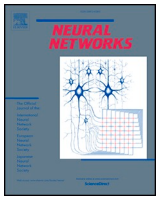
Neural Networks 180 (2024) 106731

Contents lists available at ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet



Localized estimation of event-related neural source activity from simultaneous MEG-EEG with a recurrent neural network

Jamie A. O'Reilly [a,*], Judy D. Zhu [b], Paul F. Sowman [b,c]

[a] School of Engineering, King Mongkut's Institute of Technology Ladkrabang, Bangkok, 10520, Thailand
[b] School of Psychological Sciences, Macquarie University, New South Wales, 2109, Australia
[c] School of Clinical Sciences, Auckland University of Technology, Auckland, 1142, New Zealand

ARTICLE INFO

Keywords:
Artificial neural networks
Biomedical signal processing
Computational neuroscience
Deep learning
Electroencephalography (EEG)
Neural source reconstruction
Magnetoencephalography (MEG)

ABSTRACT

Estimating intracranial current sources underlying the electromagnetic signals observed from extracranial sensors is a perennial challenge in non-invasive neuroimaging. Established solutions to this inverse problem treat time samples independently without considering the temporal dynamics of event-related brain processes.

This paper describes current source estimation from simultaneously recorded magneto- and electroencephalography (MEEG) using a recurrent neural network (RNN) that learns sequential relationships from neural data. The RNN was trained in two phases: (1) pre-training and (2) transfer learning with L1 regularization applied to the source estimation layer. Performance of using scaled labels derived from MEEG, magnetoencephalography (MEG), or electroencephalography (EEG) were compared, as were results from volumetric source space with free dipole orientation and surface source space with fixed dipole orientation. Exact low-resolution electromagnetic tomography (eLORETA) and mixed-norm L1/L2 (MxNE) source estimation methods were also applied to these data for comparison with the RNN method.

The RNN approach outperformed other methods in terms of output signal-to-noise ratio, correlation and mean-squared error metrics evaluated against reference event-related field (ERF) and event-related potential (ERP) waveforms. Using MEEG labels with fixed-orientation surface sources produced the most consistent estimates.

To estimate sources of ERF and ERP waveforms, the RNN generates temporal dynamics within its internal computational units, driven by sequential structure in neural data used as training labels. It thus provides a data-driven model of computational transformations from psychophysiological events into corresponding event-related neural signals, which is unique among MEEG source reconstruction solutions.

1. Introduction

A goal of non-invasive neuroimaging is to localize sources of brain activity observed from outside the head. Localizing sources of magnetoencephalography (MEG) and electroencephalography (EEG) signals is known as the inverse problem, which is ill-posed because there are theoretically infinite compositions of source signals that can explain observed MEG or EEG signals (Nunez & Srinivasan, 2009). Due to the non-uniqueness of inverse problem solutions, constraints based on biophysics or numerical methods are employed to reduce the search space of possible solutions. The forward problem, estimating MEG and EEG signal amplitudes from an observed brain source signal, is comparatively well posed and can be estimated with knowledge of the physical conduction of electromagnetic phenomena through biological tissues (Nunez & Srinivasan, 2009). However, brain source signals are generally inaccessible, making solutions to the inverse problem desirable for estimating brain source signals that account for observed MEG and EEG signals. A *lead field* matrix of coefficients can be determined from a forward conduction model, which describes the contributions of currents in the brain to MEG and EEG sensors if head and sensor geometry are known (Hämäläinen & Ilmoniemi, 1994). The lead field matrix provides a fundamental constraint for inverse solutions. Using simultaneously recorded MEG and EEG (i.e., MEEG) constrains the space of solutions to the inverse problem, which is considered superior to using either MEG or EEG on their own (Dale et al., 1993; Luck, 2014; Nunez & Srinivasan, 2009).

* Corresponding author at: School of Engineering, King Mongkut's Institute of Technology Ladkrabang, 1 Chalong Krung 1 Alley, Lat Krabang, Bangkok, 10520, Thailand.
E-mail address: jamie.or@kmitl.ac.th (J.A. O'Reilly).

https://doi.org/10.1016/j.neunet.2024.106731
Received 13 March 2024; Received in revised form 5 September 2024; Accepted 10 September 2024
Available online 11 September 2024

Inverse solutions are optimization problems. Equivalent current dipole fitting algorithms aim to determine the optimal location, orientation and moment of one or more dipoles that can account for observed data (Ahlfors & Hämäläinen, 2012). Drawbacks of dipole fitting methods are dependence of solutions on initial conditions and assuming a low number of active dipoles. The latter drawback of dipole fitting methods can be overcome by distributed source space methods, which attempt to estimate signals at many locations (referred to as vertices) in parallel (Ahlfors & Hämäläinen, 2012). Pseudoinverse matrices can be computed from the lead field matrix using methods from linear algebra, such as minimum-norm estimates (MNE; Hämäläinen & Ilmoniemi, 1994) and more advanced variants such as exact low-resolution electromagnetic tomography (eLORETA; Pascual-Marqui et al., 2011), that include different constraints (e.g., L2 norm in conventional MNE) to solve the inverse problem. For instance, the L1/L2 mixed-norm estimates (MxNE) method includes an L1 norm penalty to obtain a sparse solution to the inverse problem (Gramfort et al., 2012). However, the rationale and reliability of these numerical constraints for determining the true distribution of brain currents remains open for debate (Luck, 2014).

Methods of MEG and EEG source localization have also arrived from Bayesian probability theory (Das et al., 2020; David et al., 2006; Henson et al., 2010; Henson, Mouchlianitis et al., 2009; Wipf & Nagarajan, 2009). These have incorporated different priors to constrain inverse solutions, such as preconceived network topologies and interactions (Kiebel et al., 2006), constraints derived from function magnetic resonance imaging data (Henson et al., 2010), and deconvolving signals using a bank of linear "neuro-current response functions" (Das et al., 2020). Several methods have also emerged from recent developments in deep learning with artificial neural networks, reportedly having equivalent or superior performance to conventional linear inverse solutions (Hecker et al., 2021; Liang et al., 2023; Pantazis & Adler, 2021; Sun et al., 2022). However, these methods do not extract temporal structure from neural data because they are fitted to an individual or small number of time points, which limits the extent to which resulting solutions can be interpreted.

Recurrent neural network (RNN) models are designed to learn sequential relationships from data, potentially making them more useful for analysing neural time series than other artificial neural networks. Simple RNNs have been used to model event-related neural signals from mice exposed to auditory stimulation (O'Reilly, 2022a, 2022) and humans exposed to auditory and visual stimuli (O'Reilly, 2022b, 2023). A long short-term memory (LSTM; Hochreiter & Schmidhuber, 1997) type of RNN has been used to refine MEG and EEG source estimates obtained by applying MNE to individual time samples, highlighting the importance of temporal context among neural signals for improving source reconstruction (Dinh et al., 2021). Simple RNN models have also been developed for estimating localized source signals from MEG (O'Reilly, Zhu et al., 2023) and EEG (Srivastava et al., 2023) by fixing the weights of a feed-forward output layer equal to the lead field matrix. By training the RNN in two steps: (1) pretraining and (2) transfer learning with L1 regularization, estimated source signals are derived from penultimate layer hidden unit activations (O'Reilly, Zhu et al., 2023).

In this study, the RNN method of neural current source localization was developed for and applied to simultaneously-recorded MEEG signals. Estimated source signals were obtained from MEEG, MEG, and EEG data, and two linear inverse solutions (eLORETA and MxNE) were applied to the same data for comparison. Furthermore, volumetric source space with free dipole orientation and surface source space with fixed surface-normal orientation were evaluated. The results of this study will inform future developments of RNN models of event-related neural signal processing.

2. Materials and methods

2.1. MEEG recordings

MEG and EEG were simultaneously recorded at the KIT-Macquarie Brain Research Laboratory in Sydney, Australia, using the equipment described in a recent article (Zhu et al., 2022). During the recording, the subject listened to "ba" and "da" syllable sounds played in a random sequence. This experiment was part of a larger study approved by the Human Research Ethics Committee at Macquarie University. Before recording signals, an EEG cap of appropriate size was fitted to the subject's head, and five head position marker coils were attached to the cap (to locate the head position once inside the MEG sensor array). The subject's head shape information was obtained by 3D head scanning using the Structure Sensor accessory (Occipital, Inc., Boulder, CO) on an iPad, and the positions of fiducials and marker coils were marked manually. The sequence of sounds in the experiment was generated using custom code in Matlab (version R2017b; The MathWorks Inc., Natick, MA, USA) and delivered binaurally via pneumatic earphones (Etymotic ER30; Etymotic Research, Inc., Elk Grove Village, IL, USA).

The MEG data were recorded using a system with 160 axial gradiometers (Model PQ1160R-N2; Kanazawa Institute of Technology, Kanazawa, Japan) located inside a magnetically shielded room (Fujihara Co. Ltd., Tokyo, Japan). Signals were recorded from three additional reference sensors, which were used for post-hoc denoising of MEG signals using time-shift regression (de Cheveigné & Simon, 2007). Audio signals recorded synchronously with MEG were used to obtain accurate onset timing for the auditory stimuli. The subject's head position was measured before and after the experimental task and it deviated by <5 mm. Signals were acquired from the MEG system at 1000 samples per second with online band-pass filtering from 0.03 to 200 Hz.

The EEG data were recorded using an MEG-compatible 64-channel system (BrainAmp; Brain Products, Germany). The EEG cap (BrainCap-MEG; EasyCap, Germany) contained 61 EEG channels, placed according to the 10–20 system (Fp1, AF3, AF7, Fz, F1, F3, F5, F7, FC1, FC3, FC5, FT7, Cz, C1, C3, C5, T7, CP1, CP3, CP5, TP7, TP9, Pz, P1, P3, P5, P7, PO3, PO7, Oz, O1, Fpz, Fp2, AF4, AF8, F2, F4, F6, F8, FC2, FC4, FC6, FT8, C2, C4, C6, T8, CPz, CP2, CP4, CP6, TP8, TP10, P2, P4, P6, P8, POz, PO4, PO8, and O2), plus EOG, ECG, and a reference channel at FCz. Data were re-referenced offline to the average montage. An affine transform aligned default EEG channel locations with fiducials (nasion and bilateral preauricular areas) and head landmarks (vertex and occipital point) obtained from head digitization data. Channel location and head digitization alignments are plotted in Figure S1. Fiducials and head shape points were also co-registered with a template head MRI to compute forward conduction models for both MEG and EEG, as described in Section 2.2.

Independent component analysis (ICA) was used to correct artifacts related to ocular, cardiological, and non-biological sources from both signal modalities (Jung et al., 2000); three components were removed from MEG, and six components were removed from EEG. EEG and MEG were band-pass filtered from 1 to 40 Hz, and resampled to 100 Hz, before extracting epochs from -0.1 to 0.4 s relative to stimulus onset on each trial. Baseline correction was applied using the pre-stimulus interval. The resulting MEEG data matrix comprised 99 "ba" epochs and 98 "da" epochs, $\boldsymbol{Y}_{MEEG} \in \mathbb{R}^{197 \times 51 \times 221}$ (total 197 epochs, 51 time-samples, and 221 channels); MEG and EEG submatrices were extracted from this, where $\boldsymbol{Y}_{MEG} \in \mathbb{R}^{197 \times 51 \times 160}$ and $\boldsymbol{Y}_{EEG} \in \mathbb{R}^{197 \times 51 \times 61}$. These data matrices were scaled and used as labels for training RNN models.

2.2. Source spaces and forward models

Volume and surface source spaces were used, and the results from both were compared to assess the reproducibility of source signals

estimated by either approach using the RNN method. The volume source space had 30.0 mm grid spacing with 3.0 mm minimum distance from the inner skull; producing 50 vertices, each with three components for x, y, and z directions. The surface source space had "oct3" octahedron decimation, producing 66 vertices in each hemisphere, with orientation fixed to the cortical surface normal. While not matched exactly, these two source spaces provided similar degrees of freedom for fitting observed signals; 150 for the volume space and 132 for the surface space. More degrees of freedom might be assumed to yield better-fitting solutions to the data if that were the only difference between the two models. The average minimum distance between each vertex and its nearest neighbour was 30.0 mm for the volume source space and 12.9 mm for the surface source space. The location of each vertex is shown alongside estimated source waveforms in supplementary figures available from https://osf.io/dfr52.

A three-layer boundary element method (BEM) conduction model was computed from the co-registered template head, with brain, skull, and scalp conductance of 0.3, 0.006, and 0.3 S m^{-1}, respectively. Forward solutions for volume and surface source spaces were derived from this conduction model, with lead field coefficients calculated for 160 MEG and 61 EEG channels. All of these calculations were performed using MNE-Python software (Gramfort et al., 2013). The combined volume source space lead field matrix can be denoted $L_{V_{MEEG}} \in \mathbb{R}^{221 \times 150}$, with 221 channels and 50 × 3 source signal components (x, y, and z directions for each vertex). Similarly, the combined surface source space, $L_{S_{MEEG}} \in \mathbb{R}^{221 \times 132}$, accounted for 66 sources on each hemisphere with fixed surface-normal orientation. Standardized (Z-scored) lead field coefficient distributions for MEG and EEG channels in volume and surface source spaces are plotted in Figure S2.

2.3. RNN inputs and labels

Fig. 1 illustrates how an RNN was used to model ERF and ERP waveforms derived from simultaneously recorded MEEG. Inputs consisted of two signal channels: one representing "ba" and another representing "da" events; $\boldsymbol{x} \in \mathbb{R}^{51 \times 2}$. A unit step-pulse set high from 0.0 to 0.075 s on the respective channel signalled that event. Each input was paired with a single epoch of MEEG data recorded in response to the associated event (i.e. $\boldsymbol{y} \in \mathbb{R}^{51 \times 221}$), which provided labels for supervised learning. All of the single epochs in the data matrix were paired with one of two input representations, producing an input matrix, $\boldsymbol{X} \in \mathbb{R}^{197 \times 51 \times 2}$. These pairings effected a one-to-many mapping of input (encoding event type) to output label (single-trials of brain signals) time sequences, where the RNN models the function: $\boldsymbol{X} \rightarrow \boldsymbol{Y}$. As an artificial neural network minimizes mean-squared error (MSE) loss with a one-to-many relation between its inputs and labels, its output will

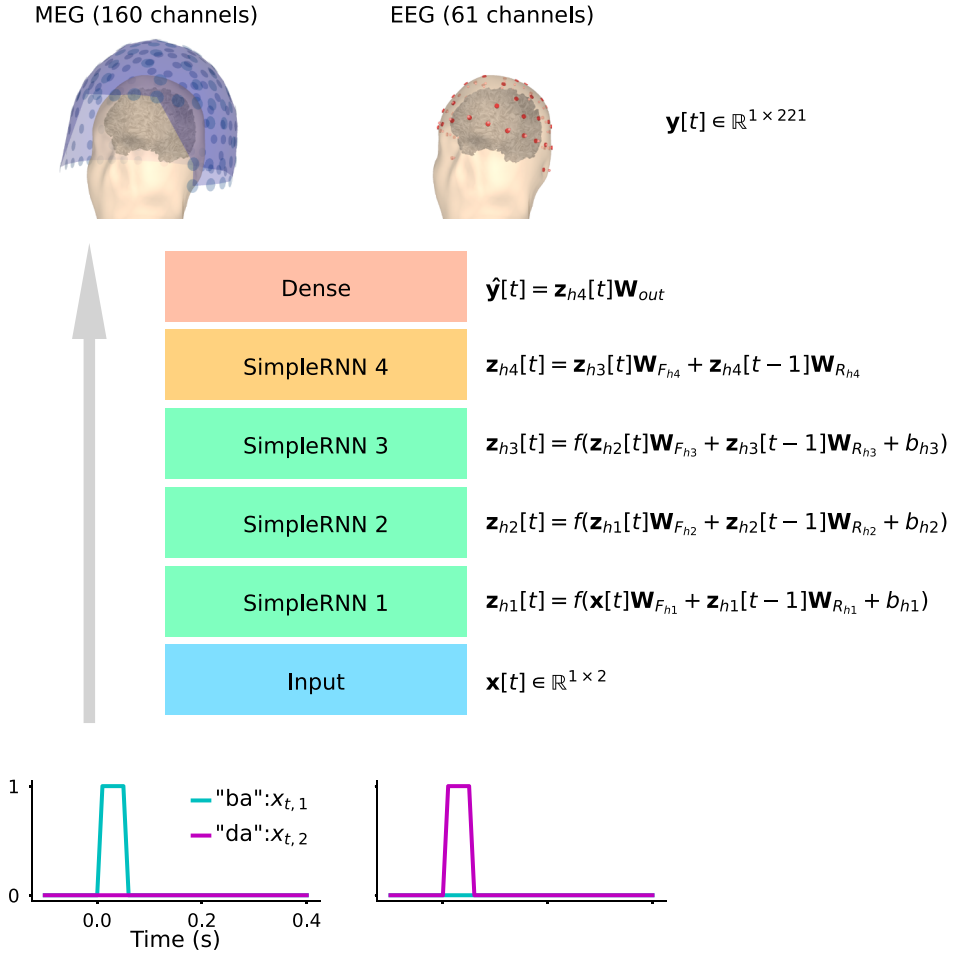


Fig. 1. Model architecture and training scheme for using an RNN to estimate sources of MEEG data. Two sets of step pulse functions (bottom) representing "ba" (left) and "da" (right) events were input to the RNN. These inputs were transformed through sequential recurrent layers (SimpleRNN 1–4), the first three with rectified linear unit activation function, the fourth with linear activation, and then a single feed-forward (Dense) layer with linear activation. The fourth SimpleRNN layer was trained without a bias and had L1-norm activity regularization when the model was fine-tuned. The layer naming convention adopted here is consistent with the TensorFlow API. Labels for supervised learning were obtained from simultaneously recorded MEG (top-left) and EEG (top-right). The final layer of this model effectively implements the forward solution. Thus, outputs from the penultimate layer ($\mathbf{z}_{h4}$) can be interpreted as source signals. Mathematical notations illustrate the data entities and computations at each stage in the modelling process for an individual time sample.

converge towards the average of those labels associated with each unique event. This behaviour is desirable in the context of modelling averaged event-related neural signals. The same approach was used to model waveforms from MEG-only and EEG-only by setting the model output weights ($\boldsymbol{W}_{out}$) using the appropriate lead field matrix and corresponding single-modality data as labels, described in Section 2.5.

2.4. RNN architecture and training parameters

As with previous RNN models of event-related neural signals (O'Reilly, 2022b, 2022, 2022; O'Reilly, Zhu et al., 2023), the architecture consisted of an input, four hidden layers, and an output layer. This basic architecture effectively reproduces ERP and ERF waveforms; however, whether it is optimal compared with alternative designs remains an open question. Time-domain signals were passed to the input layer (*Input*), transformed sequentially through four simple recurrent layers (*SimpleRNN 1–4*), and fed through a feed-forward output layer (*Dense*). Mathematical objects and operations at each layer are annotated on the diagram in Fig. 1. The first three recurrent layers had rectified linear unit activation function. The fourth recurrent layer, which generates estimated source signals, had linear activation. The output layer also had linear activation and effectively implemented the forward computation: *signals* = *sources* × *lead field*, formulated as $\widehat{\boldsymbol{y}} = \boldsymbol{z}_{h4}\boldsymbol{W}_{out}$, where $\widehat{\boldsymbol{y}}$ and $\boldsymbol{W}_{out}$ were obtained by scaling data labels and corresponding lead field coefficients, respectively, as explained in Section 2.5.

The backpropagation through time (BPTT) algorithm (Werbos, 1990) with adaptive moment estimation (Adam) optimizer (Kingma & Ba, 2015) and mean-squared error loss were used to train the RNN, all using default hyperparameter settings in TensorFlow (Abadi et al., 2016). As described previously (O'Reilly, Zhu et al., 2023; Srivastava et al., 2023), to derive source estimates, the RNN was pre-trained to fit labels without additional constraints (training step 1) before discarding the bias parameter and introducing L1-norm activity regularization to the fourth recurrent layer (training step 2). This L1-norm penalty was weighted by 10^{-4}. In both training steps the maximum number of training epochs/iterations was set to 5000. However, training ceased if the loss failed to improve for 50 consecutive epochs; all models stopped before reaching the maximum (Figure S3). A batch size of 197 was selected to include all training instances. Due to stochastic variation in network initialization and optimization algorithms, models were trained five times using different random seeds (integers 0 to 4).

2.5. Scaling data and lead field

MEG and EEG data were scaled before being used as labels to avoid the vanishing gradients problem (Bengio et al., 1994). Magnetic fields measured from the brain are commonly reported on the order of femtoteslas (fT = × 10^{-15} T). In contrast, potential differences due to electric fields of the brain are observed on the order of microvolts (μV = × 10^{-6} V). Different orders of magnitude reflect different signal conduction and acquisition physics for magnetic and electric fields, although both signals arise from biophysical processes operating on the same scale. That is to say, if synchronous brain processes are considered to generate an electric dipole that can be measured from outside the head by EEG and/or MEG, the same dipole current source density produces EEG signals in the μV range and MEG signals in the fT range. Therefore, when scaling MEG and EEG labels for supervised learning, their corresponding lead field coefficients were also scaled to ensure consistent data-to-lead field scale ratios for both. Four scaling parameters (c_{MEG}, c_{EEG}, $c_{L_{MEG}}$, $c_{L_{EEG}}$) were used to accomplish this, maintaining the relations: $c_{MEG}/c_{L_{MEG}} = c_{EEG}/c_{L_{EEG}} = c_{src} = 10^7$, which ensured that estimated source signals for MEG and EEG would have a common scale, and could be transformed into biophysically plausible estimates by dividing by c_{src}.

Five sets of these scaling parameters were devised to adjust the balance between MEG and EEG data labels, as outlined in Table S1. The first set of scaling parameters was selected based on previous single-modality studies (O'Reilly, Zhu et al., 2023; Srivastava et al., 2023) to avoid vanishing gradients. However, these produced an imbalance between MEG and EEG data, with a ratio of mean absolute values $\langle|\boldsymbol{Y}_{MEG}|\rangle/\langle|\boldsymbol{Y}_{EEG}|\rangle = 1.06/0.541$. The second set of scaling parameters was selected to partly redress this imbalance, producing a ratio of $\langle|\boldsymbol{Y}_{MEG}|\rangle/\langle|\boldsymbol{Y}_{EEG}|\rangle = 1.06/0.811$. The third set was calculated to balance this ratio, i.e., $\langle|\boldsymbol{Y}_{MEG}|\rangle/\langle|\boldsymbol{Y}_{EEG}|\rangle = 1$. The fourth set was calculated to balance the sum of MEG and EEG absolute values, i.e., $\sum|\boldsymbol{Y}_{MEG}|/\sum|\boldsymbol{Y}_{EEG}| = 1$. The fifth was calculated to balance the standard deviation of MEG and EEG data labels, i.e., $\sigma_{MEG}/\sigma_{EEG} = 1$.

While these five sets of scaling parameters influenced the performance of RNN models in terms of comparison between reference and reconstructed ERF and ERP waveforms (Table S2 and Figure S4), the relatively small variation (i.e., all had $r > 0.89$) did not provide strong evidence for preferring source signals estimated using either one set of scaling values. Weighting EEG more than MEG labels (scaling set 4) biased the model towards the former, enhancing the overall correlation between reference ERP/ERF and reconstructed waveforms; however, this evaluation is confounded by higher inter-channel correlation for EEG than MEG (Figure S5). Therefore, to calculate the final estimated source signals from the RNN method, those produced by all five sets of scaling values and five random seeds (i.e., 25 estimates) were averaged. ERF and ERP waveforms reconstructed from these RNN source estimates are plotted in Fig. 2.

Lead field matrices $L_{MEG} \in \mathbb{R}^{160\times n_{src}}$ and $L_{EEG} \in \mathbb{R}^{61\times n_{src}}$ were used to fit MEG and EEG data, respectively, where $n_{src} = 150$ (50 sources × 3 directions) for volume source space, and 132 for surface source space. Lead field matrices L_{MEG} and L_{EEG} were concatenated for MEEG; i.e., $L_{MEEG} = [L_{MEG}, L_{EEG}]$. Output layer weights ($\boldsymbol{W}_{out}$) in Fig. 1 were constructed from scaled lead field matrices, $c_{L_{MEG}}L_{MEG}$ and $c_{L_{EEG}}L_{EEG}$, and held fixed during training. Scaled data, $c_{MEG}\boldsymbol{Y}_{MEG}$ and $c_{EEG}\boldsymbol{Y}_{EEG}$ were provided as labels for supervised learning; i.e., $\boldsymbol{Y} = [c_{MEG}\boldsymbol{Y}_{MEG},\ c_{EEG}\boldsymbol{Y}_{EEG}] \in \mathbb{R}^{197\times 51\times 221}$ for MEEG data, $\boldsymbol{Y} = c_{MEG}\boldsymbol{Y}_{MEG} \in \mathbb{R}^{197\times 51\times 160}$ for MEG data, and $\boldsymbol{Y} = c_{EEG}\boldsymbol{Y}_{EEG} \in \mathbb{R}^{197\times 51\times 61}$ for EEG data. In each case, estimated source signals were extracted from the fourth recurrent layer hidden unit activations (i.e., $\widehat{sources} = \boldsymbol{z}_{h4}/c_{src}$). The number of hidden units in the SimpleRNN 4 layer was thus n_{src}, and the number of hidden units in SimpleRNN 1–3 layers was set to $2n_{src}$.

2.6. eLORETA and MxNE

Two established inverse methods were computed using averaged ERF and ERP responses to "ba" and "da" events. Functions in MNE-python were used to implement these. Noise sample covariance was estimated from the pre-stimulus baseline, and data sample covariance was estimated from the post-stimulus period of single-trial data. For the volume source space, the *depth* weight prior was 0.8, and the variance of source components tangential to the cortical surface was weighted by the *loose* parameter 1.0 for free orientation. There was no *depth* prior weighting for the surface source space, and the *loose* parameter was set to zero for fixed orientation.

Exact low-resolution electromagnetic tomography (eLORETA) is claimed to have zero localization error for EEG inverse problems under noiseless conditions (Pascual-Marqui et al., 2011). In reality, there is no such ideal situation, but eLORETA is widely used and regarded highly among linear inverse methods for EEG and MEG source reconstruction (Pascual-Marqui et al., 2018). The MNE-python implementation of eLORETA was applied with the apply_inverse function. Recommended default parameters were used, and the *lambda2* regularization parameter was set to $1/3^2$, where 3 was the estimated signal-to-noise ratio.

Mixed-norm estimates (MxNE) combine L1 and L2 norm constraints while solving the inverse problem, yielding sparse solutions (Gramfort

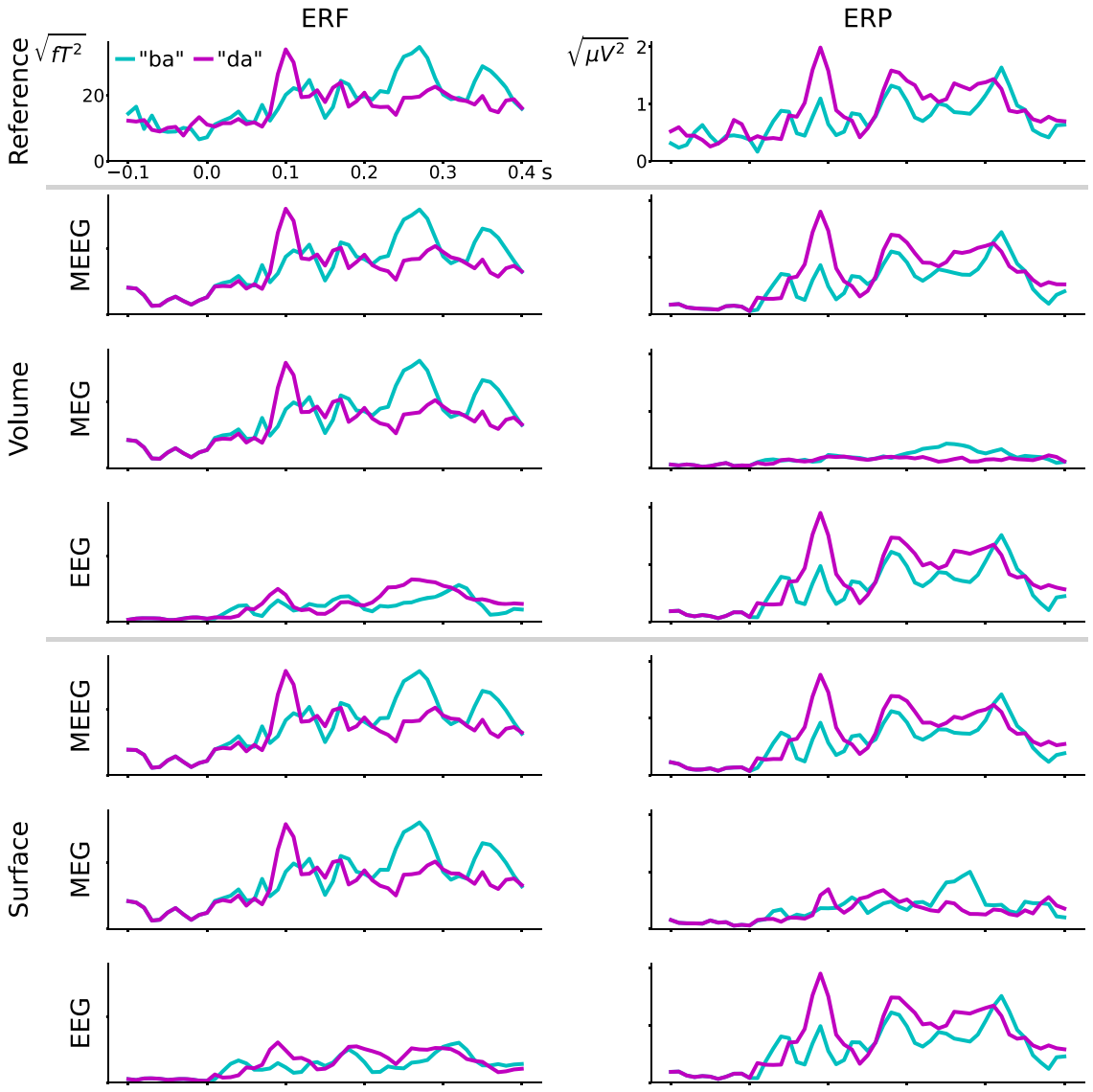


Fig. 2. ERF and ERP waveforms reconstructed from source signals estimated from RNN. Root-mean-squared amplitudes across all MEG or EEG channels (for ERF or ERP waveforms, respectively) are plotted on left and right columns. Y-axis scales are the same for each row, and x-axis scales are the same for all plots. Results from volume and surface source spaces are comparable. When fitted to unimodal data (MEG or EEG), reconstructed waveforms for the alternate modality are comparatively poor; consistent with results from eLORETA and MxNE presented in Figure S6 and Figure S7. Analyses of these waveforms are provided in Table 1.

et al., 2012). We selected MxNE to compare with the RNN approach because both methods incorporate L1 regularization of estimated source signals. It was implemented using the mixed_norm function in MNE-python with the regularization parameter *alpha* 5.0; otherwise, recommended default parameters were selected.

2.7. Source ranking analysis

Estimated source signals were ranked in reverse order by correlating their projected ERF and ERP waveforms with the reference ERF/ERP waveforms. Individual source signals were multiplied by their lead field coefficients to project them onto sensor space. Then, these projections were compared with reference ERF and ERP waveforms using Pearson's correlation coefficient. For each volume source, estimated signals for three direction components were multiplied by their corresponding lead field coefficients to generate projected waveforms. For both source spaces, each vertex received nine rankings [i.e., from three methods (RNN, eLORETA, MxNE) applied to three data types (MEEG, MEG, EEG)]. These aggregated rankings were used to evaluate the strength of evidence for each vertex's involvement in generating ERF and ERP waveforms. Only the two most highly ranked sources are plotted (Figs. 3 and 4), although all are available from an online repository.

The consistency of signals estimated by the RNN method was also quantified by computing the correlation between source waveforms estimated for the same vertex by multiple RNNs. For each source signal component, the average correlation among estimates from 25 RNN models (five sets of scaling parameters, five random seeds) after training without (step 1) and with L1-norm regularization (step 2) were evaluated for each data type (MEEG, MEG, EEG) and source space (volume, surface). For this analysis, three direction components for volume source space vertices were treated separately. Distributions of these average correlations are plotted in Fig. 5, and their descriptive statistics are provided in Table S8. Higher correlation among signals estimated for the same vertex reflects convergence towards the accurate ERP/ERF source reconstruction solution.

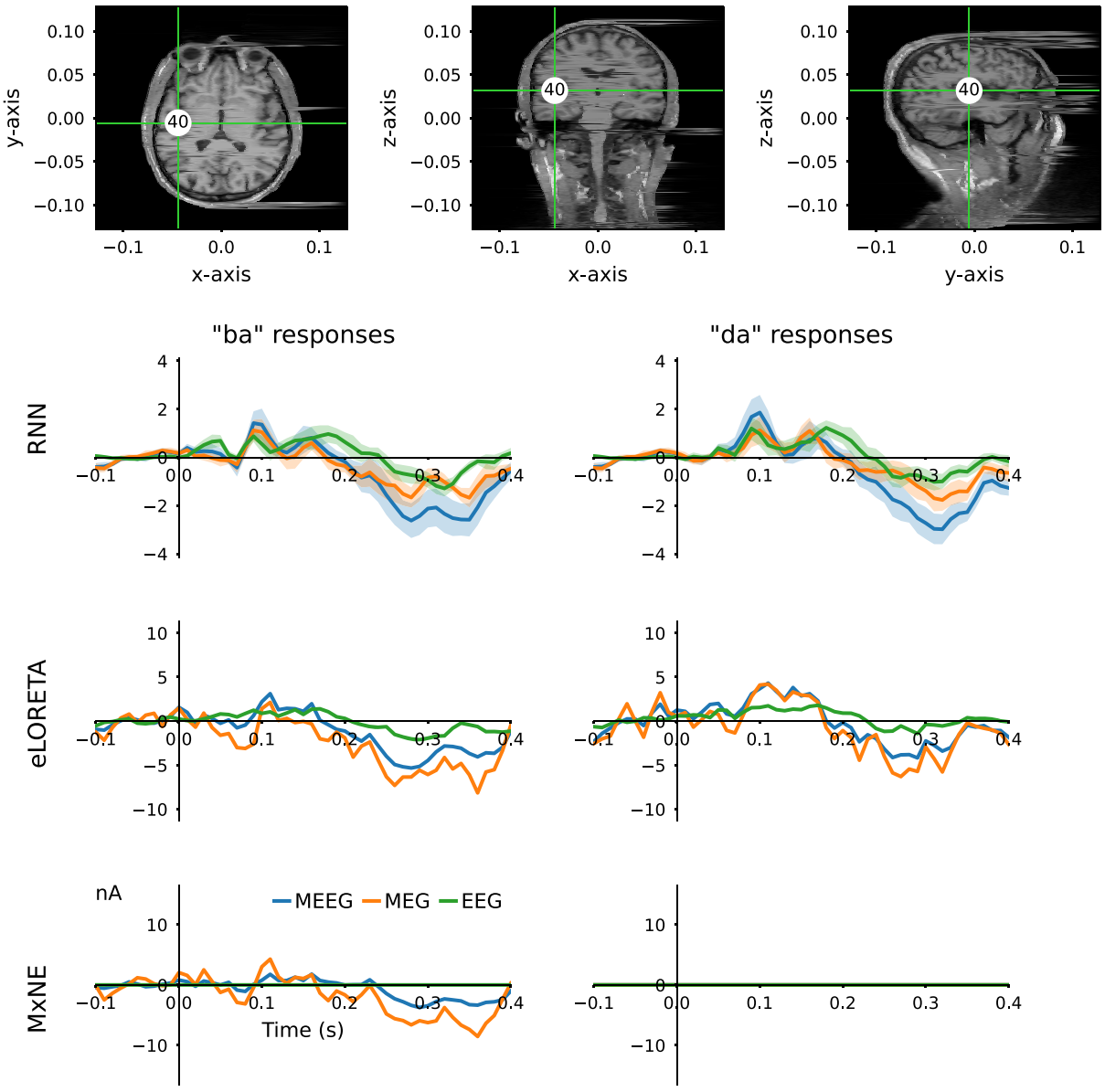


Fig. 3. Estimated signals from vertex 40 of the surface source space with fixed dipole orientation normal to the cortical surface. When projected on to MEEG channels using the lead-field matrix, this source was ranked among the top five in terms of correlation between real and projected MEEG for 8 out of 9 source estimates (i.e., three methods [RNN, eLORETA, MxNE] times three data types [MEEG, MEG, EEG]). This source is located in the left temporal lobe, consistent with regions involved in sound and language processing. Shaded error bars represent the standard deviation from 25 source estimates obtained with the RNN method. Different y-axis scales are used for plotting estimated source signals from different methods.

2.8. Data analysis and statistical tests

Reconstructed ERF and ERP waveforms were compared with reference waveforms using Pearson's correlation coefficient (r) and mean squared error (MSE). Signal to noise ratio (SNR) was calculated in decibels (dB) using the formula $\text{SNR} = 20\log_{10}(\text{signal}_{\text{RMS}}/\text{noise}_{\text{RMS}})$, where $\text{signal}_{\text{RMS}}$ was the root-mean-squared (RMS) amplitude from the post-stimulus period (>0.0 s) and $\text{noise}_{\text{RMS}}$ was the RMS amplitude in the pre-stimulus period (<0.0 s). The SNR of reconstructed waveforms and estimated source signals (SRC SNR) were evaluated. These data are reported in Table 1.

Non-parametric statistical hypothesis tests were performed to assess whether differences among samples were statistically significant because they either did not have equal variance and/or did not have a normal distribution. The Brown-Forsythe test was used to check for equal variance (Brown & Forsythe, 1974), and D'Agostino and Pearson's method was used to test for normality (D'Agostino & Pearson, 1973). Paired samples were tested with Wilcoxon signed-rank tests (Wilcoxon, 1945) and independent samples with Mann-Whitney *U tests* (Mann & Whitney, 1947), with two-sided alternatives. Descriptive statistics are presented in supplementary tables, and inferential statistics are reported in square brackets, e.g., [statistic, p-value], with p-values below 0.05 considered statistically significant.

2.9. Experiments with higher-density surface source spaces

After initially evaluating the RNN method with the relatively sparse source spaces described in Section 2.2, a series of experiments with denser surface source spaces was performed. In the mne.setup_source_space function, *spacing* parameters of "oct4", "ico3", "oct5", and "ico4" were used to generate cortical source spaces with 516, 1282, 2052, and 5120 vertices, respectively, each with fixed surface normal orientation. Otherwise, forward model computations were performed as described in Section 2.2. Lead-field matrices, EEG and MEG were scaled to balance

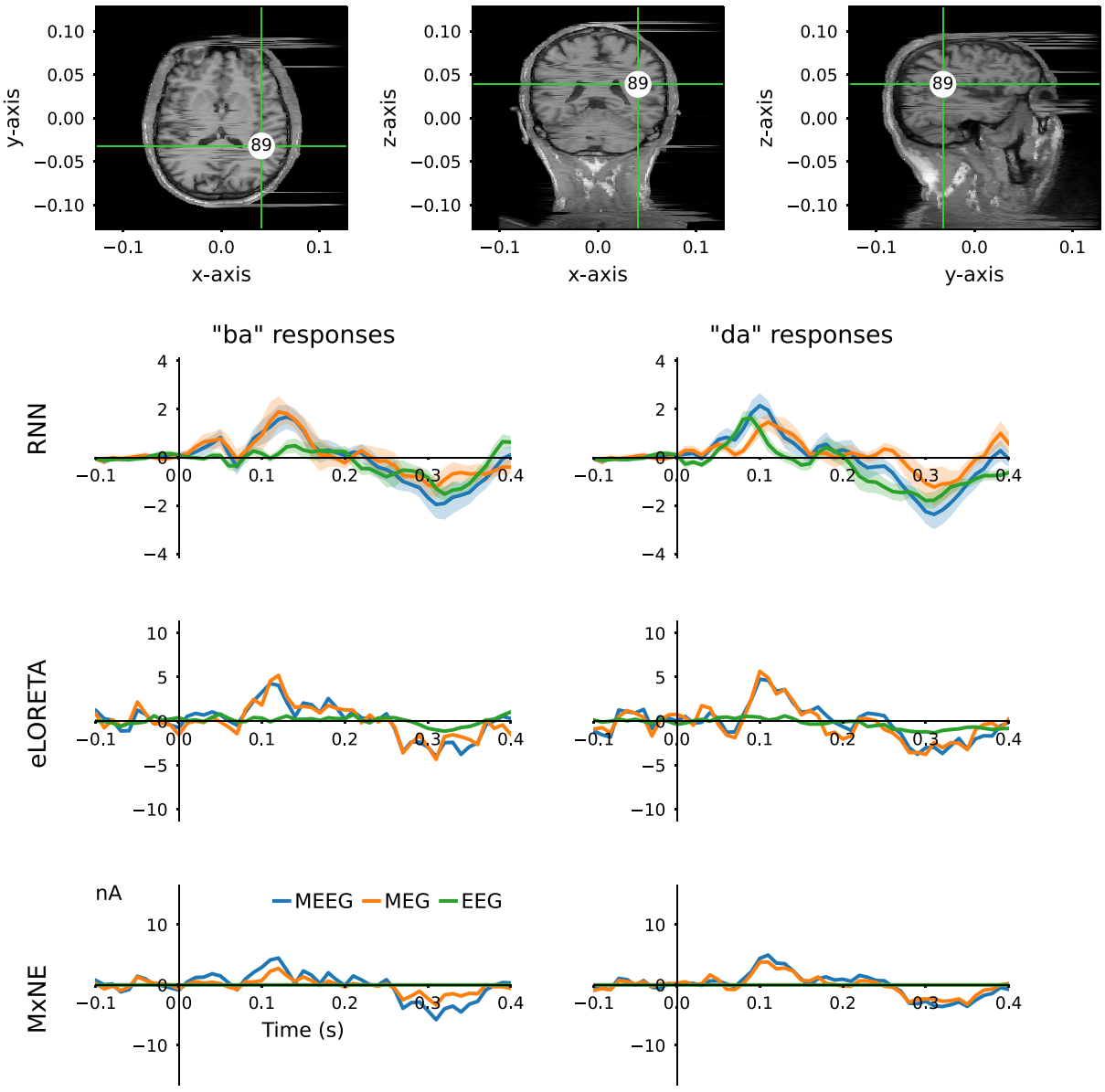


Fig. 4. Source signals estimated from vertex 89 of the fixed-orientation surface source space. This source also ranked within the top five with the highest correlation between real and projected MEEG for 8 out of 9 source estimates. Situated in the right posterior temporal lobe, this is consistent with expectations for auditory processing of syllable sounds. Standard deviations of RNN source estimates are denoted with shared error bars. Y-axes scales are different for each method. Comparable figures for all estimated source signals are available from https://osf.io/dfr52.

absolute values of EEG and MEG labels (like the fourth set described in Section 2.5) to prepare RNN output layer weights and training data.

Critically, the RNN architecture and optimization strategy were modified to accommodate these denser source spaces. The numbers of hidden units grew in proportion to the number of source space vertices; SimpleRNN 4 having n_{src} units, and SimpleRNN 1–3 having $2n_{src}$ units, for n_{src} in {516, 1282, 2052}. For the 5120-vertex source space, proportionally growing the model in this manner exceeded computational resources; therefore, the number of hidden units in SimpleRNN 1–3 layers was restricted to 500 for the 5120-vertex source space RNN, while layer 4 had 5120 units. The Adam optimizer learning rate hyperparameter was decreased to achieve stable convergence with larger models; learning rate was 0.001 (default) for 132 and 516 vertices, and 0.0001 for 1282, 2052, and 5120 vertices. Aside from these modifications, the training procedure was as described in Section 2.4. For comparison, eLORETA was also applied to forward models computed with denser source spaces, using the same parameters described for the surface source space in Section 2.6.

Coronal MRI slices at $y = -0.009$ m, including auditory regions, are shown in Fig. 6. Vertices located within ± 0.0003 m are displayed on these slices, coloured according to their instantaneous amplitude at 0.1 s when estimated using MEEG signals. Estimated source waveforms for these vertices are plotted alongside the MRI slices. Comparable figures for RNN and eLORETA methods applied EEG-only and MEG-only with different source space densities are available from https://osf.io/dfr52. The numbers of completed training iterations before RNNs stopped training were compared using the Kruskal-Wallis H-test (Kruskal & Wallis, 1952).

2.10. Software

Python 3 with matplotlib 3.5.2 (Hunter, 2007), meegkit 0.1.3, mne 1.4.0 (Gramfort et al., 2013), numpy 1.25.2 (Harris et al., 2020), pyvista 0.37.0, scipy 1.11.2 (Virtanen et al., 2020), and tensorflow 2.14.0

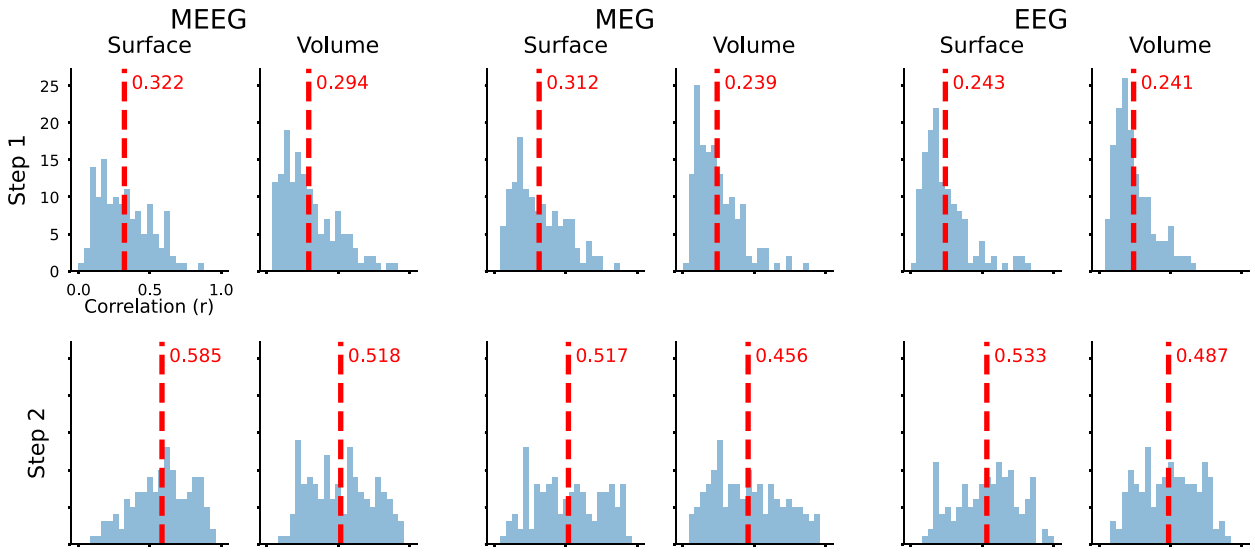


Fig. 5. Correlation among source estimates from 25 RNNs (five scales, five seeds) fitted to MEEG, MEG, or EEG data, with surface or volume source spaces. Directions (x, y, z) for volume space sources were treated separately, yielding 150 estimated source signals; the surface source space with fixed orientation had 132 estimated source signals. Step 2 source estimates were more consistent (higher r) than step 1. Surface space solutions are more consistent than those of the volume space, although this could be explained by fewer estimates in the surface space constraining the range of possible solutions. Moreover, MEEG sources were more consistent than single-modality estimates. Means are indicated by vertical dashed lines and reported in red text.

Table 1
Evaluation of reconstructed ERF and ERP waveforms and source signals estimated using different methods, source spaces, and fitting data.

Method	Source space	Fitted to	ERF			ERP			SRC[a]
			r	MSE	SNR	r	MSE	SNR	SNR
RNN	Volume	MEEG	0.929	4.92e-29	10.6	0.935	1.12e-13	15.8	13
		MEG	0.932	4.67e-29	10.2	0.0908	8.15e-13	11.3	10.9
		EEG	−0.00759	3.81e-28	16.8	0.949	8.36e-14	16.4	15.8
	Surface	MEEG	0.921	5.5e-29	10.7	0.937	1.08e-13	15.4	13.5
		MEG	0.927	5.06e-29	10.2	0.301	7.52e-13	12	12.2
		EEG	0.254	3.25e-28	17.1	0.952	7.81e-14	16.6	16
eLORETA	Volume	MEEG	0.838	1.04e-28	6.66	0.851	2.36e-13	6.89	5.23
		MEG	0.862	8.9e-29	6.48	0.0476	3.05e-12	5.24	5.07
		EEG	−0.0997	4.37e-28	5.2	0.889	2e-13	7.08	5.08
	Surface	MEEG	0.83	1.09e-28	6.83	0.805	3e-13	7.41	5.82
		MEG	0.882	7.73e-29	6.77	−0.0385	3.42e-12	3.97	4.76
		EEG	0.262	3.28e-28	4.82	0.836	2.92e-13	7.6	5.47
MxNE	Volume	MEEG	0.834	1.12e-28	6.41	0.833	2.53e-13	7.16	6.76
		MEG	0.857	1.01e-28	6.31	0.0177	1.38e-12	5.52	6.83
		EEG	0.00767	7.63e-28	5.9	0.916	1.32e-13	6.95	5.18
	Surface	MEEG	0.757	1.48e-28	6.71	0.754	3.54e-13	8.54	6.86
		MEG	0.843	1.07e-28	6.64	0.126	1.63e-12	6.1	6.44
		EEG	0.199	3.58e-28	3.81	0.755	3.74e-13	8.11	6.29

[a] SRC: source.
Gray shading is applied to highlight performance on the ERF or ERP waveform from the signal modality that was not used when deriving source estimates.

(Abadi et al., 2016) were used to conduct this research. Data and code are available from https://osf.io/dfr52.

3. Results

3.1. Scaling labels influences the correlation between reconstructed and reference ERF and ERP waveforms

Scaling MEG and EEG data labels and their corresponding lead field matrices was necessary to avoid vanishing gradients when training RNNs (Bengio et al., 1994; Pascanu et al., 2013). When label values are very small ($y \ll 1$), error gradients propagated backwards through the layers to update network parameters become minuscule (i.e., they "vanish"). Therefore, labels must be scaled appropriately for effective training. This requirement for modelling event-related neural signals with RNNs has been highlighted previously (O'Reilly, Wehrman et al., 2022). In the present study, additional consideration was given to balancing the amplitudes of labels derived from MEG and EEG because they have different raw signal magnitudes.

Different scaling coefficients for MEG and EEG data labels and lead-field matrices influenced the performance of RNNs, as detailed in Table S1. The average source estimates from five RNNs (i.e., 5 random seeds) trained using each set of scaling values were evaluated by correlation between reconstructed and reference ERF and ERP waveforms. These correlations are plotted in Figure S4. The scaled MEG to EEG standard deviation ($\sigma_{MEG}/\sigma_{EEG}$) ratio positively correlated with the ERF to ERP correlation (r_{ERF}/r_{ERP}) ratio from RNNs trained with MEEG. In contrast, the overall correlation of combined ERF and ERP waveforms (r_{both}) from RNNs trained with MEEG had an inverse relationship with $\sigma_{MEG}/\sigma_{EEG}$, indicating that the correlation with the ERP has more influence over r_{both}. Models trained with single modality labels are consistent with these observations. Outputs from RNNs trained with

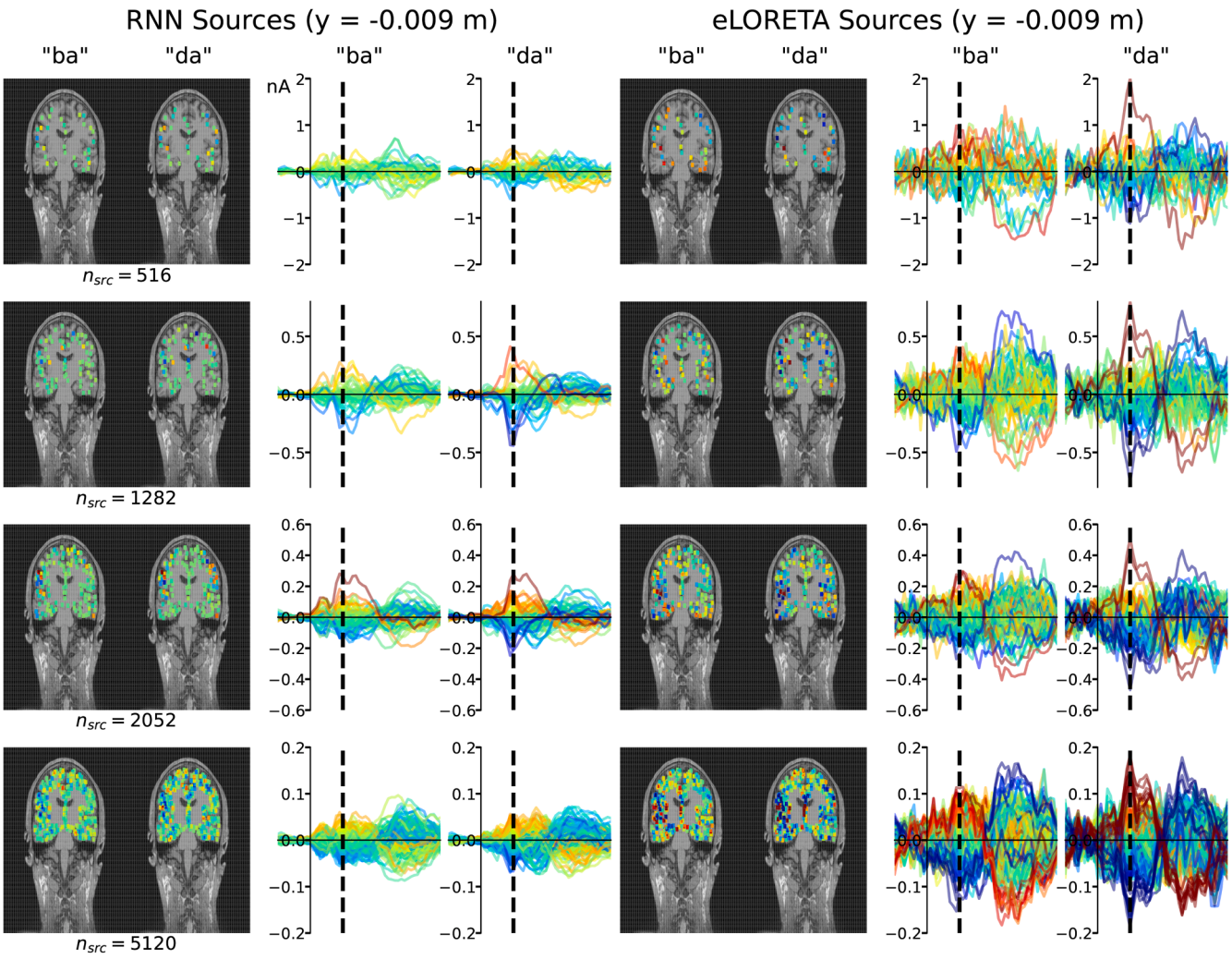


Fig. 6. Denser source spaces used with RNN and eLORETA source estimation methods for MEEG source reconstruction. Four rows of plots from top to bottom present data from surface source spaces with 516, 1282, 2052, and 5120 vertices, respectively. Panels on the left show results from the RNN method, panels on the right show results from eLORETA. Coronal MRI slices taken from $y = -0.009$ m are displayed with source space vertices coloured based on their instantaneous amplitude at 0.1 s post-stimulus. Estimated source waveforms for each of these vertices are plotted beside MRI slices for each method, coloured based on a common symmetrical colour scale (negative is dark-blue, positive is dark-red) for each source space density; ± 1 nA (516), ±0.5 nA (1282), ±0.25 nA (2052), and ±0.1 nA (5120). The time point represented by coronal vertex colouring is identified by a vertical dashed line in each source waveform plot; the time range corresponds to 0.1 s pre-stimulus to 0.4 s post-stimulus. Comparable figures for RNN and eLORETA used for EEG, MEG, and MEEG source reconstruction are available from https://osf.io/dfr52.

scaled MEG labels only, have a positive relationship between r_{ERF} and $\sigma_{MEG}/\sigma_{EEG}$. In contrast, outputs from RNNs trained with scaled EEG labels only, have a negative relationship between r_{ERP} and $\sigma_{MEG}/\sigma_{EEG}$. For RNNs trained with MEG-only or EEG-only, performance appeared to level off towards the favoured side of this ratio: higher for MEG, lower for EEG. The same pattern of results was observed for volume and surface source spaces.

These observations suggest that the overall correlation metric r_{both} is biased towards EEG data because EEG had fewer channels and higher inter-channel correlation (Figure S5) than MEG, making it "easier" for the model to fit and thereby achieve higher overall correlation. Biophysical processes cause EEG channels to have higher inter-channel correlation than MEG channels (Luck, 2014; Nunez & Srinivasan, 2009), so a model trained with larger amplitude labels derived from EEG than MEG appears to perform better at reproducing ERF and ERP waveforms overall. Despite the influence of scaling to sway the performance of RNNs to favour of fitting ERF or ERP waveforms, models trained with lead field weights and labels scaled by the five scaling coefficient sets generated highly correlated outputs ($r > 0.89$) with reference ERF and ERP waveforms. The ill-posed nature of EEG/MEG source reconstruction can make selecting one solution problematic. It is uncertain which solution is closest to the unknown pattern of ground-truth neural current sources that produced the reference ERP/ERF waveforms. Therefore, source estimates from all 25 RNN models (five scales × five seeds) were averaged to obtain the final source estimates used to evaluate reconstructed ERF and ERP waveforms against reference waveforms for the RNN method reported in Section 3.2.

3.2. *Fewer training epochs are required when fine-tuning models with L1 regularization*

Distributions of completed epochs for RNN training steps 1 and 2 are plotted in Figure S3 and corresponding descriptive statistics are provided in Table S3. In every case, training stopped before 3000 epochs were completed. There were no statistically significant differences between volume or surface source spaces, nor were there among models trained with MEEG, MEG-only, or EEG-only. In training step 2, when fine-tuning the RNN with L1 regularization applied to the source estimation layer (SimpleRNN 4 in Fig. 1), fewer epochs were completed than in training step 1 [$W = 1515.5$, $p = 7.21e\text{-}15$]. Observing a

statistically significant difference between the number of completed training epochs in step 1 and step 2 is unsurprising, given that deep neural network models generally require less training to converge when pre-trained before transfer learning (Hosna et al., 2022). Adding the L1-norm constraint to source estimation layer activations also produced more consistent source estimates for each vertex from RNNs (described in Section 3.6, Fig. 5), and caused retrograde attenuation of hidden unit activations in preceding layers (e.g., as shown in Figure S8).

The RNN was trained using the BPTT algorithm, described in Section 2.4. The network, depicted in Fig. 1, was configured as a sequence-in/sequence-out model. During training, its inputs were step-pulse signals, representing psychophysiological events with high (amplitude 1) during stimulus-on times, and low (amplitude 0) during stimulus-off times. Its labels were scaled event-related neural signals, as described in Section 2.5. Training was performed in a typical supervised learning fashion, with network parameters (weights and biases) being iteratively updated to minimise MSE loss between model outputs and labels. Output layer weights were fixed as the scaled lead-field matrix, so these weights did not change during training. Internal dynamics of the RNN shown in Figure S8 evolve from this optimization scenario, as the model transforms input step-pulse signals into estimated source waveforms at the penultimate layer (SimpleRNN 4), which are multiplied by the scaled lead-field matrix to produce model outputs.

3.3. Reconstructed ERF and ERP waveforms from RNN are "better" than eLORETA and MxNE

Reference ERF and ERP waveforms are plotted on the top row of Fig. 2 above waveforms reconstructed from the RNN method. Source estimates from RNNs fitted to MEEG data reconstructed ERF and ERP waveforms with high fidelity. In contrast, source signals estimated from RNNs trained with MEG-only or EEG-only produced comparatively poor reconstructions of contra waveforms; i.e., those from RNNs trained with EEG-only had poor ERF reconstructions, whereas those from RNNs trained with MEG-only had poor ERP reconstructions. Comparable plots for eLORETA and MxNE are in Figure S6 and Figure S7, respectively, where a similar pattern is observed. Analyses of reconstructed ERF and ERP waveforms are reported in Table 1. An example of signals generated by each hidden unit of the RNN is plotted in Figure S8. These signals generated internally by the model are analogous to neurophysiological signals generated by the brain that produce ERF and ERP waveforms. This functional distinction between the RNN method and conventional linear inverse solutions, such as eLORETA and MxNE, offers the possibility of using the RNN as a "model organism" for investigating computational mechanisms associated with event-related neural signal generation (Lindsay & Bau, 2023; O'Reilly, 2022a, 2022).

Metrics calculated for reconstructed ERF and ERP waveforms are provided in Table 1 and descriptive statistics are provided in Table S4. There were no statistically significant differences between volume and surface source spaces in terms of r [$W = 82.0$, $p = 0.899$], MSE [$W = 62.0$, $p = 0.325$], or SNR [$W = 53.0$, $p = 0.167$]. Metrics for waveforms reconstructed from source signals estimated by the RNN method were statistically significantly better, meaning higher r and SNR and lower MSE, from those of eLORETA [r: $W = 1.0$, $p = 0.000977$, MSE: $W = 0.0$, $p = 0.000488$, SNR: $W = 0.0$, $p = 0.000488$] and MxNE [r: $W = 1.0$, $p = 0.000977$, MSE: $W = 0.0$, $p = 0.000488$, SNR: $W = 0.0$, $p = 0.000488$], whereas those from eLORETA and MxNE were not statistically significantly different [r: $W = 27.0$, $p = 0.380$, MSE: $W = 32.0$, $p = 0.622$, SNR: $W = 25.0$, $p = 0.301$].

Descriptive statistics for estimated source signal to noise ratio (SRC SNR) are provided in Table S5. The SRC SNR was not statistically significantly different between volume and surface source spaces [$W = 7.0$, $p = 0.0742$]. However, they were statistically significantly different between RNN and eLORETA [$W = 0.0$, $p = 0.03125$], RNN and MxNE [$W = 0.0$, $p = 0.03125$], and eLORETA and MxNE [$W = 0.0$, $p = 0.03125$]. RNN and MxNE methods incorporate an L1 norm constraint, producing higher SRC SNR than eLORETA. However, MxNE does this by turning off sources, whereas the RNN does not eliminate any sources. Training RNNs with input signals that go high only during event-on time points (bottom of Fig. 1) induces common baseline activity for both event types because there is no distinction between input representations before stimulus onset. Therefore, baseline activity estimated by the RNN converges towards the mean of both event types and is less than that estimated by eLORETA or MxNE methods, producing higher SNR for estimated source signals and reconstructed waveforms. This can be viewed as an advantage of the RNN method with the assumption that there are no systematic differences in preparatory activity or overlap from previous events between event types that would cause their baseline activity to be different.

These results demonstrate that the RNN method can produce source signals that effectively reconstruct event-related neural signals observed from simultaneously recorded MEG and EEG. These reconstructed ERF and ERP waveforms had higher correlation and lower MSE evaluated against reference ERF and ERP waveforms and higher SNR compared with those derived from source estimates produced by eLORETA or MxNE methods (Table 1). All methods applied to MEEG data produced source signals better at reconstructing both ERF and ERP waveforms compared with single modality data alone. As seen from Fig. 2 (and Figure S6 and Figure S7), source estimates derived from MEG-only or EEG-only were comparatively poor at reconstructing event-related signals from the modality not used to derive those source estimates. This highlights that MEG and EEG sensitivity to different sources varies according to the location and orientation of those sources relative to the electrode or magnetometer (Nunez & Srinivasan, 2009). Electric potential gets smeared through the skull and the magnetic field does not. Moreover, MEG and EEG are sensitive to perpendicular components of the electromagnetic field and have different sensor geometry relative to the brain. As such, it is not unexpected that source signals estimated using either modality alone are considerably different (e.g., see Figs. 3 and 4).

3.4. Source rankings from surface source space provide stronger evidence for vertex relevance

From three source estimation techniques (RNN, eLORETA, and MxNE) applied to three data types (MEEG, MEG, and EEG), the top five highest-ranked sources based on the correlation between their projections and reference ERF and ERP waveforms are reported in Table 2. From volume source space, vertices 21 and 23 were most consistently ranked in the top five across estimation methods and data types, at 6/9 and 5/9, respectively. The binomial distribution law of selecting k from n elements gives the probability of each source ranking among the top five across nine experiments. The probability of any single volume-space vertex being ranked in the top five by random chance would be $5/50 = 0.1$, and the probabilities of vertices 21 and 23 appearing in 6/9 and 5/9 of the top-five ranked source estimates are on the order of 10^{-5} and 10^{-4}. From the surface source space, vertices 40 and 89 were ranked among the top five of 8/9 source estimates. By random chance, the probability of an individual vertex in the surface source space being ranked in the top five would be $5/132 = 0.0378$; therefore, the probability of observing these results for vertices 40 and 89 is on the order of 10^{-10}.

The lower probability of observed rankings for vertices 40 and 89 of the surface source space can be viewed as evidence that these estimates are more reliable than those of the volume source space. As such, estimated source signals from surface space vertices 40 and 89 are plotted in Figs. 3 and 4, respectively; equivalent figures for all volume and surface source space vertices are available from https://osf.io/dfr52. Vertices 40 and 89 are located in the left and right temporal lobes, respectively, where one might expect syllable sounds to excite neural tissue specialised for sound and language processing (Rong et al., 2018; Schirmer et al., 2012). Source estimates from the RNN are plotted with shaded

Table 2
Source signals ranked by correlation between projected and observed ERF and ERP waveforms.

Source space[a]	Method	Fitted to	Ranked source estimates[b] (vertex number: r_{both})[c]				
			#1	#2	#3	#4	#5
Volume	RNN	MEEG	22:	35:	23:	36:	17:
			0.737	0.706	0.703	0.66	0.657
		MEG	16:	23:	24:	17:	37:
			0.301	0.269	0.266	0.254	0.222
		EEG	32:	35:	21:	22:	36:
			0.805	0.801	0.784	0.774	0.772
	eLORETA	MEEG	22:	35:	21:	18:	23:
			0.689	0.543	0.45	0.425	0.389
		MEG	34:	7:	17:	21:	38:
			0.294	0.289	0.286	0.234	0.205
		EEG	22:	35:	18:	11:	23:
			0.736	0.677	0.605	0.598	0.587
	MxNE	MEEG	21:	23:	37:	24:	36:
			0.472	0.466	0.315	0.304	0.279
		MEG	21:	7:	17:	34:	24:
			0.362	0.326	0.28	0.265	0.214
		EEG	10:	21:	27:	48:	12:
			0.559	0.442	0.431	0.306	0.276
Surface	RNN	MEEG	89:	63:	115:	2:	40:
			0.526	0.514	0.453	0.446	0.444
		MEG	40:	89:	37:	1:	116:
			0.384	0.345	0.323	0.293	0.284
		EEG	40:	28:	89:	76:	37:
			0.623	0.588	0.583	0.559	0.531
	eLORETA	MEEG	89:	40:	108:	115:	91:
			0.419	0.358	0.339	0.322	0.278
		MEG	89:	40:	116:	22:	35:
			0.37	0.259	0.207	0.195	0.183
		EEG	40:	89:	31:	115:	108:
			0.485	0.482	0.403	0.383	0.382
	MxNE	MEEG	89:	115:	116:	2:	40:
			0.427	0.375	0.324	0.303	0.262
		MEG	89:	116:	35:	37:	40:
			0.405	0.254	0.218	0.206	0.187
		EEG	115:	63:	100:	36:	2:
			0.431	0.308	0.296	0.282	0.28

[a] Volume space projections were computed from three signal components (*x*, *y*, and *z* directions) for each of 50 vertices, whereas surface space projections were computed from individual signal components for 132 vertices.

[b] Ranked from highest to lowest correlation between projected and observed MEEG.

[c] r_{both} was calculated by comparing projections with combined ERF and ERP waveforms.

error bars that quantify variability among 25 RNN-estimated source waveforms. The error bars are more naturally estimated with RNN as the processing of temporal information is integrated with the model, while eLORETA and MxNE are solved for isolated time points, thus ignoring time dependency. The amplitude range for RNN source signals is smaller than that of eLORETA and MxNE, and RNN source signals have less pre-stimulus baseline activity. The MxNE method turns off vertices 40 and 89 for both "ba" and "da" responses when fitted to EEG and vertex 40 for "da" responses when fitted to MEEG and MEG as well. Signals from the RNN are qualitatively smoother than those produced by the other methods, reflecting dependence on the sequential structure of internal dynamics (Figure S8), in contrast with eLORETA and MxNE that can be applied to individual time samples. Aside from these differences, the estimated source signals from all three methods at these two vertices are qualitatively similar, indicating that they resemble the likely distribution of the current source density at these locations.

3.5. Analysis of individual source projections

Distributions of r, MSE, and SNR measured from waveforms projected from individual estimated sources are plotted in Figure S9, Figure S10, and Figure S11, respectively, with descriptive statistics provided in Table S6. Three signal components were projected simultaneously for the volume source space. In contrast, one signal component was projected for each source vertex for the surface source space, reflecting free and fixed dipole orientations, respectively. For r and MSE, projections were compared with combined reference ERF and ERP waveforms, and for calculating SNR, the projected ERF and ERP waveforms were also used.

Correlation between projections and reference ERF and ERP waveforms was higher for volume sources than surface sources [$U = 255{,}965.0$, $p = 1.788\text{e-}08$]. This is observed because volume sources had three projected signal components (for x, y, and z directions), whereas surface sources had only one. Borrowing terminology from statistical optimization theory, volume space vertices had more degrees of freedom than surface space vertices, and the number of degrees of freedom has a positive relationship with goodness of fit. Source estimates from the RNN produced projections that had a higher correlation with reference waveforms than those from eLORETA [$W = 3030.0$, $p = 4.857\text{e-}84$] or MxNE [$U = 133{,}944.0$, $p = 2.988\text{e-}22$]; however, those from eLORETA and MxNE were not statistically significantly different [$U = 93{,}815.0$, $p = 0.417$]. Projections of source estimates derived from EEG had higher correlation than those derived from MEEG [$U = 84{,}644.0$, $p = 2.127\text{e-}13$] or MEG [$U = 37{,}265.0$, $p = 4.936\text{e-}70$]. The same pattern of statistically significant differences was observed for MSE, but with inverted relationships; e.g., volume space projections had higher r and lower MSE than surface space projections.

The finding that source estimates obtained by fitting EEG-only produced projections with higher overall correlation can be explained by higher inter-channel correlation of EEG (mean absolute $r = 0.531$) compared with MEG (mean absolute $r = 0.315$) [$U = 6{,}299{,}769.0$, $p = 1.271\text{e-}221$]; descriptive statistics in Table S7. Distributions of inter-channel correlation coefficients from MEG and EEG are plotted in Figure S5. EEG's higher absolute inter-channel correlations reflects a greater field spread than MEG as it passes through the skull (Luck, 2014; Nunez & Srinivasan, 2009). This tendency for EEG channels to have higher correlation, and EEG amplitudes being orders of magnitude greater than MEG, causes r and MSE metrics applied to combined ERF and ERP waveforms to be biased towards the latter.

In terms of projection SNRs, volume sources were lower than surface sources [$U = 199{,}267.0$, $p = 0.0255$]. Projections from source signals estimated using the RNN method had higher SNR than those of eLORETA [$W = 96.0$, $p = 6.617\text{e-}91$] or MxNE [$U = 187{,}110.0$, $p = 1.912\text{e-}123$], whereas those from MxNE had higher SNR than those of eLORETA [$U = 87{,}648.0$, $p = 0.0152$]. The SNRs of projections from source estimates derived from MEEG were not statistically significantly different from those of MEG [$U = 129{,}501.0$, $p = 0.0556$] or EEG [$U = 108{,}850.0$, $p = 0.0796$]. However, those from EEG were statistically significantly higher than MEG [$U = 96{,}162.0$, $p = 0.000287$].

3.6. Training RNNs with MEEG labels and L1 regularization produced more consistent source estimates

Consistency of source estimates refers to correlation between source waveforms estimated for the same vertex across multiple RNNs. Distributions of mean source signal correlation among estimates from 25 RNN models are plotted in Fig. 5, with descriptive statistics provided in Table S8. The upper triangle of each correlation coefficient matrix from 25 estimates ($n = 300$) was averaged to find the mean correlation for each vertex in volume and surface source spaces. Models after training step 2 had higher correlations than after training step 1 [$W = 6928.0$, $p = 1.296\text{e-}129$], demonstrating that the L1 constraint for fine-tuning produces more consistent source signal estimates. The surface source space had higher mean correlations than the volume source space [$U = 397{,}015.0$, $p = 5.121\text{e-}05$]; potentially because the domain of all possible solutions constrained by the surface source space with fixed orientation is smaller than the domain of solutions constrained by the volume source space (Henson, Mattout et al., 2009).

Source estimates from RNNs trained with MEEG had higher mean correlations than those from RNNs trained with MEG-only [$W = 50{,}824.0$, $p = 9.396\text{e-}14$] or EEG-only [$W = 57{,}817.0$, $p = 1.672\text{e-}08$]. However, the difference between mean source signal correlations across 25 RNNs trained with MEG-only or EEG-only was not statistically significantly different [$W = 76{,}958.0$, $p = 0.484$]. These observations could be interpreted in light of the additional constraints placed on the solution space by combined MEEG, which produces more consistent source estimates than either modality individually by narrowing the search space of possible solutions (Chowdhury et al., 2015; Chowdhury et al., 2018; Dale et al., 1993). This type of analysis cannot be applied to eLORETA or MxNE because they are deterministic.

3.7. Denser source spaces

It was necessary to decrease Adam optimizer learning rate from 0.001 to 0.0001 to fit RNNs with source spaces containing a larger number of (i.e., >516) vertices. Tuning the learning rate hyper-parameter enabled the convergence of models with more hidden units and parameters (weights and biases) to update during training. Also, for the highest density source space tested (i.e. 5120 vertices), the number of hidden units in layers 1–3 was reduced from $2n_{src}$ to 500. This change did not prevent the RNN from converging, suggesting that the efficiency of models trained more sources can be improved by decreasing the number of units in earlier hidden layers. For example, models trained with 2052 units in layer 4 could have been trained with fewer than 4104 hidden units in layers 1–3. However, RNNs applied to different source densities did have statistically significant differences in the number of completed training iterations [$K = 63.385$, $p = 1.818\text{e-}13$], principally reflecting RNNs trained with the 5120-vertex source space requiring more training iterations (Table S9). This observation could be due to having only 500 units in layers 1–3 combined with a higher number of units in layer 4 of the 5120-source RNNs. This pattern of results suggests considerable scope for systematically evaluating model architectures to determine the optimal trade-off between performance and efficiency.

Reconstructed ERP and ERF waveforms from different source space densities are plotted in Figure S12 for RNN and Figure S13 for eLORETA methods. Consistent with the results from sparser source spaces in Table 1, sources estimated using MEEG reconstructed both ERF and ERP waveforms well. In contrast, sources estimated with EEG only made relatively poor reconstructions of ERF waveforms, and vice versa. Coronal MRI slices showing auditory cortical regions and estimated source amplitudes at 0.1 s are shown in Fig. 6 alongside corresponding estimated source waveforms for the vertices shown. Both RNN and eLORETA methods estimate higher amplitudes in vertices concentrated towards the temporal regions, consistent with cortical sites involved in the auditory processing of syllable stimuli (Rong et al., 2018; Schirmer et al., 2012). Source space density is inversely related to estimated source amplitudes, and eLORETA estimates tend to have higher amplitudes than RNN estimates. However, the RNN approach yields smoother, higher SNR estimates than eLORETA, consistent with results in Figs. 3 and 4.

Minimum-norm methods of EEG/MEG source reconstruction are generally considered to be biased towards superficial sources, and therefore often incorporate a depth weighting regularization parameter (Allouch et al., 2023; Fuchs et al., 1999; Lin et al., 2006). The eLORETA method implemented in MNE-Python applies depth weighting as described by Pascual-Marqui et al. (2011). An equivalent form of regularization was not included in the RNN method, which only applied L1-norm regularization of layer 4 outputs to derive sparse source estimates. However, this difference appears inconsequential for the current case because source estimates from the RNN do not exhibit undue bias towards superficial sources relative to eLORETA. On the contrary, eLORETA arguably estimates unrealistic source activations in the corpus callosum region (e.g., Fig. 6), which could be partly due to depth weighting regularization.

4. Discussion

The RNN method of MEEG source reconstruction has been demonstrated using data from a single subject; in future work, it could be expanded for application to multiple subjects simultaneously. Forward conduction models were computed from a template head BEM model that was morphed using accurate head digitization points. An anatomically accurate head model created from an MRI scan of the subject from whom MEG and EEG data were collected would have been preferable (Henson, Mattout et al., 2009). Using a template head model does not inherently invalidate the results from inverse source estimations but warrants an air of caution regarding their over-interpretation (Akalin Acar & Makeig, 2013; Das et al., 2020; Liu et al., 2023). As with all inverse methods applied to real MEG and EEG data, there are no ground-truth measurements to validate current source density estimates; nevertheless, results from the different methods can be aggregated to form a picture of the likely underlying distribution of neural activity that cannot be measured directly.

In a previous study (O'Reilly, Zhu et al., 2023), we simulated MEG data that provided ground-truth source signals for evaluating reconstruction methods. In that analysis, the RNN method performed comparably with minimum-norm estimates (Hämäläinen & Ilmoniemi, 1994) and MxNE (Gramfort et al., 2012), and outperformed beamformers (Van Veen et al., 1997). The main purpose of the current work was to extend the RNN method to fit MEEG signals, therefore the simulation component was omitted from this study. In future work, simulation methods such as those advocated by Allouch et al. (2023) may be used to systematically determine optimal RNN architecture (e.g., the number of hidden layers and units) and hyper-parameters (e.g., activation functions and autoregressive order). Specifically incorporating larger (i.e., >1) and different autoregressive orders can represent different transmission delays to mimic brain function (Sanchez-Bornot et al., 2024). Furthermore, several other established methods of inverse MEG/EEG source reconstruction have not been compared here, e.g., dipole fitting or dynamic statistical parametric mapping (Dale et al., 2000), partly for the sake of brevity, but more importantly because the results of such comparisons would not affect the main conclusions regarding the RNN's ability to estimate neural sources of ERP/ERF waveforms from MEEG.

The computational requirements of training RNNs were orders of magnitude greater than that of applying eLORETA or MxNE methods. It took approximately nine hours of computer processing time to train 300 RNN models for this study, excluding denser source space experiments; i.e., 2 source spaces (volume, surface) × 3 data label types (MEEG, MEG, EEG) × 5 scaling values × 5 random seeds × 2 training steps. In contrast, eLORETA and MxNE took less than a few minutes. Whilst more resources are required to implement the RNN method than other inverse solutions, most modern computing workstations can train an RNN in reasonable timescales for post-hoc analysis of neural signals. Moreover, the advantages of using an RNN for source estimation may outweigh its higher computational costs; specifically, it provides a better reproduction of ERF and ERP waveforms, higher SNR, and a model for investigating event-related signal generation processes.

At most, the RNN is an abstract model of neural signal processing, metaphorically representing biological neural networks that generate ERF and ERP waveforms. It lacks sufficient biological realism to make inference about the underlying neurophysiology. Details of the training algorithm, computational units, network connectivity, and objective of the RNN do not reflect biological reality. Brains develop ("are trained") through evolution and experience, consisting of genetically and neurochemically heterogeneous processing units, comprise comparatively massive interconnected neural circuits with imprecisely characterised network topologies, and implement the objective of controlling perception and behaviour. In contrast, the RNN has a simple six-layer structure (input, four hidden, and an output) with relatively homogenous computational units (the only difference being rectified vs. linear

activation function), optimized by gradient descent to reproduce labels derived from MEG and EEG signals.

5. Conclusions

The RNN method is effective at estimating event-related neural source signals from MEEG, outperforming the linear inverse solutions in every way except for computational complexity. It transforms simple time sequences representing physical/cognitive events into complex internal dynamics that effectively replicate ERF and ERP waveforms. Its dependence on sequential relationships among internal signals paints a sharp distinction with linear inverse MEEG source estimation methods applied to individual time points. This reliance on sequential dynamics might explain why estimated source signals from the RNN are smoother and have less baseline noise than those from the other methods.

While the RNN hidden unit activations are not precisely matched to biological neural signals, they demonstrate how a relatively simple representation of psychophysiological events can undergo a hierarchy of transformations to elicit ERF and ERP waveforms. This analogy to human brain function is not available from conventional inverse solutions. It opens up the possibility of using the RNN to study computations underpinning event-related neural signal generation. In future studies, biophysically-inspired constraints may be incorporated to explore how structural and functional changes to the recurrent network influence its performance at source estimation. These constraints may include more diverse patterns of connectivity among units and layers of the RNN (Perich & Rajan, 2020), controlling the relative balance of excitation and inhibition in the network (Hertäg & Clopath, 2022), incorporating different autoregressive orders (Sanchez-Bornot et al., 2024), and employing more sophisticated artificial neuron models.

This study also showed empirically that surface source space with fixed surface-normal orientation is preferable to volume source space with free orientation because surface source space produced more consistent estimates and source rankings. This finding is consistent with biophysical arguments of the likely sources of most ERF and ERP signal components being postsynaptic potentials at the distal or apical dendrites of cortical pyramidal neurons oriented radially within the cortical sheet (Luck, 2014; Nunez & Srinivasan, 2009); with some notable exceptions such as auditory brainstem potentials (Moore, 1987). Moreover, the results indicate that MEEG is superior to MEG or EEG, yielding more consistent source signal estimates by constraining the space of solutions that can fit the observed data.

Data availability

Data and code from this study are available from https://osf.io/dfr52.

Ethical review

The Human Research Ethics Committee at Macquarie University reviewed and approved the experimental protocol.

CRediT authorship contribution statement

Jamie A. O'Reilly: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Judy D. Zhu:** Writing – review & editing, Software, Methodology, Data curation, Conceptualization. **Paul F. Sowman:** Writing – review & editing, Supervision, Resources, Funding acquisition, Conceptualization.

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.

Acknowledgements

The authors acknowledge the facilities and technical assistance at the KIT-Macquarie Brain Research (MEG) Laboratory, part of the Australian National Imaging Facility. This work was financially supported by King Mongkut's Institute of Technology Ladkrabang [KREF046704].

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neunet.2024.106731.

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Zheng, X. (2016). *TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems*. http://arxiv.org/abs/1603.04467.

Ahlfors, S. P., & Hämäläinen, M. S. (2012). MEG and EEG: Source estimation. In R. Brette, & A. Destexhe (Eds.), *Handbook of neural activity measurement* (pp. 257–286). Cambridge University Press. https://doi.org/10.1017/CBO9780511979958.007.

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

Allouch, S., Kabbara, A., Duprez, J., Khalil, M., Modolo, J., & Hassan, M. (2023). Effect of channel density, inverse solutions and connectivity measures on EEG resting-state networks reconstruction: A simulation study. *NeuroImage, 271*, Article 120006. https://doi.org/10.1016/j.neuroimage.2023.120006

Bengio, Y., Simard, P., & Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks, 5*(2), 157–166. https://doi.org/10.1109/72.279181

Brown, M. B., & Forsythe, A. B. (1974). Robust tests for the equality of variances. *Journal of the American Statistical Association, 69*(346), 364. https://doi.org/10.2307/2285659

Chowdhury, R. A., Pellegrino, G., Aydin, Ü., Lina, J. M., Dubeau, F., Kobayashi, E., & Grova, C. (2018). Reproducibility of EEG-MEG fusion source analysis of interictal spikes: Relevance in presurgical evaluation of epilepsy. *Human Brain Mapping, 39*(2), 880–901. https://doi.org/10.1002/HBM.23889

Chowdhury, R. A., Zerouali, Y., Hedrich, T., Heers, M., Kobayashi, E., Lina, J. M., & Grova, C. (2015). MEG–EEG information fusion and electromagnetic source imaging: From theory to clinical application in epilepsy. *Brain Topography, 28*(6), 785–812. https://doi.org/10.1007/S10548-015-0437-3/FIGURES/8

D'Agostino, R., & Pearson, E. S. (1973). Tests for departure from normality. Empirical results for the distributions of b2 and $\sqrt{b1}$. *Biometrika, 60*(3), 613–622. https://doi.org/10.2307/2335012

Dale, A. M., Liu, A. K., Fischl, B. R., Buckner, R. L., Belliveau, J. W., Lewine, J. D., & Halgren, E. (2000). Dynamic statistical parametric mapping: Combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron, 26*(1), 55–67. https://doi.org/10.1016/S0896-6273(00)81138-1

Dale, A. M., Sereno, M. I., & Prechtl, J. C. (1993). Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: A linear approach. *Journal of Cognitive Neuroscience, 5*(2), 162–176. https://doi.org/10.1162/JOCN.1993.5.2.162

Das, P., Brodbeck, C., Simon, J. Z., & Babadi, B. (2020). Neuro-current response functions: A unified approach to MEG source analysis under the continuous stimuli paradigm. *NeuroImage, 211*, Article 116528. https://doi.org/10.1016/J.NEUROIMAGE.2020.116528

David, O., Kiebel, S. J., Harrison, L. M., Mattout, J., Kilner, J. M., & Friston, K. J. (2006). Dynamic causal modeling of evoked responses in EEG and MEG. *NeuroImage, 30*(4), 1255–1272. https://doi.org/10.1016/J.NEUROIMAGE.2005.10.045

de Cheveigné, A., & Simon, J. Z. (2007). Denoising based on time-shift PCA. *Journal of Neuroscience Methods, 165*(2), 297–305. https://doi.org/10.1016/J.JNEUMETH.2007.06.003

Dinh, C., Samuelsson, J. G., Hunold, A., Hämäläinen, M. S., & Khan, S. (2021). Contextual MEG and EEG source estimates using spatiotemporal LSTM networks. *Frontiers in Neuroscience, 15*, 119. https://doi.org/10.3389/FNINS.2021.552666/BIBTEX

Fuchs, M., Wagner, M., Köhler, T., & Wischmann, H. A. (1999). Linear and nonlinear current density reconstructions. *Journal of Clinical Neurophysiology, 16*(3), 267–295. https://doi.org/10.1097/00004691-199905000-00006

Gramfort, A., Kowalski, M., & Hämäläinen, M. (2012). Mixed-norm estimates for the M/EEG inverse problem using accelerated gradient methods. *Physics in Medicine & Biology, 57*(7), 1937. https://doi.org/10.1088/0031-9155/57/7/1937

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., & Hämäläinen, M. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience, 7*(7 DEC), 267. https://doi.org/10.3389/fnins.2013.00267

Hämäläinen, M. S., & Ilmoniemi, R. J. (1994). Interpreting magnetic fields of the brain: Minimum norm estimates. *Medical & Biological Engineering & Computing, 32*(1), 35–42. https://doi.org/10.1007/BF02512476/METRICS

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., … Oliphant, T. E. (2020). Array programming with NumPy. *Nature, 585*(7825), 357–362. https://doi.org/10.1038/s41586-020-2649-2

Hecker, L., Rupprecht, R., Tebartz Van Elst, L., & Kornmeier, J. (2021). ConvDip: A convolutional neural network for better EEG source imaging. *Frontiers in Neuroscience, 15*, 533. https://doi.org/10.3389/FNINS.2021.569918/BIBTEX

Henson, R. N., Flandin, G., Friston, K. J., & Mattout, J. (2010). A Parametric empirical bayesian framework for fMRI-constrained MEG/EEG source reconstruction. *Human Brain Mapping, 31*(10), 1512–1531. https://doi.org/10.1002/HBM.20956

Henson, R. N., Mattout, J., Phillips, C., & Friston, K. J. (2009). Selecting forward models for MEG source-reconstruction using model-evidence. *NeuroImage, 46*(1), 168–176. https://doi.org/10.1016/j.neuroimage.2009.01.062

Henson, R. N., Mouchlianitis, E., & Friston, K. J. (2009). MEG and EEG data fusion: Simultaneous localisation of face-evoked responses. *NeuroImage, 47*(2), 581–589. https://doi.org/10.1016/J.NEUROIMAGE.2009.04.063

Hertäg, L., & Clopath, C. (2022). Prediction-error neurons in circuits with multiple neuron types: Formation, refinement, and functional implications. *Proceedings of the National Academy of Sciences, 119*(13). https://doi.org/10.1073/PNAS.2115699119

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation, 9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Hosna, A., Merry, E., Gyalmo, J., Alom, Z., Aung, Z., & Azim, M. A. (2022). Transfer learning: A friendly introduction. *Journal of Big Data, 9*(1). https://doi.org/10.1186/s40537-022-00652-w

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science and Engineering, 9*(3), 90–95. https://doi.org/10.1109/MCSE.2007.55

Jung, T., Makeig, S., Humphries, C., Lee, T., McKeown, M. J., Iragui, V., & Sejnowski, T. J. (2000). Removing electroencephalographic artifacts by blind source separation. *Psychophysiology, 37*(2), 163–178. https://doi.org/10.1111/1469-8986.3720163

Kiebel, S. J., David, O., & Friston, K. J. (2006). Dynamic causal modelling of evoked responses in EEG/MEG with lead field parameterization. *NeuroImage, 30*(4), 1273–1284. https://doi.org/10.1016/J.NEUROIMAGE.2005.12.055

Kingma, D. P., & Ba, J. L. (2015). Adam: A method for stochastic optimization. In *3rd International conference on learning representations, ICLR 2015 - conference track proceedings*.

Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association, 47*(260), 583–621. https://doi.org/10.1080/01621459.1952.10483441

Liang, J., Yu, Z. L., Gu, Z., & Li, Y. (2023). Electromagnetic source imaging with a combination of sparse Bayesian learning and deep neural network. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 1. https://doi.org/10.1109/tnsre.2023.3251420

Lin, F. H., Witzel, T., Ahlfors, S. P., Stufflebeam, S. M., Belliveau, J. W., & Hämäläinen, M. S. (2006). Assessing and improving the spatial accuracy in MEG source localization by depth-weighted minimum-norm estimates. *NeuroImage, 31*(1), 160–171. https://doi.org/10.1016/J.NEUROIMAGE.2005.11.054

Lindsay, G. W., & Bau, D. (2023). Testing methods of neural systems understanding. *Cognitive Systems Research, 82*, Article 101156. https://doi.org/10.1016/j.cogsys.2023.101156

Liu, C., Downey, R. J., Mu, Y., Richer, N., Hwang, J., Shah, V. A., Sato, S. D., Clark, D. J., Hass, C. J., Manini, T. M., Seidler, R. D., & Ferris, D. P. (2023). Comparison of EEG source localization using simplified and anatomically accurate head models in younger and older adults. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 31*, 2591–2602. https://doi.org/10.1109/TNSRE.2023.3281356

Luck, S. J. (2014). *An introduction to the event-related potential technique*. MIT press.

Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics, 18*(1), 50–60. https://doi.org/10.1214/aoms/1177730491

Moore, J. K. (1987). The human auditory brain stem as a generator of auditory evoked potentials. *Hearing Research, 29*(1), 33–43. https://doi.org/10.1016/0378-5955(87)90203-6

Nunez, P. L., & Srinivasan, R. (2009). Electric fields of the brain: The neurophysics of EEG. *Electric fields of the brain: The neurophysics of EEG*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195050387.001.0001

O'Reilly, J. A. (2022a). Modelling mouse auditory response dynamics along a continuum of consciousness using a deep recurrent neural network. *Journal of Neural Engineering, 19*. https://doi.org/10.1088/1741-2552/AC9257

O'Reilly, J. A. (2022b). Recurrent neural network model of human event-related potentials in response to intensity oddball stimulation. *Neuroscience, 504*, 63–74. https://doi.org/10.1016/J.NEUROSCIENCE.2022.10.004

O'Reilly, J. A., Angsuwatanakul, T., & Wehrman, J. (2022). Decoding violated sensory expectations from the auditory cortex of anaesthetised mice: Hierarchical recurrent neural network depicts separate 'danger' and 'safety' units. *European Journal of Neuroscience*. https://doi.org/10.1111/ejn.15736

O'Reilly, J. A., Wehrman, J., Carey, A., Bedwin, J., Hourn, T., Asadi, F., & Sowman, P. F. (2023). Neural correlates of face perception modeled with a convolutional recurrent neural network. *Journal of Neural Engineering, 20*(2). https://doi.org/10.1088/1741-2552/acc35b

O'Reilly, J. A., Wehrman, J., & Sowman, P. F. (2022). A guided tutorial on modelling human event-related potentials with recurrent neural networks. *Sensors, 22*(23), 9243. https://doi.org/10.3390/s22239243

O'Reilly, J. A., Zhu, J. D., & Sowman, P. F. (2023). Localized estimation of electromagnetic sources underlying event-related fields using recurrent neural networks. *Journal of Neural Engineering, 20*(4), Article 046035. https://doi.org/10.1088/1741-2552/acef94

Pantazis, D., & Adler, A. (2021). MEG source localization via deep learning. *Sensors, 21*(13), 4278. https://doi.org/10.3390/S21134278

Pascanu, R., Mikolov, T., & Bengio, Y. (2013). On the difficulty of training recurrent neural networks. In *30th International conference on machine learning, ICML 2013, PART 3* (pp. 2347–2355). https://proceedings.mlr.press/v28/pascanu13.html.

Pascual-Marqui, R. D., Faber, P., Kinoshita, T., Kochi, K., Milz, P., Nishida, K., & Yoshimura, M. (2018). Comparing EEG/MEG neuroimaging methods based on localization error, false positive activity, and false positive connectivity. *BioRxiv*, 1–18. https://doi.org/10.1101/269753

Pascual-Marqui, R. D., Lehmann, D., Koukkou, M., Kochi, K., Anderer, P., Saletu, B., Tanaka, H., Hirata, K., John, E. R., Prichep, L., Biscay-Lirio, R., & Kinoshita, T. (2011). Assessing interactions in the brain with exact low-resolution electromagnetic tomography. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 369*(1952), 3768–3784. https://doi.org/10.1098/rsta.2011.0081

Perich, M. G., & Rajan, K. (2020). Rethinking brain-wide interactions through multi-region 'network of networks' models. *Current Opinion in Neurobiology, 65*, 146–151. https://doi.org/10.1016/J.CONB.2020.11.003

Rong, F., Lisette Isenberg, A., Sun, E., & Hickok, G. (2018). The neuroanatomy of speech sequencing at the syllable level. *PloS one, 13*(10), Article e0196381. https://doi.org/10.1371/journal.pone.0196381

Sanchez-Bornot, J., Sotero, R. C., Kelso, J. A. S., Şimşek, Ö., & Coyle, D. (2024). Solving large-scale MEG/EEG source localisation and functional connectivity problems simultaneously using state-space models. *NeuroImage, 285*, Article 120458. https://doi.org/10.1016/j.neuroimage.2023.120458

Schirmer, A., Fox, P. M., & Grandjean, D. (2012). On the spatial organization of sound processing in the human temporal lobe: A meta-analysis. *NeuroImage, 63*(1), 137–147. https://doi.org/10.1016/j.neuroimage.2012.06.025

Srivastava, C. K., Gupta, R., & O'Reilly, J. A (2023). Neural generators of intensity mismatch negativity modelled with a recurrent neural network: A pilot study on the role of sound level transitions. In *BMEiCON 2023 - 15th biomedical engineering international conference*. https://doi.org/10.1109/BMEiCON60347.2023.10322021

Sun, R., Sohrabpour, A., Worrell, G. A., & He, B. (2022). Deep neural networks constrained by neural mass models improve electrophysiological source imaging of spatiotemporal brain dynamics. *Proceedings of the National Academy of Sciences of the United States of America, 119*(31), Article e2201128119. https://doi.org/10.1073/pnas.2201128119

Van Veen, B. D., Van Drongelen, W., Yuchtman, M., & Suzuki, A. (1997). Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Transactions on Biomedical Engineering, 44*(9), 867–880. https://doi.org/10.1109/10.623056

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., … Vázquez-Baeza, Y. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods, 17*(3), 261–272. https://doi.org/10.1038/s41592-019-0686-2

Werbos, P. J. (1990). Backpropagation through time: What it does and how to do it. *Proceedings of the IEEE, 78*(10), 1550–1560. https://doi.org/10.1109/5.58337

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin, 1*(6), 80. https://doi.org/10.2307/3001968

Wipf, D., & Nagarajan, S. (2009). A unified Bayesian framework for MEG/EEG source imaging. *NeuroImage, 44*(3), 947–966. https://doi.org/10.1016/J.NEUROIMAGE.2008.02.059

Zhu, J. D., Blanco-Elorrieta, E., Sun, Y., Szakay, A., & Sowman, P. F. (2022). Natural vs forced language switching: Free selection and consistent language use eliminate significant performance costs and cognitive demands in the brain. *NeuroImage, 247*, Article 118797. https://doi.org/10.1016/J.NEUROIMAGE.2021.118797