image (we tried using the T2-weighted image instead but with disappointing results).

[8] We used the master branch from the GitHub repository (https://github.com/fieldtrip) at commit 4b50f71105f6bb250018b815bb628e6f3e9de2b6 (from 22.11.2021).

with generally high anatomical accuracy; FieldTrip-SPM is a FEM model with decent accuracy of brain tissue but simplified non-brain tissue (e.g., bone); MNE-FS is a BEM model with simplified anatomy (no distinction between white matter, gray matter, and CSF) and limited accuracy; MNI-Template is a FEM model based on average anatomy which, in general, is quite dissimilar to the individual anatomy but may estimate CSF and bone reasonably well.

In each case, we used the default conductivities specified by the particular pipeline[9].

We did not do a formal comparison of computation times, however, we provide timings for a single subject in section S4 in the supplementary material. Overall, computation times were comparable for all methods.

2.3.2. Electrode positions

To model electrode positions we used an EasyCap BC-TMS64-X21 transcranial magnetic stimulation (TMS) compatible cap with a modified M10 (equidistant) layout (https://www.easycap.de). See Section 2.4.1 for details on how these positions were obtained.

We placed the electrodes on the individual reference models (i.e., the models based on the manual segmentations) by registering the template positions (in MNI space) to subject MRI space by a seven parameter affine registration (one uniform scaling parameter) based on coregistration of landmarks (nasion, left preauricular (LPA), right preauricular (RPA)). For SimNIBS-CHARM, MNE-FS, and FieldTrip-SPM, the electrodes were then projected onto the skin surface of the respective model. For the MNI-Template, the mesh was deformed nonlinearly using thin-plate splines (Darvas et al., 2006) to head points sampled around the electrodes (four points per electrode resulting in approximately 200 points in total) similar to the procedure used in Brainstorm[10] (Tadel et al., 2011). Given that the skin surface was well modeled in all pipelines, we do not anticipate major effects due to different electrode positions (in fact, the average projection distance was below 1 mm for all electrodes which we consider very low).

2.3.3. Source space

In the MNE-FS and FieldTrip-SPM models, we defined the source space as the central gray matter surface extracted from the reference model using CAT12 and downsampled to approximately 10,000 nodes per hemisphere. In the event that a source was not inside the brain (defined as the brain compartment in MNE-FS and white matter or gray matter in FieldTrip-SPM) this was excluded from analysis[11]. In SimNIBS-CHARM and MNI-Template, we used the central gray matter surface created during the respective runs. To compare the gain matrices (leadfields) we mapped them to *fsaverage* before computing the evaluation metrics. *fsaverage* is the vl template included in FreeSurfer to which the cortical sheet of individual subjects may be registered (e.g., facilitating group analysis) (Fischl et al., 1999).

2.4. Impact of electrode positions on the forward solution

The dataset used in this study is a subset of the dataset described in Madsen et al. (2019) and Karabanov et al. (2021) in which noninvasive brain stimulation was performed conditional on the brain state as

determined by EEG measurements. Of particular relevance to the current study, structural MRI scans were acquired and electrode positions digitized using an infrared optic stereo tracking system from Localite (https://localite.de). For this study, we extracted 32 subjects for which T1- and T2-weighted images as well as digitized electrode positions were available. During our analysis, we found that some digitization runs contained electrode positions which were clearly the result of some kind of equipment malfunction (e.g., a situation where an electrode was placed in the center of the head). To formalize the identification of such errors, and possibly "correct" them, we first transformed all electrode positions to standard space by nonlinear deformation. We then defined a "spurious digitization" as a position which was more than five centimeters away from the group median and we decided to correct a maximum of one electrode per subject. Corrections were performed by replacing the erroneous coordinates with the group median. Spurious digitization errors were detected in four of 32 subjects. In each of these subjects there was only a single error which we then corrected.

2.4.1. EEG Custom template generation

We generated a custom template in standard space of the EasyCap BC-TMS64-X21 EEG cap using the following procedure[12]. We reconstructed the skin surface of an augmented MNI152 template (which included the jaw and neck as well) from the final segmentation of the MNI-Template described in Section 2.3.1 using the marching cubes implementation from *scikit-image* (van der Walt et al., 2014) and smoothed the result using *PyVista* (Sullivan and Kaszynski, 2019). Since the MNI152 model is quite large (approximately 60 cm circumference of the head), we scaled it to a circumference of 57 cm which allowed us to comfortably fit a 58 cm cap on it. This scaled model was 3D printed on an Ultimaker 2 Extended +, the aforementioned cap was placed on it, and electrode positions were digitized using a Polaris Spectra infrared camera from NDI with custom software (https://www.ndigital.com). To securely fasten the sensor responsible for tracking the head position, we augmented and flattened the nasal bridge such that it could be easily attached here. The final template was scaled back to the original space of the MNI template.

To facilitate coregistration between the coordinate system of the digitizer and MNI space, we added 13 small dents (which served as "landmarks") in the surface (nasion, LPA, RPA, and ten points spread across the top of the head). We placed the EEG cap according to standard guidelines (e.g., Homölle and Oostenveld, 2019) and digitized the electrode positions as well as the landmarks. We repeated this three times, registered the electrode positions to MNI space and used the median position for the final template. In the final template we noticed a few irregularities. First, two parietal channels, 50 and 52, were not exactly equidistant to the surrounding electrodes. Upon inspection of the different cap sizes, we found that these channels were placed on one side of a stitching in caps smaller than 56 cm whereas they were on the opposite side in caps larger than or equal to 58 cm. Since this effect was evident mostly in the lateral direction, channels 50 and 52 were corrected by replacing the x coordinate with the mean of electrodes (49, 51) and (51, 53), respectively. Likewise, we saw variations in channel 22 between caps. Channel 22 was corrected by replacing the x and y coordinate with the means from 21 and 23. Finally, all channels were (re)projected onto the MNI skin surface. These steps were taken to improve the average fit of the template to subjects of different head sizes.

2.4.2. Volume conductor models

To study the effect of different ways of estimating electrode positions, we first constructed a model of the head for each subject using the CHARM pipeline in SimNIBS. To assess the errors in electrode positions

---

[9] In SimNIBS, the default conductivities (Sm^{-1}) are: 0.275 (gray matter), 0.126 (white matter), 1.654 (CSF), 0.008 (compact bone), 0.025 (spongy bone), 0.465 (scalp), 0.5 (eye balls), 0.6 (blood), 0.16 (muscle). In FieldTrip, we used conductivities from https://www.fieldtriptoolbox.org/tutorial/headmodel_eeg_fem : 0.33 (gray matter), 0.14 (white matter), 1.79 (CSF), 0.01 (skull), 0.43 (scalp). In the MNE-Python models, default conductivities are 0.3, 0.006, 0.3 for brain, skull, and scalp, respectively.

[10] See https://neuroimage.usc.edu/brainstorm/Tutorials/TutWarping for an example of how this can be done in Brainstorm.

[11] When sources were excluded, smoothing was employed to ensure that all nodes in *fsaverage* space were covered. To avoid unnecessary smoothing, this was only done for the nodes where it was required.

[12] This template will be available in SimNIBS and also from the following repository https://github.com/jduemose/eeg-templates. Over time, we plan to measure and add templates for other cap montages as well.

when using template layouts, we transformed the layout of the EEG cap as defined by the manufacturer[13] as well as our custom positions from standard space to subject space. The former was transformed using an affine transformation obtained by coregistering fiducials (nasion, LPA, and RPA) between subject space and the space of the manufacturer layout (after mapping the positions to a sphere). The latter was transformed in three different ways using (1) the nonlinear deformation field calculated by CHARM as part of the segmentation, (2) the affine registration between subject space and MNI space estimated by CHARM before performing the segmentation, and (3) the affine transformation obtained by coregistration of fiducials in subject space and MNI space. Subject specific fiducials were identified on the T1-weighted MRI of each subject. As reference, we used the digitized positions available in the data set.

We also included a model based on standard anatomy (similar to the one described in Section 2.3.1). As in the reference model, we used the digitized electrode positions from the data set. This was done to model a scenario where electrode positions have been accurately defined but structural information is unavailable.

As we did not find any appreciable differences in the accuracy of the channel positions between different ways of transforming the custom template (see Section 3.3), we chose to only use the nonlinearly deformed positions in the forward simulations as this does not require identification of landmark positions for each individual. Thus, we ran simulations using the models with the digitized positions, nonlinear deformation of custom template (Custom-Template), with affine registration of the manufacturer layout (Man-Template), and with standard anatomy (MNI-Digitized). All simulations were carried out in SimNIBS similarly to what was described for SimNIBS-CHARM in Section 2.3.1.

2.4.3. Source space

Source spaces were generated as described for SimNIBS-CHARM and MNI-Template in Section 2.3.3 for models based on individual and standard anatomy, respectively, and results mapped to the *fsaverage* template.

2.5. Evaluation metrics

To evaluate differences in electrode position, we used the Euclidean distances to the reference positions. We show the full distribution of these errors and also how the mean errors are distributed over channels. Additionally, we also show the spatial distribution of the absolute error along each axis (x, y, z) to explore the direction of the errors. For evaluation, we always used channel positions after projecting onto the skin surface.

All channel positions were evaluated in the head coordinate system used by MNE-Python. Here, the x-axis goes through the LPA to the RPA, the y-axis is normal to the x-axis and passes through the nasion, and the z-axis is normal to the xy-plane and points upwards according to the right-hand rule. Thus, this constitutes a right-anterior-superior (RAS) oriented coordinate system.

To evaluate the differences between forward models we use the relative difference measure (RDM)

$$\text{RDM} = \left\| \frac{u}{\|u\|} - \frac{u_{\text{ref}}}{\|u_{\text{ref}}\|} \right\| \qquad (10)$$

and the logarithm of the magnitude error (lnMAG)

$$\text{lnMAG} = \ln \frac{\|u\|}{\|u_{\text{ref}}\|} \qquad (11)$$

which computes the relevant measure of u with respect to u_{ref} and where $\|\cdot\|$ is the 2-norm (Meijs et al., 1989; Vorwerk et al., 2014). RDM measures the topographic error and is bounded between 0 and 2 where

0 means no difference whereas lnMAG measures the magnitude errors with a value of 0 corresponding to no difference. As u_{ref}, we used the model based on manual segmentations in the anatomy study and the model based on digitized electrode positions in the electrode study.

All forward solutions were re-referenced to an average reference and converted to a fixed orientation (normal to the cortical surface). The final gain matrix consisted of approximately 20,000 columns. Before computing the evaluation metrics, gain matrices were morphed to the fifth subdivision of the *fsaverage* template ("*fsaverage5*") which consists of 20,484 nodes in total (10,242 per hemisphere). This was done in order to facilitate comparison between models and subjects.

We show the cumulative relative frequency (CRF) of the error metrics using the data from all subjects and also the spatial distribution of the average of each metric. CRF is computed by sorting the values of interest and plotting these against their normalized indices. For example, a point on one of the curves in Fig. 1 indicates the fraction of sources (y-axis) with an RDM smaller than a certain value (x-axis). As such, the value at CRF = 0.5 is the median. All surface-based plots were created using *PyVista* 0.35.1 (Sullivan and Kaszynski, 2019). Remaining plots were created with *Matplotlib* 3.5.1 (Hunter, 2007).

3. Results

3.1. Validation of SimNIBS for EEG

3.1.1. Spherical model

Fig. 1 shows the CRF of RDM and lnMAG (calculated wrt. an analytical solution) for each model and Fig. 2 shows RDM and lnMAG as a function of source eccentricity. Both figures illustrate that increasing the resolution of the model is beneficial for numerical accuracy. Errors get larger with increasing source eccentricity. However, for the high density models, the RDM stays below 0.02 and the lnMAG between ±0.01. From Fig. 2, we see a small bias in the magnitude in that it is smaller in the SimNIBS simulations.

3.1.2. Realistic geometry

Fig. 3 shows CRF of RDM and lnMAG in a realistic three-layer model using the BEM solution from MNE-Python as reference. We see that errors of the FEM model with a resolution used in practice, i.e., the high resolution model, are generally low (below 0.05) except for a few outliers (heavy tail). The numerical accuracy generally deteriorates as sources get closer to the sensors (see section S1 and fig. S1.1 in the supplementary material). The BEM implementations are very similar in terms of topography but less so for magnitude.

Fig. 4 shows CRF of RDM and lnMAG for different model comparisons. We see that purely numerical errors (as evidenced by the comparison of a model with its refined counterpart, e.g., FEM 0.5 vs FEM 0.5r) are lower than errors due to modeling the anatomy in greater detail (i.e., FEM 0.5 vs. FEM 1.0) which again is much lower than the errors incurred by simplifications of the model (i.e., the full FEM models vs. three-layer BEM). An example of the effect of resolution and refinement on the geometry of the FEM models are illustrated in Fig. 5. Errors due to model simplifications are about ten times those of the errors due to numerical inaccuracies. The pattern is the same for both RDM and lnMAG. Contrary to the simple three-layer model evaluated above, source eccentricity does not seem to be as good a predictor of the overall errors (numerical and otherwise) in more complicated geometries (see section S1 and figs. S1.2 and S1.3 in the supplementary material).

3.2. Impact of anatomical accuracy on the forward solution

Figs. 6 and 7 show the surfaces extracted using each pipeline overlaid on the T1-weighted image for two subjects. It is not our intention to do a systematic assessment of the anatomical accuracy of the models as this has been done previously (Nielsen et al., 2018; Puonti et al., 2020b).
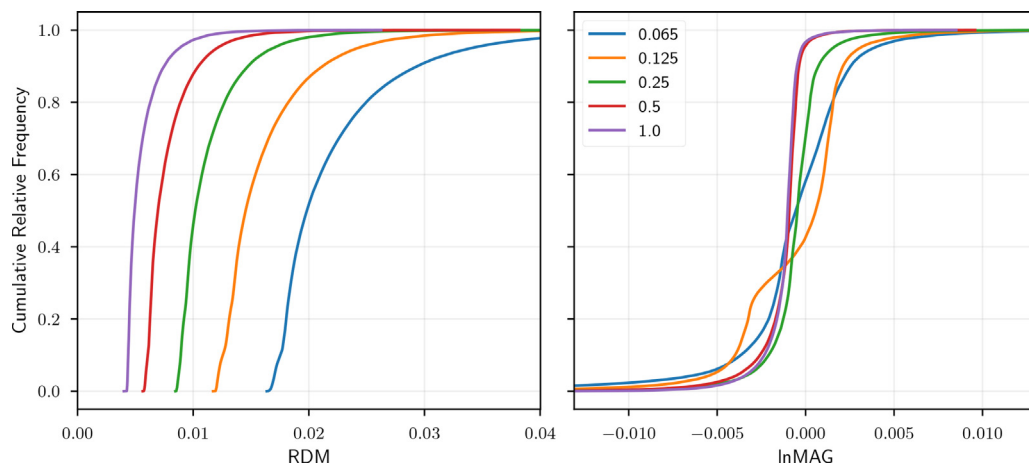
---

[13] The cap used here has an equidistant M10 layout and includes two additional channels in the temporal area. Here we use the manufacturer layout for this cap shipped with MNE-Python augmented by these two channels. To adhere to the equidistant properties of the layout, we defined the angles (theta, phi) of these electrodes as (-138, -22) and (138, 22).

Fig. 1. Spherical model. CRF of RDM and lnMAG for different model densities ($\mathrm{nodes/mm^2}$) in a three-layer sphere model. An analytical solution was used as reference.
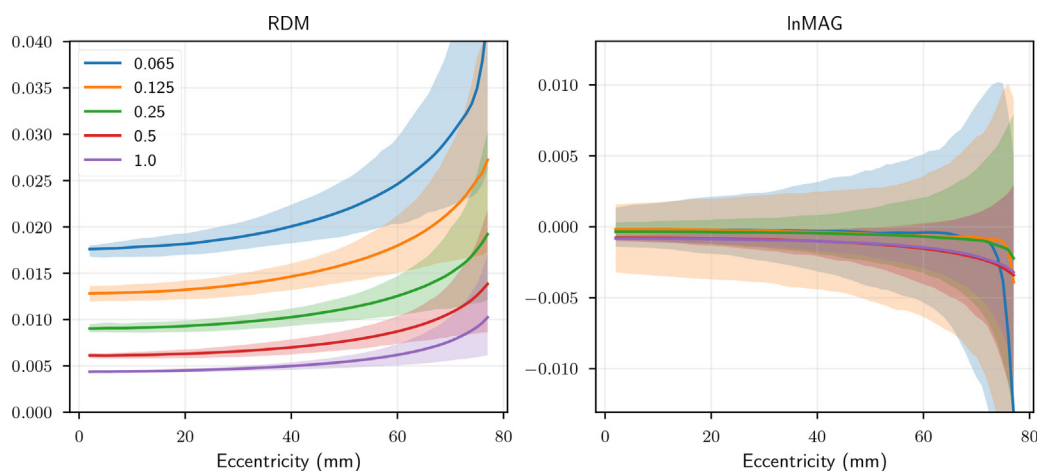

Fig. 2. Spherical model. RDM and lnMAG (wrt. an analytical solution) as a function of source eccentricity for different model densities in a three-layer sphere model. An analytical solution was used as reference. Shaded area denote the interval from the 0.5th to 99.5th percentile. Solid line denotes the mean.

However, we believe this to be representative of the differences between the pipelines we have seen in the data.

Generally, we found high anatomical accuracy of SimNIBS-CHARM whereas substantial simplifications of the modeling of the CSF and bone compartments are apparent in FieldTrip-SPM. As a result, FieldTrip-SPM is not particularly accurate for non-brain tissues. MNE-FS collapses white matter, gray matter, and CSF into one brain compartment. Additionally, we found that the BEM surfaces were not very accurate (as evidenced by the need for us to correct intersections between these surfaces). MNI-Template does not capture the individual structure of the brain tissues very well. However, in several cases it actually provided a decent fit of the skull compartment, whereas in others it was not accurate at all. We provide similar figures to Fig. 6 and 7 for more subjects in section S2 figs. S2.4 to S2.7 in the supplementary material.

Fig. 8 shows CRF of the error metrics, RDM and lnMAG, using the forward solution associated with the manual segmentations as reference. In terms of topographical errors, SimNIBS-CHARM shows best performance followed by FieldTrip-SPM, MNI-Template, and finally MNE-FS. As to the magnitude errors, SimNIBS-CHARM and MNI-Template are both centered around zero with the latter having a larger standard deviation than the former. FieldTrip-SPM and MNE-FS generally show decreased and increased sensitivity, respectively. The effect is most severe for MNE-FS both in terms of the mean and standard deviation.

Figs. 9 and 10 present the spatial distribution of the mean RDM and lnMAG on the *fsaverage* surface. From Fig. 9 we see slightly elevated

errors in the orbitofrontal area for SimNIBS-CHARM but otherwise low errors. The spatial pattern for FieldTrip-SPM is similar, however, the effects are more pronounced and also extends into the temporal cortex. MNE-FS shows high errors on the gyral crowns both for superior and deep sources. For MNI-Template we see mostly errors comparable to those of FieldTrip-SPM, but with a few hot spots and also slightly increased errors in parts of the parietal cortex. From Fig. 10 we see that SimNIBS-CHARM generally has the lowest overall errors. FieldTrip-SPM shows decreased sensitivity to deep, occipital, and temporal sources whereas MNE-FS shows increased sensitivity everywhere, particularly for shallow sources on the gyral crowns.

We expect the magnitudes to be sensitive to the ohmic tissue conductivities. Since we used the default values in each pipeline—and because the conductivities are the same in SimNIBS-CHARM and MNI-Template as in the reference model—it is not surprising that these are similar in terms of sensitivity. On the other hand, differences in conductivity is a potential confounder for the comparison of the FieldTrip-SPM and MNE-FS models to the reference. In an attempt to disentangle such effects, we created another reference model with conductivities similar to those of FieldTrip-SPM[14] and computed RDM and lnMAG. From Fig. 8, we see

---

[14] The conductivity of compact and spongy bone were both set to that of bone in the FieldTrip-SPM model and thus a minor geometrical simplification of the reference was introduced.
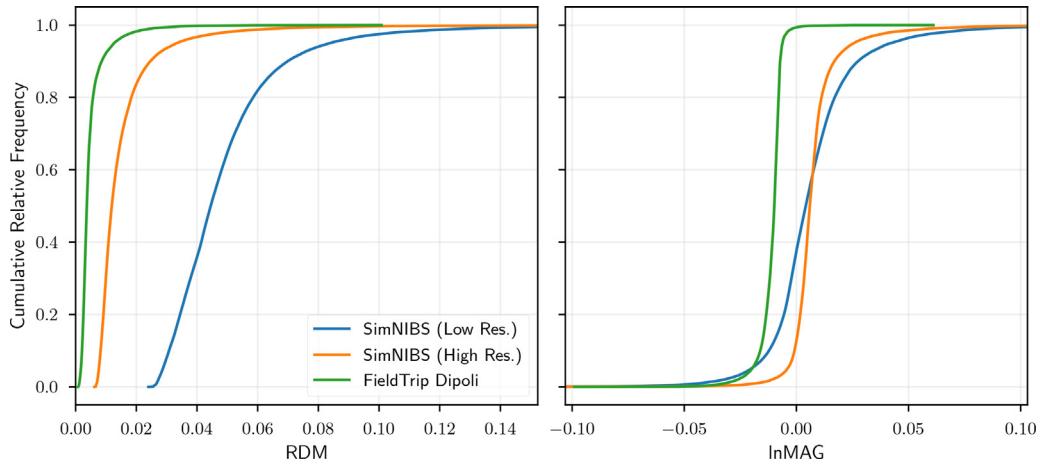
Fig. 3. Realistic geometry. CRF of RDM and lnMAG for a high and low resolution FEM model and a BEM model in a realistically shaped three-compartment model. The BEM solution from MNE-Python was used as a reference.
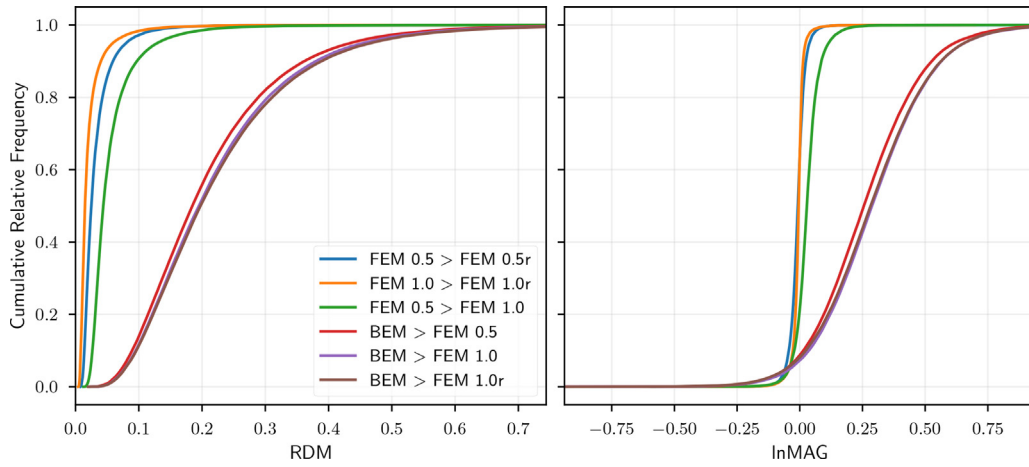

Fig. 4. Realistic geometry. CRF of RDM and lnMAG for FEM comparing five-compartment FEM models of different resolution (0.5 and 1.0 nodes/mm²), refined (i.e., upsampled) versions of these models, and a three-layer BEM model. The model on the right-hand side of the greater-than sign is the reference.
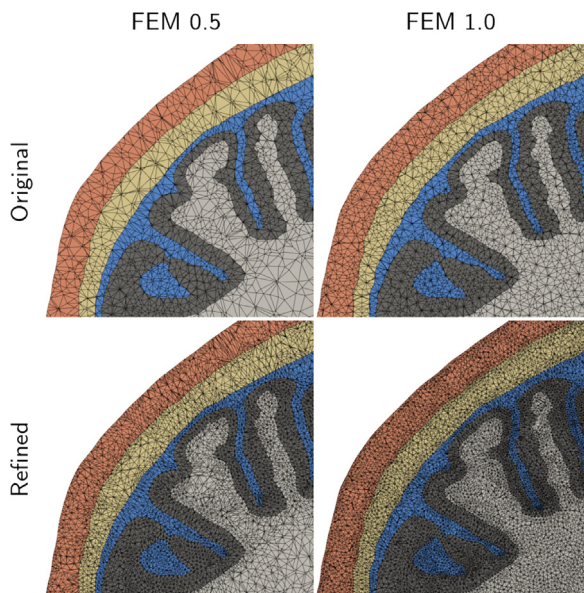
Fig. 5. Realistic geometry. Zoomed view of a coronal slice showing the effect of resolution and refinement on the head models. The corresponding effects on the forward solution can be found in Fig. 4.

that most of the differences remain even after this correction, and so we believe that our statements above are indeed valid. Due to the vast geometrical simplifications in the MNE-FS model (collapsing brain tissue and CSF into one compartment) and their known effects on the forward solution (Vorwerk et al., 2014), we did not do a similar analysis with the MNE-FS model as we feel the geometrical adjustments to the reference model would be too large.

As SimNIBS-CHARM was the only model to use both a T1- and a T2-weighted image, we also ran this pipeline using only a T1-weighted image. From Fig. 8, it is clear that performance is not substantially decreased which is consistent with previous results showing that the segmentation procedure is fairly robust even without a T2-weighted image (Puonti et al., 2020a).

3.3. Impact of electrode positions on the forward solution

We first report errors on the channel positions before describing the resulting effects on the forward solution.

Fig. 11 is a density plot of the Euclidean distance between electrode positions estimated using the different methods and the digitized positions across all subjects. It is evident that Man-Template generally performs worse than Custom-Template and also that the particular transformation method of Custom-Template does not have a strong influence on the accuracy of the electrode positions.
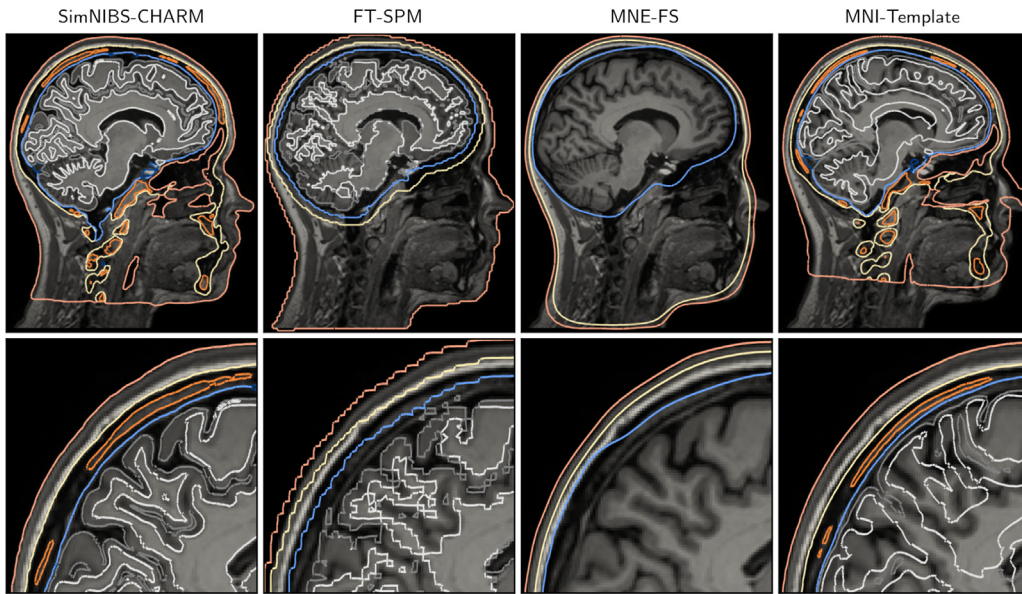
Fig. 6. Tissue compartments as estimated by each of the pipelines in the anatomy study for subject one. It is apparant that MNE-FS does not capture the inner skull very well for this subject. The skull thickness varies substantially but this does not seem to agree with the structural MRI scan. Also, the lower part of the head is mostly modeled as skull as the outer skull surface is obtained by eroding the skin surface. Furthermore, the skull thickness in FieldTrip-SPM is almost constant throughout the volume (e.g., thickness is overestimated around the occipital cortex). MNI-Template captures the skull surprisingly well in this subject but the brain tissues are obviously not accurate.
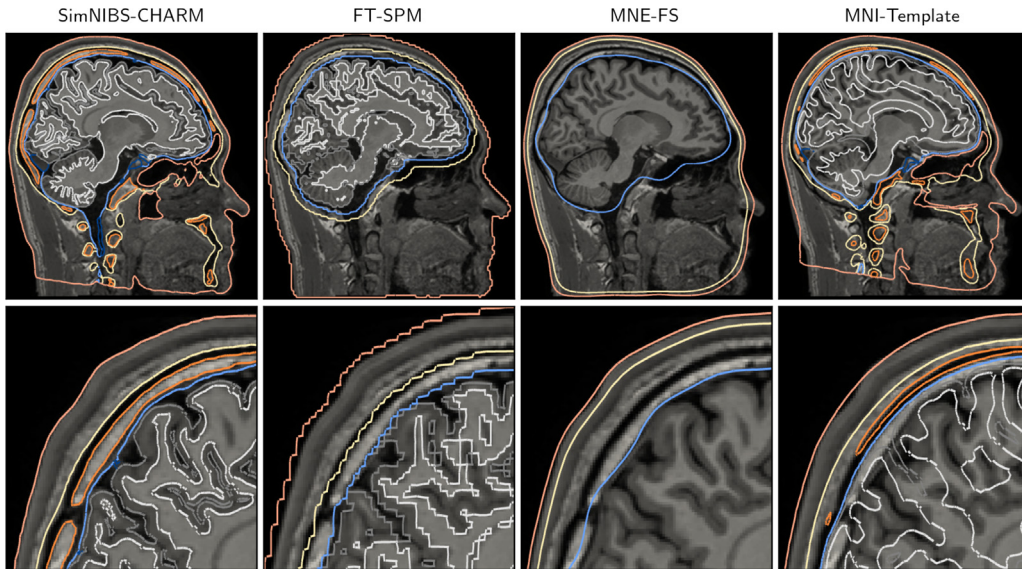
Fig. 7. Tissue compartments as estimated by each of the pipelines in the anatomy study for subject two. MNE-FS captures the inner skull better in this subject compared to Fig. 6, however, the skull thickness is not well modeled. FieldTrip-SPM captures inner skull and thickness well except for the prefrontal part. MNI-Template displaces the skull outwards in the zoomed-in view whereas other parts are modeled more accurately. The spongy bone, however, is not.

Fig. 12 shows the spatial distribution of the mean error over channels. We see that the errors in the posterior areas (mostly occipital but also parietal and temporal) are large for Man-Template whereas this is much less severe for the Custom-Template. Likewise, the standard deviation is also larger for Man-Template, again in posterior but also anterior areas. Interestingly, although the mean error is similar across the different transformation methods for Custom-Template, the standard deviation is not. The MRI-based coregistrations show larger variability in the anterior parts of the head (suggesting that other parts of the image are driving the registration). The landmark-based registration exhibits larger variability in the posterior regions (perhaps unsurprisingly given that the landmarks used here are the nasion, LPA, and RPA).

To determine potential spatial bias in the template registrations, we show the absolute errors along each axis in Fig. 13. In general, it seems that errors in anterior and posterior areas are largest in the z direction and errors in the superior areas are largest in the y direction. This suggests that the errors here are mostly due to an anterior-posterior (AP) shift (i.e., a rotation around the x-axis). Interestingly, errors in the most inferior electrodes in the temporal areas are primarily due to shifts along the AP direction and not up or down. Finally, please note the different scales of Custom-Template and Man-Template and that errors may be exaggerated close to the edges of the head contour due to extrapolation.

For both Custom-Template and Man-Template, we found that channel errors in occipital and frontal regions were mostly along the z-axis and along the y-axis for superior areas. To investigate the extent to
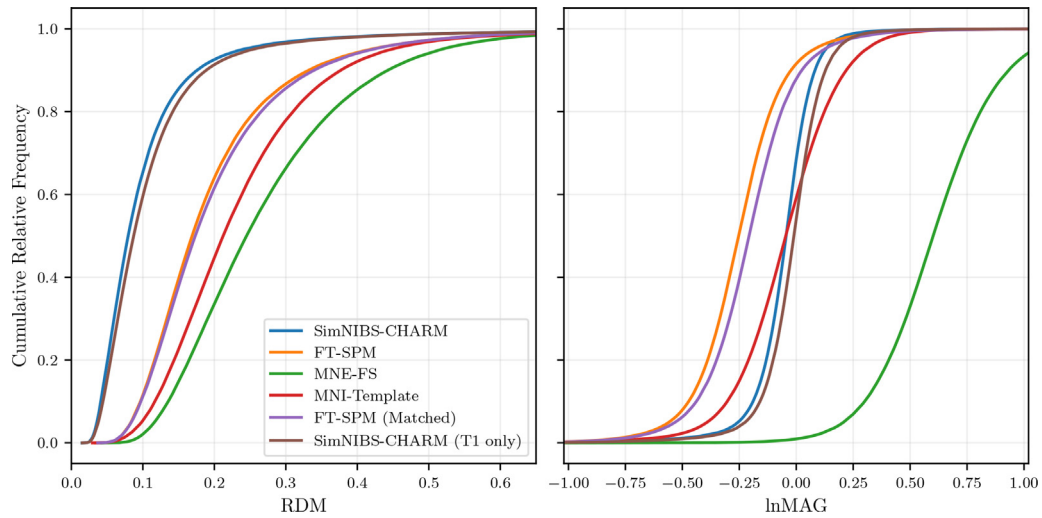
Fig. 8. CRF of RDM and lnMAG for each model. The model based on manual segmentation was used as reference. FieldTrip-SPM (Matched) is a comparison of the FieldTrip-SPM model with a reference with matched conductivities (see Section 4.2.2) allowing us to disentangle effects of geometry and conductivity. SimNIBS-CHARM (T1 only) is a model constructed using only a T1-weighted image as input.
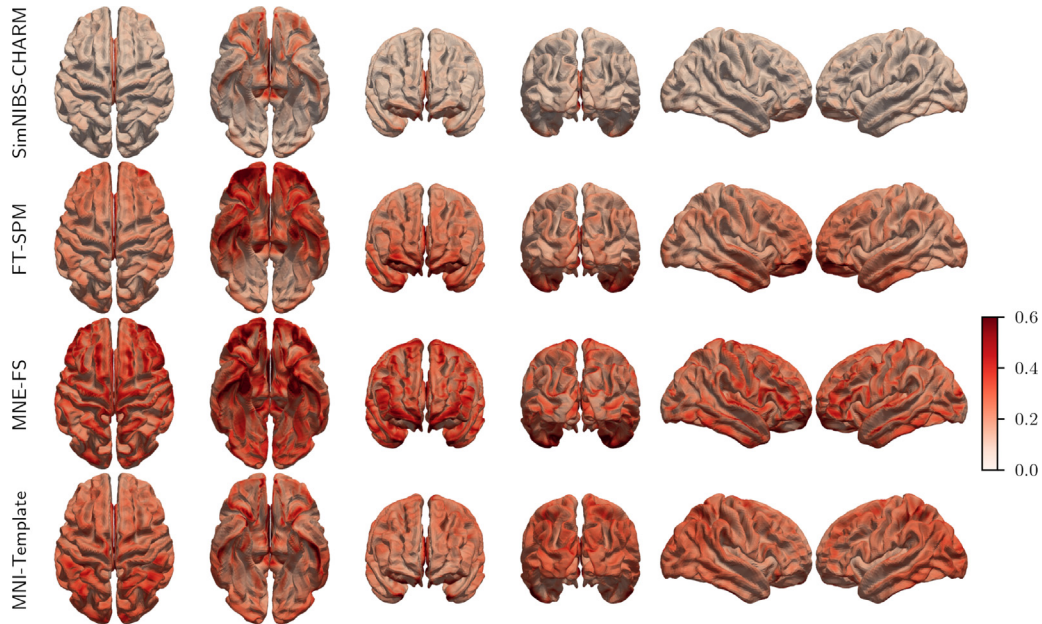

Fig. 9. Spatial distribution of mean RDM for each model. The model based on manual segmentation was used as reference.

which errors were due to bias (i.e., a systematic difference between the template and the average positions across subjects) or variability in the reference positions (i.e., the standard deviation of each position, e.g., due to inter-individual differences in cap placement), we transformed the reference positions for all subjects to standard space and computed bias and standard deviation for each position. The standard deviations (Fig. 14) showed the same pattern as that observed in Fig. 13, suggesting more variability in AP alignment of the cap compared to aligning it laterally. On the other hand, we also found a higher bias at some positions than others (Fig. 15). For example, the digitized positions of the electrodes between the ears and the eyes were, generally speaking, more anterior compared to Custom-Template.

We also show all digitized electrode positions in standard space on the MNI surface template (Fig. 16) from which it is apparent that the positions towards the back are lower in the left side compared to the right side. This points to some kind of misalignment, perhaps due to the way the chin strap was fastened being different between subjects and

the MNI template head. As such, it is not clear whether the observed bias is due to suboptimal (biased) cap placement in the data set used here or because the template fails to capture how a cap fits on an actual human head.

Fig. 17 shows CRF of the error metrics, RDM and lnMAG, for the forward solutions using the solution associated with the digitized positions as reference. It is evident that the RDM errors are generally higher for Man-Template than for Custom-Template whereas the errors for MNI-Digitized fall between the two. As to lnMAG, Custom-Template is again better than Man-Template but here MNI-Digitized is clearly worse than the other two as evidenced by the heavy tails.

Figs. 18 and 19 present the spatial distribution of the mean RDM and lnMAG on the *fsaverage* surface. From Fig. 18, we see slightly elevated topographical errors for Custom-Template in orbitofrontal and occipital errors. However, this is more pronounced in Man-Template which also shows increased errors in parietal and temporal regions. The errors of MNI-Digitized are generally similar to those of Custom-
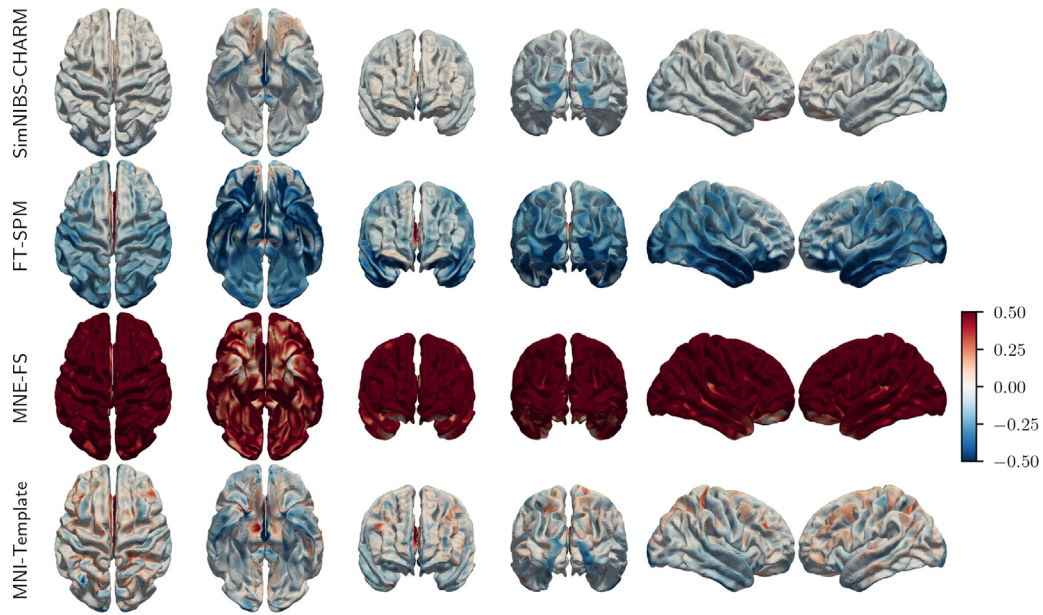
Fig. 10. Spatial distribution of mean lnMAG for each model. The model based on manual segmentation was used as reference.
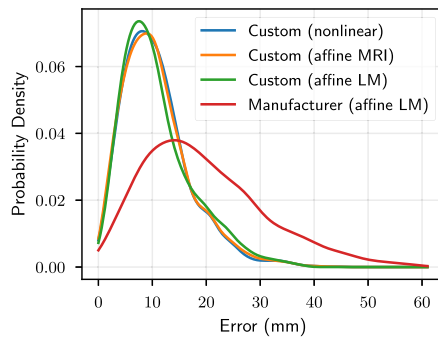

Fig. 11. Density of errors in electrode position as measured by the Euclidean distance between each model and the digitized positions. The parenthesis denotes the registration method where "LM" means that the registration was based on landmarks.

Template. However, we see several hot spots of increased error, often on the gyral crowns. From Fig. 19, we see slightly increased magnitudes for Custom-Template in frontal and right temporal regions. Generally, Man-Template shows increased and decreased magnitudes for deep and shallow sources, respectively. Again the magnitudes are elevated in the right temporal region. MNI-Digitized shows large magnitude errors with increased and decreased sensitivity to sources on the gyral crowns and sulcal walls, respectively. Decreased sensitivity is also observed for orbitofrontal and occipital sources.

4. Discussion

4.1. Validation of SimNIBS for EEG

The simulations in spherical models, assuming sufficient model resolution of about 0.5 or 1.0 nodes/mm^2, showed that the numerical errors in topography and magnitude are low and similar to those of existing FEM implementations for EEG (Vorwerk et al., 2012). The same was true for simulations using a realistic geometrical models. In addition, by comparing numerical errors of our FEM to errors incurred by not distinguishing different tissue compartments in the BEM head models (see Figs. 3 and 4), we show that the latter are about ten times those of the former. Hence, we suggest that in general the decrease in numerical accuracy associated with FEM is offset by its flexibility when it comes to modeling complex anatomical structures.

Close inspection of Fig. 2 reveals that there is a small bias in the magnitude in the SimNIBS simulations. Decreasing the thickness of
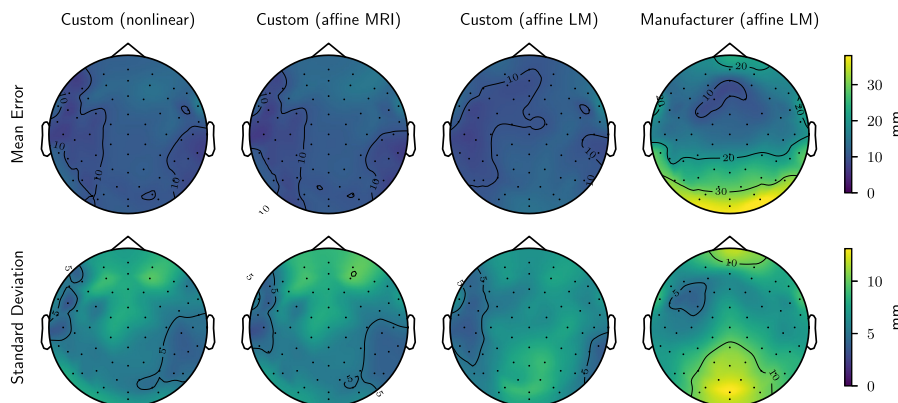
Fig. 12. Spatial distribution of errors (mean and standard deviation) in electrode position as measured by the Euclidean distance between each model and the digitized positions. The parenthesis denotes the registration method where "LM" means that the registration was based on landmarks.
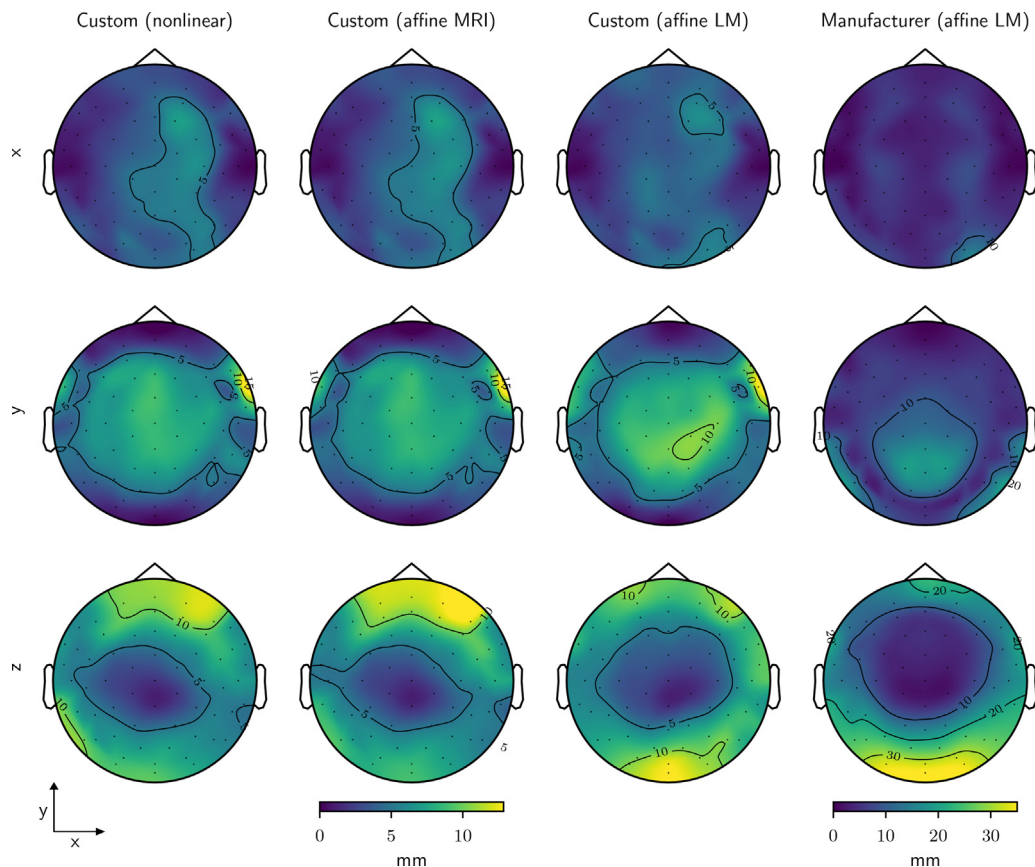
Fig. 13. Spatial distribution of mean errors in electrode position as measured by the absolute difference between each model and the digitized positions. Please note the different scale of the rightmost column. The parenthesis denotes the registration method where "LM" means that the registration was based on landmarks. Positive z direction is out the page.
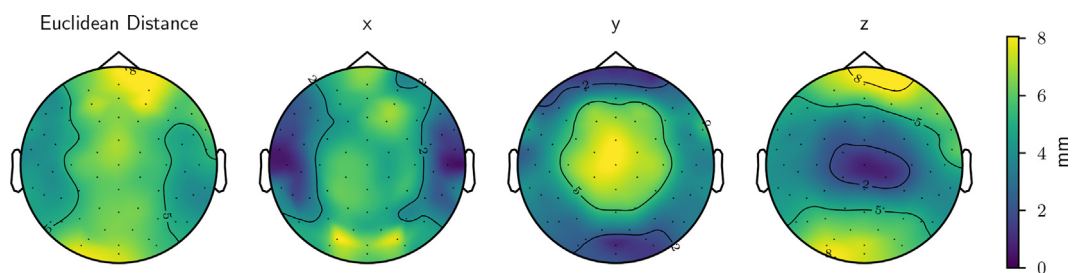

Fig. 14. Standard deviations of the digitized positions in the electrode study calculated using the overall Euclidean distance or distances along each major axis. This illustrates the variability of the reference positions.
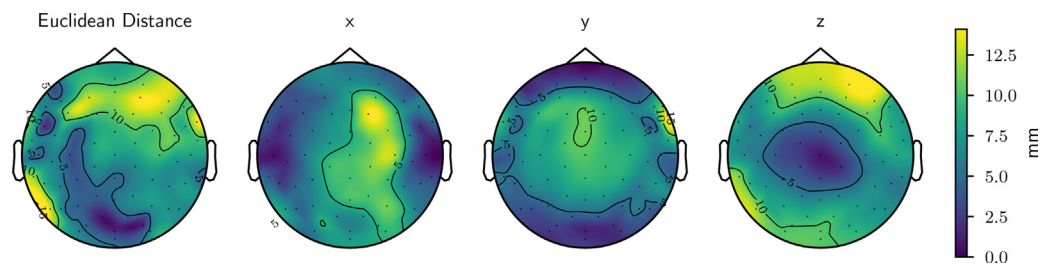
Fig. 15. Distance between the average of the digitized positions in the electrode study and the positions of the custom template calculated using the overall Euclidean distance or distances along each major axis. This illustrates the bias of the template compared to the reference positions.
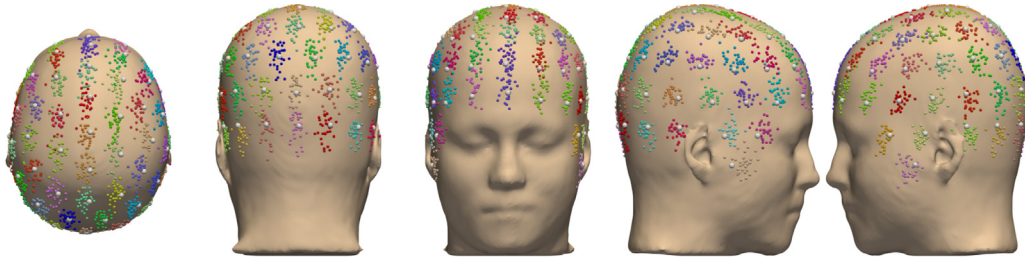
Fig. 16. Our custom template positions (white spheres) and the digitized positions for each subject (colored spheres) in MNI coordinates.
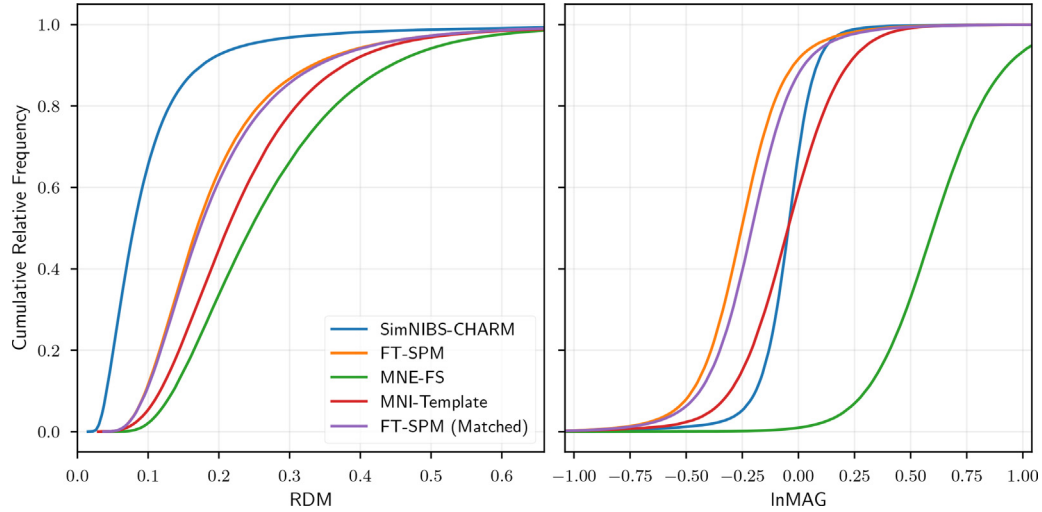
Fig. 17. CRF of RDM and lnMAG for each model. The model based on digitized electrode positions was used as reference.
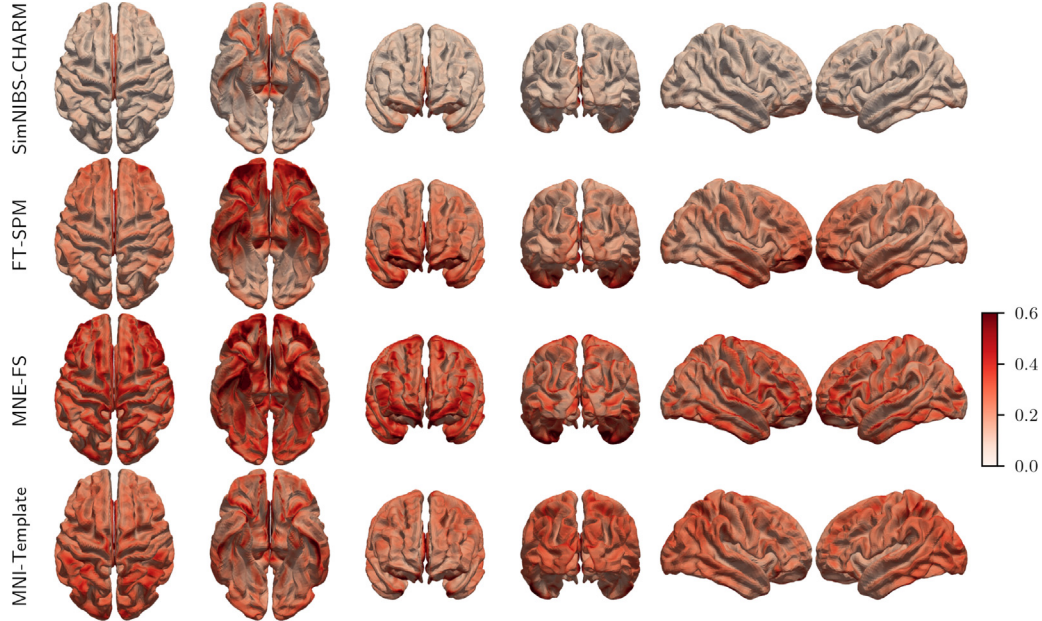
Fig. 18. Spatial distribution of mean RDM for each model. The model based on digitized electrode positions was used as reference.

the electrodes (thus minimizing potential drop over the electrode itself) and matching the conductivity to that of skin (since high conducting electrodes effectively makes the skin compartment a better conductor) removed this effect. Thus, we attribute this to the electrodes being meshed in SimNIBS whereas the analytical solution assumes point sensors. One could argue that the correct approach is to model the electrodes. However, the effects are very small in practice.

From Fig. 2 it is clear that errors increase as sources get closer to the electrodes and tissue interfaces. As sources get close to the electrodes, the St. Venant principle will become less valid and therefore a blurred dipole model (as used here by SimNIBS and FieldTrip/DUNEuro) will approximate the mathematical dipole less well. This holds irrespective of whether there are tissue boundaries or not. The particular source model will also be important and when nodes are excluded because they are not in gray matter, there will be less nodes to model the dipole (and the
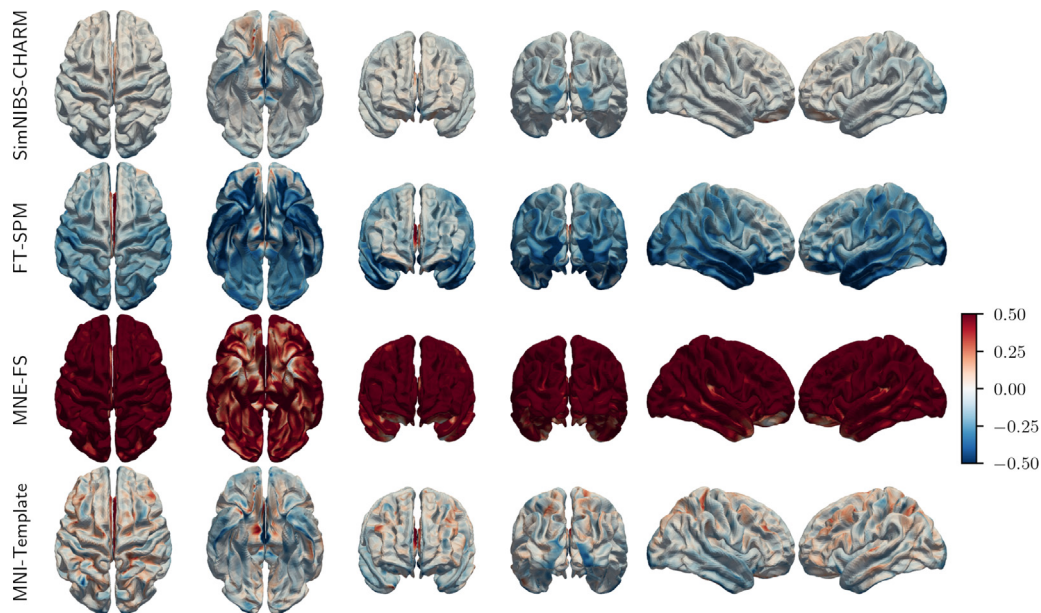
Fig. 19. Spatial distribution of mean lnMAG for each model. The model based on digitized electrode positions was used as reference.

cloud of nodes available will be spatially biased away from the interface) (Medani et al., 2015). In addition, the current flow tends to change more rapidly at tissue interfaces because of the change in conductivity. Therefore we speculate that one would require higher resolution to model this well. Currents tend to be strongest close to sources, hence when sources are placed close to an interface with a (high) conductivity gradient we might expect that the simulations will be less precise unless resolution is increased correspondingly.

As to the source model used in SimNIBS, calculating the electric field in each element is similar to the partial integration approach (Vorwerk, 2016) with dipoles along each axis. However, to improve the accuracy of the estimate, we use a postprocessing procedure (SPR) which consists in fitting a smooth function such that the value of a node is given by a linear combination of its neighboring elements. Although different from the St. Venant approach (Vorwerk, 2016), we found the numerical accuracy to be similar.

4.2. Impact of anatomical accuracy on the forward solution

4.2.1. Topographical errors

Interpreting the topographical errors reflected by RDM can be difficult since the quality of the anatomical models may vary from subject to subject. However, we will summarize some general observations about the different models.

Both FieldTrip-SPM and MNE-FS showed elevated topographical errors in orbitofrontal and deep temporal areas. Given the close proximity of such sources to the facial area and the fact that neither of these models attempt to characterize the complex anatomy in this region, we suspect that such simplifications may explain the increased RDM values here. Since the anatomy is simpler around the occipital cortex, the modeling decision of FieldTrip-SPM to simply enclose the brain with a certain amount of bone (and CSF) seems more appropriate here. The same holds for MNE-FS which also showed large RDM values for superficial sources in the superior part of the brain. Similar observations were made by Vorwerk et al. (2014) when distinguishing CSF and brain tissues. The inaccuracy of the inner and outer skull surfaces in MNE-FS likely exacerbates this issue (Lanfer et al., 2012).

In line with the better anatomical accuracy, the topographical errors of SimNIBS-CHARM are generally small compared to the other models (see Figs. 6 and 7). MNI-Template seems slightly better than MNE-FS

but worse than the others. This seems to highlight the importance of accurately modeling CSF since the brain tissue compartments in MNI-Template are generally very different from the individual anatomy.

4.2.2. Magnitude errors

As with the topographical errors, the magnitude errors of SimNIBS-CHARM are low whereas they are slightly higher for MNI-Template (heavier tails in the CRF curve). The errors of MNE-FS and FieldTrip-SPM are larger still and biased towards higher and lower sensitivity, respectively.

The magnitude errors in FieldTrip-SPM suggest that sensitivity is generally decreased, particularly for deep sources. From Figs. 6 and 7 it is evident that the models tend to contain too much CSF, particularly in the deep and temporal areas where bone is also too thick, which will increase shunting effects. MNE-FS shows the opposite effect, i.e., an overall increase in magnitude, most prominently for sources close to the sensors. MNE-FS does not model the CSF compartment explicitly but rather collapses white matter, gray matter, and CSF into a single compartment with an adjusted conductivity. Thus, the conductivity gradient between the brain and bone compartments is smaller and hence we expect less current shunting. These results are in line with Vorwerk et al. (2014) in terms of the relevance of distinguishing CSF from brain tissue as well as white matter from gray matter.

4.3. Impact of electrode positions on the forward solution

Similar to Homölle and Oostenveld (2019), we also found that channel errors were generally much larger for Man-Template than Custom-Template in occipital and parietal areas suggesting that approximating the head shape with a sphere is problematic when the sphere is aligned using frontal and temporal landmarks. Including the inion in the coregistration could potentially help, although we did not test this. On average, we did not find similar effects in posterior regions when transforming Custom-Template using landmarks (although standard deviations were higher), suggesting that the misalignment of Man-Template is in fact due to this layout being a poor fit of the actual positions in these areas.

We also investigated the extent to which the errors observed in Custom-Template were due to bias (i.e., systematic differences to the reference positions) or variability (in the reference positions). We found that the errors were partly due to a bias of the template but also that

some electrodes were less consistently placed on the individual subjects. We did not do a similar analysis for Man-Template.

Given this bias and variability in the electrode positions, it is not reasonable to expect a perfect fit of the template to all subjects. We noticed that the fit was particularly bad in a few cases where the cap was clearly not aligned well (e.g., rotated around the *x* or *z* axis).

4.3.1. Topographical errors

In line with the mean error observed on the channel level, topographical errors in the forward solutions, as estimated by RDM were more severe for Man-Template than Custom-Template. For the former, errors were high in occipital and posterior areas. Errors were higher on the gyral crowns (i.e., where the sensitivity is high) than the sulcal walls and valleys, something which is most clearly seen for Man-Template.

Topographical errors for MNI-Digitized are slightly higher than Custom-Template but smaller than Man-Template suggesting that electrode positions are important in determining such errors and that template electrode positions as supplied by the manufacturer result in higher RDM values compared to a template anatomy with digitized positions. This is in line with the results of Homölle and Oostenveld (2019).

We also see a few RDM hot spots in MNI-Digitized which, as they stand out on a group level, are likely due to local inaccuracies in this model. We were not able to identify any obvious errors upon visual inspection of the final segmentation and mesh though.

4.3.2. Magnitude errors

We observed that Man-Template showed increased and decreased magnitudes for deep and superior sources, respectively. Visual inspection of the final positions showed that electrodes were generally placed further down on the head model, thus giving increased sensitivity to deep (or ventral) sources, and, conversely, decreased sensitivity to superior sources (most notably parietal, occipital, and prefrontal) as the montage was spread more thinly across the head. Increased sensitivity to temporal sources was also observed as these are generally not very well covered by electrodes.

Magnitude errors were much larger for MNI-Digitized. In particular, we observed increased sensitivity to sources on the gyral crowns and decreased sensitivity to sources on the sulcal walls. Since the MNI152 T1-weighted template represents a standard anatomy, it is generally quite blurred. As a result, fine details such as gray matter curvature and sulcal structure are lost. Visual inspection (see Fig. 20) showed that the gray matter compartment of the final head model was much smoother than for individual subjects, hence sources on sulcal walls and in sulcal valleys were surrounded almost exclusively by gray matter and not highly conducting CSF as was the case in the models based on the individual anatomy. Additionally, we also found that the CSF layer was much thinner compared to the individualized models resulting in less current shunting, particularly for sources on the gyral crowns to which sensitivity was increased (see Fig. 20).

4.4. Effect on inverse solution

Despite clear differences in the forward solutions obtained from the different pipelines investigated in this study, the impact on the inverse solution still needs to be established. From an EEG perspective, these results are only relevant to the extent that they translate to differences in inverse solutions. There is, however, already a body of literature suggesting that the forward model can indeed be an important factor that influences source reconstructions.

In the anatomy study, we saw clear differences in skull reconstructions between the models and several studies have shown how such errors may result in decreased source localization accuracy (Chauveau et al., 2004; Dannhauer et al., 2011; Fiederer et al., 2016; Lanfer et al., 2012; McCann and Beltrachini, 2022; Montes-Restrepo et al., 2014). For example, it is apparent that FieldTrip-SPM and MNE-FS simplifies the modeling of the inferior part of the skull substantially
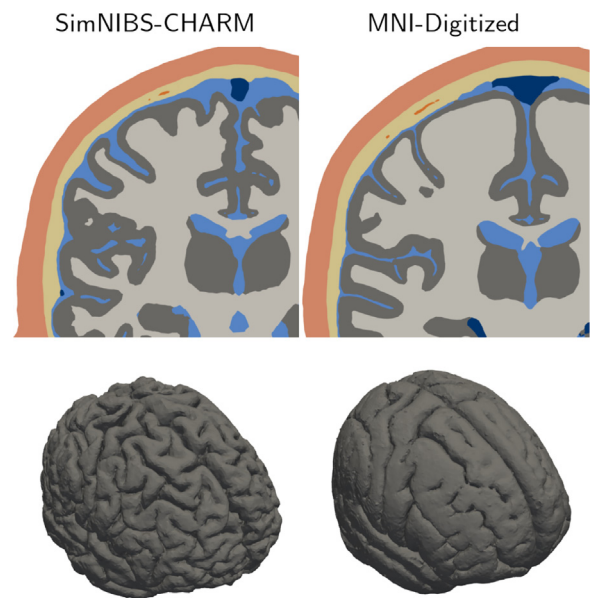

Fig. 20. The top row illustrates the reduced amount of CSF between the cortex and the inner skull in MNI-Digitized compared to the SimNIBS-CHARM for a particular subject (the MNI-Digitized model is almost identical for all subjects except for the nonlinear deformation). The bottom row shows the corresponding gray matter compartments.

which was shown by Lanfer et al. (2012) to result in large errors (more than 1 cm) but only for sources close to the base of the skull. Likewise, we see that the skull thickness is not accurately captured in FieldTrip-SPM and MNE-FS whereas SimNIBS-CHARM is able to delineate the skull with decent accuracy (Puonti et al., 2020b). However, previous studies report only minor source localization errors (less than 5 mm) due to such erroneous thickness estimation (Chauveau et al., 2004; Lanfer et al., 2012). Thus, not all aspects of the geometry may be equally important for source analysis.

Another important consideration when constructing a head model is whether or not to model the CSF compartment separately. Again, this has been shown to be important for accurately localizing sources. Ramon et al. (2006) concluded that the CSF compartment is important for shaping the observed potential distribution on the scalp and neglecting this increases average source localization errors by a few millimeters. Likewise, Conte and Richards (2021) found that including CSF was the most important addition to a model consisting initially of brain, skull, and scalp and that when doing so, it performed similarly to more complex models distinguishing even more tissue types. The latter suggests that the increased complexity of the SimNIBS-CHARM model (including for example muscle tissue and eyes) may not matter much in terms of source localization accuracy. Similarly, differentiating between white and gray matter was found to be less important compared to CSF (Ramon et al., 2006).

Finally, constructing a head model based on template anatomy has been found to increase localization errors approximately 5 mm on average (Acar and Makeig, 2013).

There are also studies which have shown how source localization accuracy deteriorates with increasing errors in electrode positions, e.g., due to digitization errors (Dalal et al., 2014; Shirazi and Huang, 2019), using template positions (Homölle and Oostenveld, 2019) or simulating noise on the positions (Acar and Makeig, 2013; Wang and Gotman, 2001). In this study, we found spatially correlated errors concentrated around occipital and parietal areas. This was true for both Custom-Template and Man-Template, but was more prominent for the latter. Similar results were reported by Homölle and Oostenveld (2019), who found average dipole localization errors of 8.5 and

11 mm for custom and manufacturer templates, respectively. Again, particularly high errors for the latter occurred over the above mentioned areas.

While most of the literature reviewed here suggests that errors in the head model and electrode positions used for forward modeling affect the accuracy of the inverse solution, this influence may be modulated by other factors such as the SNR of the data and the inverse modeling approach used. For example, Montes-Restrepo et al. (2014) showed that source localization accuracy benefits from accurate skull modeling based on CT, however, they also found that differences between CT and MRI based skull models disappeared with increasing noise levels (e.g., going from SNR of 10 dB to SNR < 5 dB). Ramon et al. (2006) investigated the effect of distinguishing different tissue compartments (e.g., white matter from gray matter or merging CSF, white matter, and gray matter) on source localization accuracy. They found that differences between models only started to emerge above an SNR of approximately 5dB. Thus, a sufficient SNR seems to be required to benefit from more complex forward modeling. Both of these studies used dipole fitting for localizing sources. On the other hand, distributed inverse methods are generally expected to be more robust to forward modeling errors at the expense of being less focal (even in ideal conditions) as well as suffering from depth bias (Stenroos and Hauk, 2013).

Taken together, we expect the errors which we report here to be relevant in a source localization context although it is clear that differences in forward solutions will not translate one-to-one to differences in the corresponding inverse solution as the latter is determined by other factors as well, e.g., the noise level in the data (Montes-Restrepo et al., 2014; Ramon et al., 2006). That being said, we expect that accurately capturing as many aspects of the forward problem as possible will ultimately result in better source estimates (everything else being equal) even though effects may be small.

4.5. Limitations

The comparison of different pipelines is likely biased in favor of SimNIBS-CHARM since the tissue priors used by the CHARM pipeline was built from this dataset. Although we used alternative priors from a four-fold CV split such that a particular subject never contributed to the prior used for generating its own segmentation, we might still expect some bias. On the other hand, it is clear from simple visual inspection of the segmentations that SimNIBS-CHARM is more accurate than FieldTrip-SPM and MNE-FS. As such, we believe the observed differences are robust and will generalize to other MRI datasets as well. Similar to SimNIBS-CHARM, the results of the model based on a template anatomy might also be biased given that it contains the same tissue classes and was constructed from the same pipeline. To minimize bias, we again used priors from a four-fold CV split.

The reference models used in the anatomy study, although based on manual segmentations, were still not perfect and certain simplifications were made. For example, we failed to model brain conductivity anisotropy. However, assuming that the effect is similar across different pipelines, we do not expect it to strongly affect relative differences. Besides, white matter anisotropy seems to mostly affect certain deep sources and effects are generally small compared to other modeling aspects (Vorwerk et al., 2014).

In the anatomy study, we compared each pipeline as is. This means that the models differed not only in terms of anatomy but also other aspects such as numerical method (BEM and FEM), solver, source model, element type (tetrahedra, hexahedra, triangles) and resolution. As to the effects of element type and resolution, we refer to section S3 in the supplementary material where we show that these effects are small compared to the overall differences we observe. Likewise, since all the methods used here have been numerically validated, we believe that the major factor contributing to the differences that we report here is in fact the anatomical accuracy of the models.

A final point relates to the use of additional contrasts for segmentation. All methods use a T1-weighted image, however, only SimNIBS-CHARM uses a T2-weighted image as well. We expect head model accuracy to improve when a T2-weighted image is used as well (Nielsen et al., 2018; Puonti et al., 2020a). In the FieldTrip pipeline, this can be achieved by doing the segmentation in SPM12 beforehand and then feed the results to FieldTrip. However, the use of additional contrasts is currently not supported through the FieldTrip interface which is why we did not use a T2-weighted image in this study. As a control condition, we also tested SimNIBS-CHARM using only a T1-weighted image. As shown in Fig. 8, this does not decrease the performance of CHARM notably. Also, previous results suggest that CHARM performs better than SPM12 in terms of segmentations accuracy even when a T2-weighted image is used (Puonti et al., 2020a).

In the electrode study, we used a dataset collected at our department as part of a previous project. Since electrode positions were digitized in this study, the experimenter may have been less careful with cap placement than if it was known beforehand that a template was to be used to model the electrode positions. As such, they might not constitute the best possible case when evaluating the fit of an electrode template. For example, visual inspection (as well as the standard deviation some of the electrodes across the dataset) suggested that in some subjects the cap was clearly misaligned in the anterior-posterior direction or twisted left or right. On the other hand, we believe that this is likely to happen in an actual experimental setting and as such our results may give a realistic impression of the errors that might be incurred using such methods.

5. Conclusion

In summary, it seems that topographical errors, as measured by RDM, are sensitive to errors in the modeled anatomy and the sensor positions, whereas magnitude errors, as measured by lnMAG, are sensitive to anatomical errors only. Consequently, magnitudes from a template model or simplified models cannot be trusted. Based on the results presented here, we suggest to use a forward model as realistic as possible. The availability of pipelines able to generate realistic geometrical descriptions of the human head have so far been limited (although see Medani et al., 2023, which describes the current FEM pipeline in Brainstorm). SimNIBS lowers the practical hurdles of creating detailed head models as the pipeline (segmentation, meshing, etc.) is fully automated. Compute time is a few hours which is similar to the other tools used in this work (FreeSurfer with MNE-Python and FieldTrip with DUNEuro) but typically more than what is required to construct a three-layer BEM model (which is often used in EEG source analysis).

If digitized electrode positions are unavailable, we suggest to use a template created by measuring the positions on a realistically shaped head template. Alternatively, average positions from a previous study using the same cap could be used. The template described here will be included in SimNIBS and we plan to add other caps in the future. Perhaps somewhat surprisingly, using a geometrical model based on template anatomy did not result in larger errors than coarse anatomical modeling of the individual anatomy. This suggests that if individual anatomical information is lacking then it might still be possible to generate a usable forward model based on standard anatomy. Even though we did not evaluate the impact of the observed effects on source localization accuracy, previous studies suggest that our main findings will translate to source localization although this association will likely be modulated by the noise level in the data and the particular choice of inverse solver.

Data availability

The data used in this study has been presented previously and we refer to the original articles for statements about data availability. The

```
# Create head model based on T1- and (optionally) T2-weighted MRI scan.
↪   This creates
#   ./m2m_subid (directory)
charm subid t1.nii.gz t2.nii.gz

# Create montage file. This creates
#   ./eeg_montage.csv
prepare_eeg_montage mne eeg_montage.csv info.fif trans.fif

# Run TDCS leadfield calculations. This creates
#   ./fem_subid/subid_leadfield_eeg_montage.hdf5
prepare_tdcs_leadfield subid eeg_montage.csv -s 10000 -o fem_subid

# Convert TDCS leadfield to forward solution in MNE-Python format. This
↪   creates
#   ./fem_subid/subid_leadfield_eeg_montage_subsampling-10000-fwd.fif
#   ./fem_subid/subid_leadfield_eeg_montage_subsampling-10000-morph.h5
#   ./fem_subid/subid_leadfield_eeg_montage_subsampling-10000-src.fif
prepare_eeg_forward mne subid fem_subid/subid_leadfield_eeg_montage.hdf5
↪   info.fif trans.fif --fsaverage 10
```

Listing A.1. Generating an EEG forward solution for MNE-Python with SimNIBS. `info.fif` can be a Raw, Epochs, Evoked or Info object. `trans.fif` is a Trans object containing an affine mapping between head coordinates and subject MRI coordinates. `-fwd.fif`, `-morph.fif`, and `-src.fif` are standard MNE-Python objects containing the forward solution, the sparse matrix to morph to *fsaverage* space, and the source space definition, respectively.

```
# Create head model based on T1- and (optionally) T2-weighted MRI scan.
↪   This creates
#   ./m2m_subid (directory)
charm subid t1.nii.gz t2.nii.gz

# Create montage file. This creates
#   ./eeg_montage.csv
prepare_eeg_montage fieldtrip eeg_montage.csv info.mat -t trans.mat

# Run TDCS leadfield calculations. This creates
#   ./fem_subid/subid_leadfield_eeg_montage.hdf5
prepare_tdcs_leadfield subid eeg_montage.csv -s 10000 -o fem_subid

# Convert TDCS leadfield to forward solution in FieldTrip format. This
↪   creates
#   ./fem_subid/subid_leadfield_eeg_montage_subsampling-10000-fwd.mat
#   ./fem_subid/subid_leadfield_eeg_montage_subsampling-10000-morph.mat
#   ./fem_subid/subid_leadfield_eeg_montage_subsampling-10000-src.mat
prepare_eeg_forward fieldtrip subid
↪   fem_subid/subid_leadfield_eeg_montage.hdf5 --fsaverage 10
```

Listing A.2. Generating an EEG forward solution for FieldTrip with SimNIBS. `info.mat` contains information about the electrodes and `trans.mat` contains an affine mapping between the coordinate system of the electrodes and MRI space. `-fwd.mat` contains the forward solution, `-morph.mat` is a sparse matrix which can be used to "morph" from subject space to *fsaverage* space, and `-src.mat` contains the source space definition.

data used to compare the different forward modeling pipelines was presented by Farcito et al. (2019). The data used to compare electrode specification strategies was presented by Madsen et al. (2019) and Karabanov et al. (2021).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Credit authorship contribution statement

Jesper Duemose Nielsen: Conceptualization, Formal analysis, Investigation, Methodology, Software, Writing – original draft, Writing – review & editing, Visualization, Funding acquisition. **Oula Puonti:** Methodology, Software, Writing – original draft, Writing – review & editing, Supervision. **Rong Xue:** Writing – review & editing, Supervision. **Axel Thielscher:** Conceptualization, Methodology, Software, Writing – original draft, Writing – review & editing, Supervision. **Kristoffer Hougaard Madsen:** Conceptualization, Methodology, Software, Investigation, Supervision, Writing – original draft, Writing – review & editing, Project administration, Funding acquisition.

Appendix A. Generating EEG forward solutions with SimNIBS

Here we show how to generate EEG forward solutions with SimNIBS which can be used with MNE-Python (Listing A.1) or FieldTrip (Listing A.2). In both examples, the central gray matter surface of each hemisphere is subsampled to 10,000 nodes per hemisphere and a mapping to *fsaverage5* (which contains 10,242 nodes per hemisphere) is constructed. In both cases, the procedure is the same but the output format is different: when exporting to MNE-Python format (Listing A.1), the forward solution is in the head coordinate system used by MNE-Python whereas when exporting to FieldTrip (Listing A.2), the forward solution is in subject MRI coordinate system.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2023.120259

References

Acar, Z.A., Acar, C.E., Makeig, S., 2016. Simultaneous head tissue conductivity and EEG source location estimation. Neuroimage 124, 168–180. doi:10.1016/j.neuroimage.2015.08.032.

Acar, Z.A., Makeig, S., 2013. Effects of forward model errors on EEG source localization. Brain Topogr 26 (3), 378–396. doi:10.1007/s10548-012-0274-6.

Ary, J.P., Klein, S.A., Fender, D.H., et al., 1981. Location of sources of evoked scalp potentials: corrections for skull and scalp thicknesses. IEEE Trans. Biomed. Eng. BME-28 (6), 447–452. doi:10.1109/tbme.1981.324817.

Attene, M., 2010. A lightweight approach to repairing digitized polygon meshes. Vis. Comput. 26 (11), 1393–1406. doi:10.1007/s00371-010-0416-3.

Aydin, U., Vorwerk, J., Küpper, P., Heers, M., Kugel, H., Galka, A., Hamid, L., Wellmer, J., Kellinghaus, C., Rampp, S., Wolters, C.H., 2014. Combining EEG and MEG for the reconstruction of epileptic activity using a calibrated realistic volume conductor model. PLoS ONE 9 (3), 1–17. doi:10.1371/journal.pone.0093154.

Azizollahi, H., Aarabi, A., Wallois, F., 2016. Effects of uncertainty in head tissue conductivity and complexity on head tissue forward modeling in neonates. Hum. Brain Mapp. 37 (10), 3604–3622. doi:10.1002/hbm.23263.

Baillet, S., Mosher, J.C., Leahy, R.M., 2001. Electromagnetic brain mapping. IEEE Signal Process. Mag. 18 (6), 14–30. doi:10.1109/79.962275.

Bangera, N.B., Schomer, D.L., Dehghani, N., Ulbert, I., Cash, S., Papavasiliou, S., Eisenberg, S.R., Dale, A.M., Halgren, E., 2010. Experimental validation of the influence of white matter anisotropy on the intracranial EEG forward solution. J. Comput. Neurosci. 29 (3), 371–387. doi:10.1007/s10827-009-0205-z.

Baumann, S.B., Wozny, D.R., Kelly, S.K., Meno, F.M., 1997. The electrical conductivity of human cerebrospinal fluid at body temperature. IEEE Trans. Biomed. Eng. 44 (3), 220–223. doi:10.1109/10.554770.

Buzsáki, G., Anastassiou, C.A., Koch, C., 2012. The origin of extracellular fields and currents — EEG, ECog, LFP and spikes. Nat. Rev. Neurosci. 13 (6), 407–420. doi:10.1038/nrn3241.

Camacho, D.L.A., Hopper, R.H., Lin, G.M., Myers, B.S., 1997. An improved method for finite element mesh generation of geometrically complex structures with application to the skullbase. J. Biomech. 30 (10), 1067–1070. doi:10.1016/s0021-9290(97)00073-0.

Chauveau, N., Franceries, X., Doyon, B., Rigaud, B., Morucci, J.P., Celsis, P., 2004. Effects of skull thickness, anisotropy, and inhomogeneity on forward EEG/ERP computations using a spherical three-dimensional resistor mesh model. Hum. Brain Mapp. 21 (2), 86–97. doi:10.1002/hbm.10152.

Conte, S., Richards, J.E., 2021. The influence of the head model conductor on the source localization of auditory evoked potentials. Brain Topogr. 34 (6), 793–812. doi:10.1007/s10548-021-00871-z.

Dahnke, R., Yotter, R.A., Gaser, C., 2013. Cortical thickness and central surface estimation. Neuroimage 65, 336–348. doi:10.1016/j.neuroimage.2012.09.050.

Dalal, S.S., Rampp, S., Willomitzer, F., Ettl, S., 2014. Consequences of EEG electrode position error on ultimate beamformer source reconstruction performance. Front. Neurosci. 8. doi:10.3389/fnins.2014.00042.

Dale, A.M., Fischl, B., Sereno, M.I., 1999. Cortical surface-based analysis. Neuroimage 9 (2), 179–194. doi:10.1006/nimg.1998.0395.

Dannhauer, M., Lanfer, B., Wolters, C.H., Knösche, T.R., 2011. Modeling of the human skull in EEG source analysis. Hum. Brain Mapp. 32 (9), 1383–1399. doi:10.1002/hbm.21114.

Darvas, F., Ermer, J.J., Mosher, J.C., Leahy, R.M., 2006. Generic head models for atlas-based EEG source analysis. Hum. Brain Mapp. 27 (2), 129–143. doi:10.1002/hbm.20171.

Delorme, A., Makeig, S., 2004. EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods 134 (1), 9–21. doi:10.1016/j.jneumeth.2003.10.009.

Farcito, S., Puonti, O., Montanaro, H., Saturnino, G.B., Nielsen, J.D., Madsen, C.G., Siebner, H.R., Neufeld, E., Kuster, N., Lloyd, B.A., Thielscher, A., 2019. Accurate anatomical head segmentations: a data set for biomedical simulations. In: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, pp. 193–208. doi:10.1109/embc.2019.8857041.

Fiederer, L.D.J., Vorwerk, J., Lucka, F., Dannhauer, M., Yang, S., Dümpelmann, M., Schulze-Bonhage, A., Aertsen, A., Speck, O., Wolters, C.H., Ball, T., 2016. The role of blood vessels in high-resolution volume conductor head modeling of EEG. Neuroimage 128, 193–208. doi:10.1016/j.neuroimage.2015.12.041.

Fischl, B., Sereno, M.I., Tootell, R.B., Dale, A.M., 1999. High-resolution intersubject averaging and a coordinate system for the cortical surface. Hum. Brain Mapp. 8, 272–284. doi:10.1002/(sici)1097-0193(1999)8:4<272::aid-hbm10>3.0.co;2-4.

Geuzaine, C., Remacle, J.-F., 2009. Gmsh: a 3-d finite element mesh generator with built-in pre- and post-processing facilities. Int. J. Numer. Methods Eng. 79 (11), 1309–1331. doi:10.1002/nme.2579.

Gramfort, A., 2013. MEG And EEG data analysis with MNE-python. Front. Neurosci. 7. doi:10.3389/fnins.2013.00267.

Gramfort, A., Luessi, M., Larson, E., Engemann, D.A., Strohmeier, D., Brodbeck, C., Parkkonen, L., Hämäläinen, M.S., 2014. MNE Software for processing MEG and EEG data. Neuroimage 86, 446–460. doi:10.1016/j.neuroimage.2013.10.027.

Gramfort, A., Papadopoulo, T., Olivi, E., Clerc, M., 2010. OpenMEEG: opensource software for quasistatic bioelectromagnetics. Biomed. Eng. Online 9 (1), 45. doi:10.1186/1475-925x-9-45.

Hallez, H., Vanrumste, B., Grech, R., Muscat, J., De Clercq, W., Vergult, A., D'Asseler, Y., Camilleri, K.P., Fabri, S.G., Van Huffel, S., 2007. Review on solving the forward problem in EEG source analysis. J. NeuroEng. Rehabil. 4 (1), 46. doi:10.1186/1743-0003-4-46.

Hämäläinen, M., Hari, R., Ilmoniemi, R.J., Knuutila, J., Lounasmaa, O.V., 1993. Magnetoencephalographytextemdashtheory, instrumentation, and applications to noninvasive studies of the working human brain. Rev Mod. Phys. 65 (2), 413–497. doi:10.1103/revmodphys.65.413.

Hämäläinen, M.S., Sarvas, J., 1989. Realistic conductivity geometry model of the human head for interpretation of neuromagnetic data. IEEE Trans. Biomed. Eng. 36 (2), 165–171. doi:10.1109/10.16463.

Haueisen, J., Tuch, D.S., Ramon, C., Schimpf, P.H., Wedeen, V.J., George, J.S., Belliveau, J.W., 2002. The influence of brain tissue anisotropy on human EEG and MEG. Neuroimage 15 (1), 159–166. doi:10.1006/nimg.2001.0962.

Homölle, S., Oostenveld, R., 2019. Using a structured-light 3d scanner to improve EEG source modeling with more accurate electrode positions. J. Neurosci. Methods 326, 108378. doi:10.1016/j.jneumeth.2019.108378.

Hunter, J.D., 2007. Matplotlib: a 2d graphics environment. Comput. Sci. Eng. 9 (3), 90–95. doi:10.1109/mcse.2007.55.

Kaiboriboon, K., Lüders, H.O., Hamaneh, M., Turnbull, J., Lhatoo, S.D., 2012. EEG Source imaging in epilepsy—practicalities and pitfalls. Nature Rev. Neurol. 8 (9), 498–507. doi:10.1038/nrneurol.2012.150.

Kaipio, J., Somersalo, E., 2007. Statistical inverse problems: discretization, model reduction and inverse crimes. J. Comput. Appl. Math. 198 (2), 493–504. doi:10.1016/j.cam.2005.09.027.

Karabanov, A.N., Madsen, K.H., Krohne, L.G., Siebner, H.R., 2021. Does pericentral mu-rhythm "power" corticomotor excitability? – a matter of EEG perspective. Brain Stimul. 14 (3), 713–722. doi:10.1016/j.brs.2021.03.017.

Kirschstein, T., Köhling, R., 2009. What is the source of the EEG? Clin. EEG. Neurosci. 40 (3), 146–149. doi:10.1177/155005940904000305.

Lanfer, B., Scherg, M., Dannhauer, M., Knösche, T.R., Burger, M., Wolters, C.H., 2012. Influences of skull segmentation inaccuracies on EEG source analysis. Neuroimage 62 (1), 418–431. doi:10.1016/j.neuroimage.2012.05.006.

Li, J., Wang, K., Zhu, S., He, B., 2007. Effects of holes on EEG forward solutions using a realistic geometry head model. J. Neural Eng. 4 (3), 197–204. doi:10.1088/1741-2560/4/3/004.

Lu, H., Lam, L.C.W., Ning, Y., 2019. Scalp-to-cortex distance of left primary motor cortex and its computational head model: implications for personalized neuromodulation. CNS Neurosci. Therapeut. 25 (11), 1270–1276. doi:10.1111/cns.13204.

Madsen, K.H., Karabanov, A.N., Krohne, L.G., Safeldt, M.G., Tomasevic, L., Siebner, H.R., 2019. No trace of phase: corticomotor excitability is not tuned by phase of pericentral mu-rhythm. Brain Stimul. 12 (5), 1261–1270. doi:10.1016/j.brs.2019.05.005.

Mahjoory, K., Nikulin, V.V., Botrel, L., Linkenkaer-Hansen, K., Fato, M.M., Haufe, S., 2017. Consistency of EEG source localization and connectivity estimates. Neuroimage 152, 590–601. doi:10.1016/j.neuroimage.2017.02.076.

Makarov, S.N., Hamalainen, M., Okada, Y., Noetscher, G.M., Ahveninen, J., Nummenmaa, A., 2021. Boundary element fast multipole method for enhanced modeling of neurophysiological recordings. IEEE Trans. Biomed. Eng. 68 (1), 308–318. doi:10.1109/tbme.2020.2999271.

Marin, G., Guerin, C., Baillet, S., Garnero, L., Meunier, G., 1998. Influence of skull anisotropy for the forward and inverse problem in EEG: simulation studies using FEM on realistic head models. Hum. Brain Mapp. 6 (4), 250–269. doi:10.1002/(sici)1097-0193(1998)6:4<250::aid-hbm5>3.0.co;2-2.

McCann, H., Beltrachini, L., 2022. Impact of skull sutures, spongiform bone distribution, and aging skull conductivities on the EEG forward and inverse problems. J. Neural Eng. 19 (1), 016014. doi:10.1088/1741-2552/ac43f7.

Medani, T., Garcia-Prieto, J., Tadel, F., Antonakakis, M., Erdbrügger, T., Höltershinken, M., Mead, W., Schrader, S., Joshi, A., Engwer, C., Wolters, C.H., Mosher, J.C., Leahy, R.M., 2023. Brainstorm-DUNEuro: an integrated and user-friendly finite element method for modeling electromagnetic brain activity. Neuroimage 267, 119851. doi:10.1016/j.neuroimage.2022.119851.

Medani, T., Lautru, D., Schwartz, D., Ren, Z., Sou, G., 2015. FEM METHOD FOR THE EEG FORWARD PROBLEM AND IMPROVEMENT BASED ON MODIFICATION OF THE SAINT VENANT'S METHOD. Progr. Electromagnet. Res. 153, 11–22. doi:10.2528/pier15050102.

Meijs, J.W.H., Weier, O.W., Peters, M.J., Van Oosterom, A., 1989. On the numerical accuracy of the boundary element method (EEG application). IEEE Trans. Biomed. Eng. 36 (10), 1038–1049. doi:10.1109/10.40805.

van Mierlo, P., Vorderwülbecke, B.J., Staljanssens, W., Seeck, M., Vulliémoz, S., 2020. Ictal EEG source localization in focal epilepsy: review and future perspectives. Clin. Neurophysiol. 131 (11), 2600–2616. doi:10.1016/j.clinph.2020.08.001.

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 Topogr. 27 (1), 95–111. doi:10.1007/s10548-013-0313-y.

de Munck, J.C., 1992. A linear discretization of the volume conductor boundary integral equation using analytically integrated elements (electrophysiology application). IEEE Trans. Biomed. Eng. 39 (9), 986–990. doi:10.1109/10.256433.

Nguyen-Danse, D.A., Singaravelu, S., Chauvigné, L.A.S., Mottaz, A., Allaman, L., Guggisberg, A.G., 2021. Feasibility of reconstructing source functional connectivity with low-density EEG. Brain Topogr. 34 (6), 709–719. doi:10.1007/s10548-021-00866-w.

Nielsen, J.D., Madsen, K.H., Puonti, O., Siebner, H.R., Bauer, C., Madsen, C.G., Saturnino, G.B., Thielscher, A., 2018. Automatic skull segmentation from MR images

for realistic volume conductor models of the head: assessment of the state-of-the-art. Neuroimage 174, 587–598. doi:10.1016/j.neuroimage.2018.03.001.

Oostendorp, T.F., van Oosterom, A., 1989. Source parameter estimation in inhomogeneous volume conductors of arbitrary shape. IEEE Trans. Biomed. Eng. 36 (3), 382–391. doi:10.1109/10.19859.

Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.-M., 2011. Fieldtrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput. Intell. Neurosci. 2011, 1–9. doi:10.1155/2011/156869.

Puonti, O., Iglesias, J.E., Leemput, K.V., 2016. Fast and sequence-adaptive whole-brain segmentation using parametric bayesian modeling. Neuroimage 143, 235–249. doi:10.1016/j.neuroimage.2016.09.011.

Puonti, O., Leemput, K.V., Saturnino, G.B., Siebner, H.R., Madsen, K.H., Thielscher, A., 2020. Accurate and robust whole-head segmentation from magnetic resonance images for individualized head modeling. Neuroimage 219, 117044. doi:10.1016/j.neuroimage.2020.117044.

Puonti, O., Saturnino, G.B., Madsen, K.H., Thielscher, A., 2020. Value and limitations of intracranial recordings for validating electric field modeling for transcranial brain stimulation. Neuroimage 208, 116431. doi:10.1016/j.neuroimage.2019.116431.

Ramon, C., Schimpf, P.H., Haueisen, J., 2006. Influence of head models on EEG simulations and inverse source localizations. Biomed. Eng. Online 5 (1). doi:10.1186/1475-925x-5-10.

Ruffini, G., et al., 2015. Application of the reciprocity theorem to EEG inversion and optimization of EEG-driven transcranial current stimulation (tCS, including tDCS, tACS, tRNS). 10.48550/ARXIV.1506.04835

Saturnino, G.B., Madsen, K.H., Thielscher, A., 2019. Electric field simulations for transcranial brain stimulation using FEM: an efficient implementation and error analysis. J. Neural. Eng. 16 (6), 066032. doi:10.1088/1741-2552/ab41ba.

Schoffelen, J.-M., Gross, J., 2009. Source connectivity analysis with MEG and EEG. Hum. Brain. Mapp. 30 (6), 1857–1865. doi:10.1002/hbm.20745.

Schrader, S., Westhoff, A., Piastra, M.C., Miinalainen, T., Pursiainen, S., Vorwerk, J., Brinck, H., Wolters, C.H., Engwer, C., 2021. DUNEuro—A software toolbox for forward modeling in bioelectromagnetism. PLoS ONE 16 (6), e0252431. doi:10.1371/journal.pone.0252431.

Shirazi, S.Y., Huang, H.J., 2019. More reliable EEG electrode digitizing methods can reduce source estimation uncertainty, but current methods already accurately identify brodmann areas. Front. Neurosci. 13. doi:10.3389/fnins.2019.01159.

Stenroos, M., Hauk, O., 2013. Minimum-norm cortical source estimation in layered head models is robust against skull conductivity error. Neuroimage 81, 265–272. doi:10.1016/j.neuroimage.2013.04.086.

Sullivan, C., Kaszynski, A., 2019. Pyvista: 3d plotting and mesh analysis through a streamlined interface for the visualization toolkit (VTK). J. Open Source Softw. 4 (37), 1450. doi:10.21105/joss.01450.

Tadel, F., Baillet, S., Mosher, J.C., Pantazis, D., Leahy, R.M., 2011. Brainstorm: a user-friendly application for MEG/EEG analysis. Comput. Intell. Neurosci. 2011, 1–13. doi:10.1155/2011/879716.

Taubin, G., 1995. Curve and surface smoothing without shrinkage. In: Proceedings of IEEE International Conference on Computer Vision. IEEE Comput. Soc. Press, pp. 852–857. doi:10.1109/iccv.1995.466848 .

Thielscher, A., Antunes, A., Saturnino, G.B., 2015. Field modeling for transcranial magnetic stimulation: A useful tool to understand the physiological effects of TMS? In: 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Institute of Electrical & Electronics Engineers (IEEE), pp. 222–225. doi:10.1109/embc.2015.7318340.

Vallaghe, S., Clerc, M., 2009. A global sensitivity analysis of three- and four-layer EEG conductivity models. IEEE Trans. Biomed. Eng. 56 (4), 988–995. doi:10.1109/tbme.2008.2009315.

Vorwerk, J., 2016. New Finite Element Methods to Solve the EEG/MEG Forward Problem. University of Münster phdthesis.

Vorwerk, J., Aydin, U., Wolters, C.H., Butson, C.R., 2019. Influence of head tissue conductivity uncertainties on EEG dipole reconstruction. Front. Neurosci. 13. doi:10.3389/fnins.2019.00531.

Vorwerk, J., Cho, J.-H., Rampp, S., Hamer, H., Knösche, T.R., Wolters, C.H., 2014. A guideline for head volume conductor modeling in EEG and MEG. Neuroimage 100, 590–607. doi:10.1016/j.neuroimage.2014.06.040.

Vorwerk, J., Clerc, M., Burger, M., Wolters, C.H., 2012. Comparison of boundary element and finite element approaches to the EEG forward problem. Biomedical Engineering / Biomedizinische Technik 57 (SI-1 Track-O). doi:10.1515/bmt-2012-4152.

Vorwerk, J., Oostenveld, R., Piastra, M.C., Magyari, L., Wolters, C.H., 2018. The fieldtrip-simbio pipeline for EEG forward solutions. Biomed. Eng. Online 17 (1). doi:10.1186/s12938-018-0463-y.

van der Walt, S., Schönberger, J.L., Nunez-Iglesias, J., Boulogne, F., Warner, J.D., Yager, N., Gouillart, E., Yu, T., 2014. Scikit-image: image processing in python. PeerJ 2, e453. doi:10.7717/peerj.453.

Wang, Y., Gotman, J., 2001. The influence of electrode location errors on EEG dipole source localization with a realistic head model. Clin. Neurophysiol. 112 (9), 1777–1780. doi:10.1016/s1388-2457(01)00594-6.

Weinstein, D., Zhukov, L., Johnson, C., 2000. Lead-field bases for electroencephalography source imaging. Ann. Biomed. Eng. 28 (9), 1059–1065. doi:10.1114/1.1310220.

Wolters, C.H., Anwander, A., Berti, G., Hartmann, U., 2007. Geometry-adapted hexahedral meshes improve accuracy of finite-element-method-based EEG source analysis. IEEE Trans. Biomed. Eng. 54 (8), 1446–1453. doi:10.1109/tbme.2007.890736.

Wolters, C.H., Anwander, A., Tricoche, X., Weinstein, D., Koch, M.A., MacLeod, R.S., 2006. Influence of tissue conductivity anisotropy on EEG/MEG field and return current computation in a realistic head model: asimulation and visualization study using high-resolution finite element modeling. Neuroimage 30 (3), 813–826. doi:10.1016/j.neuroimage.2005.10.014.

Wolters, C.H., Grasedyck, L., Hackbusch, W., 2004. Efficient computation of lead field bases and influence matrix for the FEM-based EEG and MEG inverse problem. Inverse Probl. 20 (4), 1099–1116. doi:10.1088/0266-5611/20/4/007.

Zienkiewicz, O.C., Zhu, J.Z., 1992. The superconvergent patch recovery anda posteriori error estimates. part 1: the recovery technique. Int. J. Numer. Methods. Eng. 33 (7), 1331–1364. doi:10.1002/nme.1620330702.