
This is an electronic reprint of the original article.
This reprint may differ from the original in pagination and typographic detail.

Haxel, Lisa; Ahola, Oskari; Kapoor, Jaivardhan; Ziemann, Ulf; Macke, Jakob H.

Personalized real-time inference of momentary excitability from human EEG

Published in:
NeuroImage

DOI:
[10.1016/j.neuroimage.2025.121547](https://doi.org/10.1016/j.neuroimage.2025.121547)

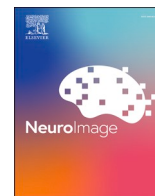
Published: 15/11/2025

Document Version
Publisher's PDF, also known as Version of record

Published under the following license:
CC BY

Please cite the original version:
Haxel, L., Ahola, O., Kapoor, J., Ziemann, U., & Macke, J. H. (2025). Personalized real-time inference of momentary excitability from human EEG. *NeuroImage*, 322, 1-14. Article 121547.
<https://doi.org/10.1016/j.neuroimage.2025.121547>

This material is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.



Personalized real-time inference of momentary excitability from human EEG

Lisa Haxel^{a,b,c,d,*} , Oskari Ahola^{c,d,e}, Jaivardhan Kapoor^{a,b},
Ulf Ziemann^{c,d,*}, Jakob H. Macke^{a,b,f,g,*}

^a Excellence Cluster Machine Learning, University of Tübingen, Tübingen, Germany

^b Tübingen AI Center, University of Tübingen, Tübingen, Germany

^c Department of Neurology and Stroke, University of Tübingen, Tübingen, Germany

^d Hertie Institute for Clinical Brain Research, University of Tübingen, Tübingen, Germany

^e Department of Neuroscience and Biomedical Engineering, Aalto University, Espoo, Finland

^f Department of Empirical Inference, Max Planck Institute for Intelligent Systems, Tübingen, Germany

^g Hertie Institute for AI in Brain Health, University of Tübingen, Germany

ARTICLE INFO

Keywords:

Transcranial magnetic stimulation
Electroencephalography
Cortical excitability
Real-time prediction
Closed-loop stimulation

ABSTRACT

The efficacy of transcranial magnetic stimulation (TMS) is often limited by non-adaptive protocols that disregard instantaneous brain states, potentially constraining therapeutic outcomes. Current EEG-guided approaches are hindered by their reliance on motor-evoked potentials (MEPs), which confound cortical and spinal excitability and restrict applications to the motor cortex, and a dependence on static biomarkers that cannot adapt to changing neurophysiological patterns. We introduce PRIME (Personalized Real-time Inference of Momentary Excitability), a deep learning framework that predicts cortical excitability, quantified by TMS-evoked potential (TEP) amplitude, from raw EEG signals. By targeting cortical excitability directly, PRIME provides a framework that could potentially extend brain state-dependent stimulation beyond the motor cortex, though validation in other cortical regions remains to be established. PRIME incorporates transfer learning and continual adaptation to automatically identify personalized biomarkers, allowing stimulation timing to be adapted across individuals and sessions. PRIME successfully predicts cortical excitability with minimal latency, providing a computational foundation for next-generation, personalized closed-loop TMS interventions.

1. Introduction

Transcranial magnetic stimulation (TMS) offers a non-invasive approach to modulate cortical excitability and connectivity. Its growing therapeutic applications in neurological and psychiatric disorders include depression, stroke recovery, and movement disorders (Bai et al., 2022; Hoogendam et al. 2010; Kricheldorf et al. 2022; Lefaucheur et al. 2020). However, despite widespread clinical adoption, patient outcomes remain highly variable, with more than half of treated individuals showing minimal or no clinical improvement (Vida et al. 2023; Ziemann and Siebner 2015; Ridding and Ziemann 2010). This variability partly stems from a fundamental limitation in conventional TMS design.

Standard protocols operate as "open-loop" systems, delivering pre-determined, static patterns of magnetic pulses that remain uniform across sessions and patients. This approach treats the brain as a passive,

static recipient of stimulation. Yet, modern neuroscience has established that the brain is better modeled as a dynamic system, with its internal state—characterized by neural oscillations, regional activity, and metabolic demands—varying over coarse-to-fine temporal scales (John et al. 2022). The neurophysiological and behavioral effects of TMS are therefore governed by principles of state-dependence, where the impact of stimulation depends not only on stimulus parameters, but also on the interaction between the stimulus and the brain's internal state (Karabanov et al. 2015; Bergmann 2018). Delivering fixed pulse trains into a dynamically fluctuating brain means that therapeutic effects become contingent on the chance alignment of pulses with neurophysiologically receptive moments.

To overcome the limitations of open-loop designs, a system must account for the brain's excitability state in real time. TMS-evoked potentials (TEPs), recorded via electroencephalography (EEG), directly

* Corresponding authors.

E-mail addresses: lisa.haxel@uni-tuebingen.de (L. Haxel), ulf.ziemann@uni-tuebingen.de (U. Ziemann), jakob.macke@uni-tuebingen.de (J.H. Macke).

<https://doi.org/10.1016/j.neuroimage.2025.121547>

Received 1 September 2025; Received in revised form 13 October 2025; Accepted 21 October 2025

Available online 22 October 2025

1053-8119/© 2025 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

index local cortical excitability and can be measured from any targeted brain region. Historically, extracting reliable TEPs on a single-trial basis has been challenging due to stimulus-related artifacts and a low signal-to-noise ratio (SNR). Recent advances, however, mitigate these issues by employing individualized, source-level dipole modeling to isolate genuine TEP signals from noise and enable their accurate single-trial extraction (Ahola et al. 2025). In contrast, motor-evoked potentials (MEPs), recorded peripherally via electromyography (EMG), provide robust measures of corticospinal excitability but are restricted to motor cortex stimulation. The clinical relevance of both biomarkers is underscored by findings that larger baseline evoked responses at standardized, individually calibrated intensities tend to show better therapeutic outcomes (Pellicciari et al. 2018; Bembenek et al. 2020; Strafella et al. 2023).

The substantial trial-to-trial variability observed in TEPs and MEPs is driven primarily by fluctuations in pre-stimulus neural activity (Voineskos et al. 2019). First-generation, brain state-dependent TMS protocols aimed to enhance efficacy by timing stimulation to predefined EEG biomarkers, such as particular alpha-band phases or power thresholds (Zrenner et al. 2018; Bergmann et al. 2019; Baur et al. 2020). Such approaches remain limited by their reliance on fixed biomarkers, which cannot capture the full complexity of moment-to-moment excitability fluctuations. Emerging evidence demonstrates that excitability arises from dynamic, network-level interactions and that its electrophysiological signatures are highly individualized, varying significantly across individuals and sessions (Metsomaa et al. 2021; Hussain and Quentin 2022; Ermolova et al. 2024; Khatri et al. 2025; Haxel, Ahola, et al. 2025; Ahola et al. 2025). While advanced closed-loop protocols propose to overcome the use of fixed biomarkers with adaptive feedback (Humaïdan et al. 2024), their implementation remains limited by a reliance on MEPs, restricting applications to the motor cortex and precluding their use in clinical contexts that target non-motor regions (Humaïdan et al. 2025).

To address these gaps—the reliance on fixed biomarkers, restriction to motor cortex applications, and inability to capture individualized excitability patterns—we introduce PRIME (Personalized Real-time Inference of Momentary Excitability), a deep learning framework that predicts TEP amplitudes from raw EEG data at the single-trial level in real-time. Unlike existing systems, PRIME processes EEG signals end-to-

end, automatically identifying and adapting to individualized biomarkers that directly reflect cortical responsiveness. Through pretraining on large-scale EEG-TMS datasets and real-time supervised continual learning (Haxel, Kapoor, et al. 2025), PRIME dynamically adjusts to changing brain activity patterns (Fig. 1).

We demonstrate the real-time predictive performance of PRIME under realistic hardware constraints using previously collected EEG-TMS data from single-pulse stimulation of the primary motor cortex (M1) in healthy adults. While the TEP-based approach is not inherently restricted to motor cortex—TEPs can theoretically be recorded from any cortical target, unlike MEP-based methods—empirical validation in association cortices, clinical populations, and repetitive stimulation protocols remains an important direction for future research. By eliminating dependence on peripheral muscle responses and fixed EEG features, PRIME provides a computational foundation for personalized closed-loop TMS systems that could potentially extend across different therapeutic contexts pending appropriate validation.

2. Methods

2.1. Participants and data acquisition

We analyzed single-pulse TMS-EEG data acquired as part of the European Research Council-funded ConnectToBrain project (grant no. 810,377). The data were collected at two research centers (University of Tübingen, Germany and Aalto University, Finland) and comprise recordings from 50 healthy, right-handed adults (26 female; mean age 27 ± 6 years).

The original study protocol was approved by the ethics committees of the University of Tübingen (810/2021BO2) and the Helsinki University Hospital (HUS/1198/2016) and was conducted in compliance with the Declaration of Helsinki. All participants provided written informed consent.

The dataset contains concurrent TMS-EEG and electromyography (EMG) recordings. The experimental protocol consisted of four blocks of 300 biphasic single pulses (1200 total) targeted at the motor hotspot of the left primary motor cortex (M1). Stimulation was delivered at an intensity of 110 % of the resting motor threshold (rMT), with a randomized inter-trial interval of 4.0–4.5 s (Aalto) or 2.5–3.5 s (Tübingen).

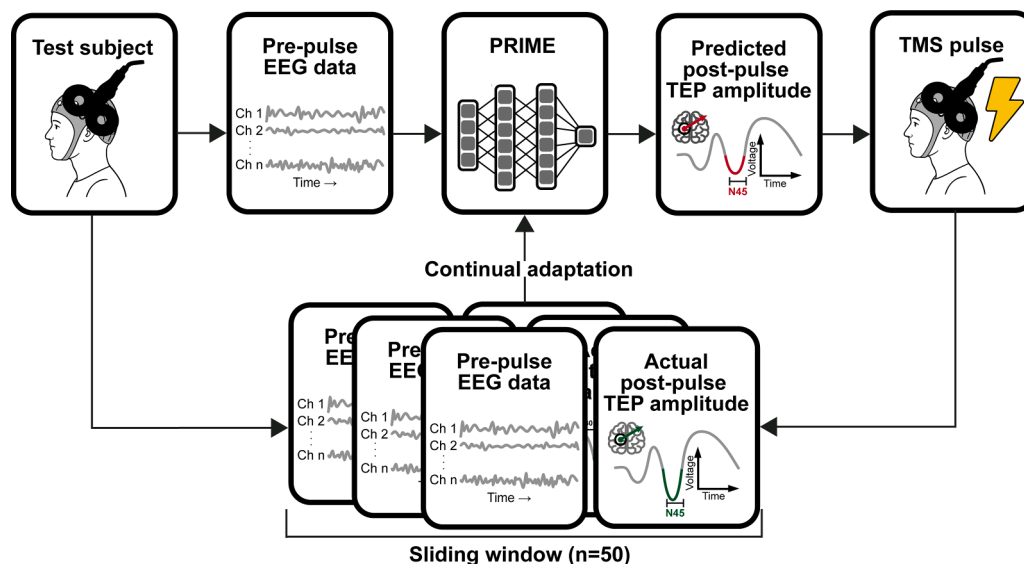


Fig. 1. Framework for real-time prediction of TMS-evoked potential amplitudes. Multi-channel electroencephalography (EEG) signals are continuously streamed and preprocessed in real-time. PRIME uses pre-stimulus EEG data to predict the amplitude of the N45 component of the upcoming TMS-evoked potential (TEP) waveform. Immediately following TMS, the actual TEP amplitude is extracted from post-stimulus EEG signals using a current-dipole approach and compared with the prediction. The resulting prediction error, together with recent pre-stimulus EEG data and corresponding TEP markers, are fed back to PRIME for continual updating of the predictive model between pulses.

High-density EEG data were recorded with TMS-compatible NeurOne systems (64-channel at Aalto, 128-channel at Tübingen; Bittium Ltd, Finland) at a sampling rate of 5 kHz. EMG was recorded from the right abductor pollicis brevis (APB) and first dorsal interosseus (FDI) muscles.

Site-specific equipment included a Nexstim NBS 5.2.4 or NBT 2.2.4 system (Nexstim Plc, Finland) with a Nexstim cooled coil at Aalto, and a MagVenture R30 or X100 stimulator (MagVenture Inc, United States) with a Cool-B65 coil at Tübingen. Neuronavigation was performed using individual T1-weighted magnetic resonance images (MRIs) with Nexstim (Aalto) or Localite (Tübingen; Localite GmbH, Germany) systems. To mask the TMS click sound, individually calibrated noise was delivered via ER-3C insert earphones (Etymotic Research Ltd, United States). T1-weighted MRIs were acquired using MPRAGE sequences with 3T Siemens Skyra or Prisma scanners.

2.2. Data preprocessing

Our preprocessing pipeline was designed to simulate a real-time analysis workflow and consisted of two primary stages: an offline *calibration* stage to define subject-specific parameters and a simulated online *application* stage where these fixed parameters were applied to subsequent trials.

2.3. Calibration phase

The first 125 trials per participant were initially selected to ensure a minimum of 100 valid trials post-rejection, from which all preprocessing parameters were derived. The trial count was increased until at least 100 valid trials were acquired.

2.3.1. Downsampling and epoching

EEG data, originally sampled at 5 kHz, were polyphase downsampled to 1 kHz for pre-stimulus ($[-505, -5]$ ms) and post-stimulus ($[-30, 100]$ ms) epochs separately. EMG data ($[-500, 200]$ ms) were likewise downsampled to 1 kHz.

2.3.2. Channel rejection

Bad channels were iteratively rejected from both pre- and post-stimulus data independently. Pre-stimulus criteria included z-scored median absolute deviation (thresholds $[-3, 3]$), high-frequency power (30–47 Hz; threshold > 5), and lag-1 autocorrelation (thresholds $[-4, 4]$). Post-stimulus channels were rejected based on the area under the curve in the average response (20–60 ms; threshold > 3). Rejected channels were reconstructed using spherical spline interpolation.

2.3.3. Independent component analysis (ICA)

Ocular artifacts were identified using Infomax ICA (Bell and Sejnowski 1995; Lee et al. 1999) on pre-stimulus data ($[-1100, -5]$ ms) that was bandpass-filtered (1–100 Hz, 2nd-order Butterworth). ICA was done using the minimum number of principal components explaining $> 99\%$ of the variance. Ocular components with a probability > 0.9 (via ICLabel (Pion-Tonachini et al. 2019)) were removed; this threshold was lowered to 0.7 if fewer than two components were initially identified (often two components exist; one for blinks and one for lateral eye movement). Trials with high ocular activity (z-score > 2 in the IC time series) in the post-stimulus interval ($[15, 100]$ ms) were rejected.

2.3.4. Pre-stimulus EEG preprocessing

Data were bandpass-filtered (2–47 Hz, 2nd-order Butterworth), average-referenced, and subjected to iterative trial rejection based on median absolute deviation (z-score thresholds $[-8, 4]$) and single-channel deviation (z-score threshold ± 5).

2.3.5. Post-stimulus EEG preprocessing

Prior to ICA, the data were baseline-corrected ($[-25, -15]$ ms), the pulse artifact window ($[-14, 14]$ ms) was tapered using a 1-hanning

window, and average referenced after interpolating bad channels. Following ICA, data were again baseline-corrected ($[-25, -15]$ ms), average-referenced, and cropped to $t > 0$ ms. We suppressed non-neural artifacts using the source-utilizing noise-discarding (SOUND) algorithm (Mutanen et al. 2018) ($\lambda = 0.01$, tolerance = 10^{-9}) and TMS-related muscle artifacts with the fixed-component Signal-Space Projection and Source-Informed Reconstruction (SSP-SIR) algorithm (Mutanen et al. 2016). For SSP-SIR, we removed components explaining 99% of the variance in supra-100-Hz-filtered average data weighted by the artifact kernel estimated using 50-ms sliding windows. Then, an extended TMS pulse artifact window ($(0, 15)$ ms) was tapered. Finally, trials were iteratively rejected based on global (median absolute deviation z-score > 5) and local (single-channel z-score ± 5) amplitude criteria in the $[20, 60]$ ms window.

2.3.6. EMG preprocessing

EMG data were high-pass filtered (2 Hz, 4th-order Butterworth). Powerline noise was removed by iteratively subtracting 50- and 100-Hz sine fits calibrated on the pre-stimulus interval ($[-200, -15]$ ms). Trials with pre-stimulus muscle activation (min-max difference $> 50 \mu\text{V}$) were rejected. For analysis, we selected the single EMG channel with the highest average MEP amplitude ($[20, 50]$ ms).

2.4. Application phase and real-time compatibility

The application phase was designed to emulate online, trial-by-trial processing. All preprocessing parameters, including channel interpolation matrices, trial rejection statistics, and frequency-based, ICA, SOUND, and SSP-SIR filters, were fixed during the calibration phase and applied without modification to all subsequent data. This approach ensures that TEP amplitude estimates are not biased by changing preprocessing characteristics over time. The computational latency of each step in this pipeline was assessed to confirm its compatibility with real-time experimental requirements.

2.5. TEP component extraction

To analyze transcranial magnetic stimulation (TMS)-evoked potentials (TEPs), we constructed individualized spatiotemporal filters following (Ahola et al. 2025) based on current dipole modeling (Henderson et al. 1975; Sarvas 1987). This approach mitigates the low signal-to-noise ratio (SNR) and various neural and non-neural artifacts present in single-trial TEPs. Our analysis focused on the N45 component, thought to reflect the balance of cortical excitation and inhibition (Darmani and Ziemann 2019). We defined a target time interval of $\mathbf{T} = [38, 50]$ ms post-TMS to avoid contamination of early-latency TMS-related artifacts and later-latency sensory-evoked activity (Gordon et al. 2021; Ahola et al. 2025).

2.5.1. Dipole estimation

Individualized dipole models were derived from the first 100 valid calibration trials. For each time point $t \in \mathbf{T}$ in the averaged EEG data $\bar{\mathbf{m}}_t$, we estimated the optimal dipole position $\hat{\mathbf{p}}_t$ on the cortex and its corresponding moment $\hat{\mathbf{Q}}_t$ as

$$\{\hat{\mathbf{Q}}_t, \hat{\mathbf{p}}_t\} = \underset{\mathbf{p} \in \mathbb{R}^3}{\operatorname{argmax}} R^2(\bar{\mathbf{m}}_t, \mathbf{L}_p \mathbf{Q}_t),$$

where \mathbf{L}_p is the leadfield matrix for a dipole at position \mathbf{p} , and $\mathbf{Q}_t = \mathbf{L}_p^\dagger \bar{\mathbf{m}}_t$ is the single-dipole solution obtained via the Moore-Penrose pseudoinverse. The leadfield was derived from the fsaverage brain template (Fischl et al. 1999), comprising 20,484 sources, and mapped to EEG sensors using the standard_1005 electrode locations.

2.5.2. Individualized time-window selection

To account for inter-individual variability in response timing, we

defined a personalized, stable, high-amplitude time window $\hat{\mathbf{w}}$ for each participant. This window was identified by penalizing deviations in dipole position and orientation while rewarding high signal amplitude as

$$\hat{\mathbf{w}} = \underset{\mathbf{w} \in \mathcal{T}}{\operatorname{argmax}} \frac{r_a(|\mathbf{w}|)^\alpha}{r_a(\Delta \bar{p}_{\mathbf{w}}) + r_a(\Delta \bar{\phi}_{\mathbf{w}}) + r_d(\min(\|\mathbf{Q}\|_{\mathbf{w}}))},$$

where the window duration $|\mathbf{w}|$ was constrained to be between 3 and 6 ms. In the equation, $\|\mathbf{Q}\|_{\mathbf{w}}$ represents the dipole amplitudes, while $\Delta \bar{p}_{\mathbf{w}}$ and $\Delta \bar{\phi}_{\mathbf{w}}$ denote the average positional and angular dipole deviances from the mean within window \mathbf{w} . The function $r_{a/d}(\cdot)$ ranks windows in ascending or descending order, and α was set to 1.5. The final dipole filter position $\hat{\mathbf{p}}$ and orientation $\hat{\mathbf{o}}$ were defined as the amplitude-weighted average over the optimal window $\hat{\mathbf{w}}$.

2.5.3. Single-trial amplitude extraction

For each trial i , we estimated the dipole amplitude \hat{a}_i using two approaches. For the **fixed-orientation** model, the scalar amplitude was estimated as

$$\hat{a}_i = \underset{a \in \mathbb{R}, t \in \hat{\mathbf{w}}}{\operatorname{argmax}} R^2(\mathbf{m}_{i,t}, a_{i,t} \mathbf{L}_{\mathbf{p}} \hat{\mathbf{o}}),$$

where $a_{i,t}$ is the scalar solution to the least-squares problem $\mathbf{m}_{i,t} = a_{i,t} \mathbf{L}_{\mathbf{p}} \hat{\mathbf{o}}$. Given that the estimated scalar amplitude can be negative due to noise or polarity reversal, we used the absolute amplitude as the final measure of cortical excitability. For the **free-orientation** model, the amplitude was estimated as

$$\hat{a}_i = \|\hat{\mathbf{Q}}_i\| = \underset{\mathbf{Q}, t \in \hat{\mathbf{w}}}{\operatorname{argmax}} R^2(\mathbf{m}_{i,t}, \mathbf{L}_{\mathbf{p}} \mathbf{Q}_{i,t}).$$

A similar procedure was performed for the P60 component within $T = [55, 70]$ ms, using a window duration constraint of $4 \leq |\mathbf{w}| \leq 7$ ms.

2.6. Target label generation

We implemented a causal, adaptive normalization procedure to transform single-trial TEP and MEP amplitudes into a stable excitability metric that is robust to slow temporal drifts. This process mirrored the main pipeline, with a calibration phase to derive statistics and an application phase to generate labels.

2.6.1. Calibration: Normalization parameter estimation

During the calibration phase, we derived a fixed set of normalization parameters using the single-trial amplitudes extracted from the same set of initial trials used for the main preprocessing (≥ 100 valid trials). First, this amplitude time series was causally detrended using an Exponentially Weighted Moving Average (EWMA) with a 25-trial span. To ensure stable parameter estimation, the initial 25 trials were disregarded to account for the filter’s warm-up period. From the remaining stable, detrended values, we computed and stored three key parameters: the mean (μ_{cal}), standard deviation (σ_{cal}), and an empirical cumulative distribution function (ECDF_{cal}).

2.6.2. Application: Real-time label generation

In the application phase, these calibrated parameters were used to generate a quantile-normalized label for each new trial. For a given trial’s raw amplitude, a_{raw} , we first obtained its detrended value, $a_{detrend}$, from the EWMA filter. This value was then standardized using the stored calibration statistics to produce a z-score, $z = (a_{detrend} - \mu_{cal}) / \sigma_{cal}$. Finally, the z-score was transformed via the stored ECDF_{cal} to yield the final label $y = \text{ECDF}_{cal}(z)$. The resulting label $y \in [0, 1]$ represents the trial’s excitability as a percentile rank relative to the initial, stable calibration distribution.

3. Model architecture

The PRIME architecture (Table 1) integrates a convolutional front-end for spatiotemporal feature extraction with a state-space core for modeling long-range dependencies. The model consists of three primary components:

- Spatiotemporal convolutional front-end:** Inspired by DeepConvNet (Schirmeister et al. 2017), this module uses a temporal convolution to extract initial time-domain features, followed by a spatial convolution to learn channel correlations. Pooling and dropout are applied to regularize the features and reduce dimensionality.
- State-space modeling:** The extracted feature sequence is processed by a bidirectional S4 block (Gu et al. 2022; Kapoor et al. 2024) to capture long-range temporal dependencies within the EEG signals within each 50 ms pre-stimulus EEG epoch. Because these fixed-length segments are fully available at prediction time, bidirectional processing remains compatible with real-time use while enabling the model to integrate both forward- and backward-directed dynamics. This design provides richer temporal feature representations than strictly unidirectional processing and is well suited to the short, windowed inputs typical of real-time EEG decoding applications.
- Prediction head:** The processed temporal features are passed through an adaptive global average pooling layer followed by a linear classifier. Global pooling provides invariance to small temporal shifts within the epoch and reduces the dimensionality of the feature maps, thereby lowering overfitting risk and computational cost. In our configuration, this reduces the parameters of the final layer from 626 (flattened 25×25 features plus bias) to 26 (25 pooled features plus bias), contributing to PRIME’s fast inference times (< 2 ms per trial; Fig. 2f). The classifier output is optimized using a binary cross-entropy (BCE) loss.

3.1. Training and adaptation framework

PRIME employs a hierarchical training strategy designed to achieve robust cross-subject generalization while enabling rapid online adaptation to individual neural characteristics. The framework operates through three sequential phases: population-level pretraining establishes general EEG-to-TEP mappings across subjects, subject-specific calibration adapts the model to individual neural patterns, and continuous trial-by-trial adaptation refines predictions during online deployment (see Table 2 for key hyperparameters for each phase).

3.2. Phase 1: Population-level pretraining

The initial phase establishes robust EEG-to-TEP mappings using data from a large population of subjects. To enhance statistical homogeneity across subjects, we apply an optional two-step alignment process to each training subject’s data. The process first whitens each subject’s data based on its own covariance matrix, then rotates the whitened data from all subjects to a common space defined by the global average covariance

Table 1

PRIME architecture summary. The input shape corresponds to (channels, timepoints).

Component	Output shape	Parameters
Input	(60, 50)	–
Spatiotemporal front-end	(25, 25)	,800
S4 Block	(25, 25)	,725
Prediction head	(1)	,26
Total		46,551

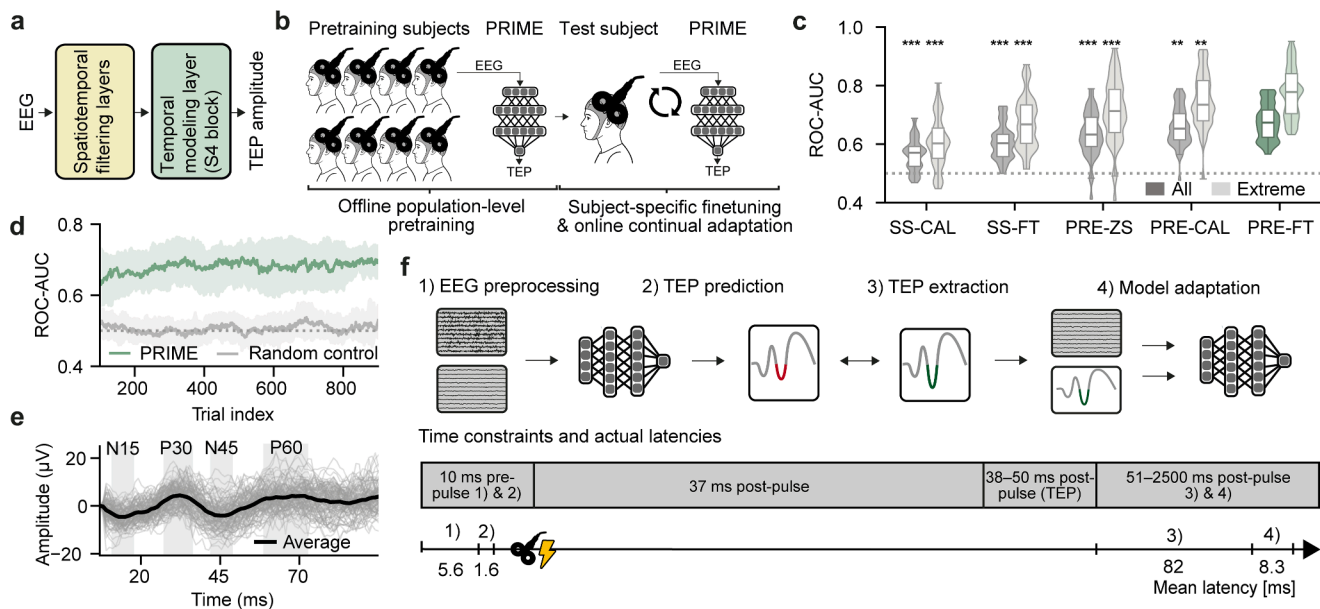


Fig. 2. PRIME achieves robust real-time prediction of cortical excitability through multi-stage learning and personalization. **a.** PRIME processes EEG data using spatial-temporal convolutions to identify local patterns in the signal, followed by structured state-space (S4) modeling to capture long-range temporal dependencies and predict upcoming TEP amplitudes. **b.** The training process involves three stages: first, a general model is pretrained on a population of subjects. This model is then personalized for a new subject through an initial calibration, followed by continual real-time adaptation on a trial-by-trial basis. **c.** Predictive performance (ROC-AUC) across 50 datasets (approximately 955 trials each) comparing five training conditions: subject-specific calibration (SS-CAL), subject-specific finetuning (SS-FT), pretrained zero-shot (PRE-ZS, no subject-specific calibration), pretrained calibrated (PRE-CAL, calibration without further finetuning), and pretrained finetuned (PRE-FT, calibration plus finetuning). Performance is shown for all trials (*All*) and extreme TEP amplitudes (*Extreme*, highest and lowest 25 %). PRIME (PRE-FT) significantly outperformed all other conditions (** $p < 0.01$, *** $p < 0.001$, paired one-sided t -tests, Benjamini-Hochberg corrected). **d.** Predictive accuracy (median ROC-AUC with interquartile range) computed using a sliding window (size = 100) shows rapid model improvement after calibration. PRIME consistently outperformed random controls with shuffled labels. **e.** Representative TEP waveforms from channel F3 (near stimulation site) during the first 100 trials. Individual trials (grey) and their average (black) are shown for the post-stimulus window (10–99 ms). **f.** Computational latency per trial (median across 50 datasets): EEG preprocessing (5.6 ms on CPU), prediction (1.6 ms on GPU), TEP extraction (82 ms on CPU), and model adaptation (8.3 ms on GPU). Pre-pulse processing (7.2 ms) meets real-time constraints (10 ms window), while all post-pulse computations fit within the 2.5 s inter-stimulus interval.

Table 2

Training configuration across the three learning phases.

Parameter	Population pretraining	Subject calibration	Online adaptation
<i>Data setting</i>			
Subject split	2-fold CV (50 % train)	From test split	From test split
Trials used	All from training subjects	First 100 trials	Sliding window (50 trials)
<i>Training hyperparameters</i>			
Epochs	100	50	1 (per trial)
Batch size	64	50	50 (window size)
Learning rate	3×10^{-4}	1×10^{-4}	1×10^{-4}
Warm-up trials	N/A	N/A	50
<i>Optimization & Loss</i>			
Optimizer	AdamW	AdamW	AdamW
Loss Function	BCE with Logits	BCE with Logits	BCE with Logits

of the entire training population. We employ a 2-fold cross-subject validation scheme where the model trains on half of the subjects and evaluates on the held-out half for each fold. The model undergoes pre-training for 100 epochs using the AdamW optimizer with a binary cross-entropy with logits loss function.

3.3. Phase 2: Subject-specific calibration

Before online deployment, the pretrained model undergoes brief calibration using an initial set of data from the target subject. The model finetunes on the first 100 trials for 50 epochs, serving two critical

purposes. Model specialization allows the weights, initialized from population-level pretraining, to adapt to the specific neural patterns of the new subject. During this phase, model parameters are updated regardless of the online adaptation strategy selected for the subsequent phase. Alignment initialization uses the same 100-trial calibration set to compute an initial reference covariance matrix that seeds the optional unsupervised alignment process in the online phase, providing a subject-specific statistical baseline.

3.4. Phase 3: Trial-by-trial online adaptation

Following calibration, the model transitions to continuous online adaptation through three sequential steps for each incoming trial: prediction, optional unsupervised alignment, and supervised finetuning.

1. Prediction

The model first generates a prediction for the current, unseen trial based on its pre-stimulus EEG features.

2. Unsupervised euclidean alignment (Optional)

To investigate the utility of mitigating statistical distribution shifts between subjects, we implement an optional unsupervised alignment step. This preprocessing approach standardizes the second-order statistics of input data by whitening each trial $\mathbf{X} \in \mathbb{R}^{C \times T}$ (where C and T represent the number of channels and time points, respectively) as

$$\tilde{\mathbf{X}} = \mathbf{M}\mathbf{X}$$

where $\mathbf{M} = \mathbf{C}_{\text{ref}}^{-1/2}$ represents the inverse square root of a reference covariance matrix \mathbf{C}_{ref} . This reference covariance updates dynamically after each trial using an Exponential Moving Average (EMA) over the covariances of the 50 most recent trials

$$\mathbf{C}_{\text{ref}}^{(t)} = \beta \cdot \mathbf{C}_{\text{ref}}^{(t-1)} + (1 - \beta) \cdot \mathbf{C}_{\text{trial}}^{(t)}$$

where $\mathbf{C}_{\text{trial}}^{(t)}$ represents the covariance of the current trial and the smoothing factor β is set to 0.99. This alignment update occurs after prediction using only the pre-stimulus trial's EEG data, preventing label information leakage into the adaptation process.

3. Supervised continual finetuning (CFT)

After the true label for a trial is revealed, the model parameters undergo supervised updating. This finetuning begins only after collecting an initial buffer of 50 trials to ensure stable gradient estimation. Subsequently, for each new trial, the model trains for a single epoch on data within this 50-trial sliding window. We evaluate three distinct CFT strategies to understand the trade-offs between adaptability and representational stability. Full finetuning updates all model parameters, allowing the entire network to adapt to the incoming data stream and offering maximum flexibility. Decision-only finetuning preserves the learned feature hierarchy from pretraining and mitigates catastrophic forgetting by freezing feature extraction layers and updating only the parameters of the final learnable layer (either linear or convolutional) in the network. Threshold-only finetuning represents the most restrictive and computationally efficient strategy, learning only a single scalar parameter that acts as an adaptive bias applied to the model's output logits, effectively shifting the decision threshold without altering feature extractor or classifier weights.

3.5. Evaluation metrics

We assess model performance using the Area Under the Receiver Operating Characteristic Curve (ROC-AUC) due to its robustness to class imbalance. To dissect the contributions of each framework component, we evaluate performance at three distinct stages that isolate different aspects of the learning process.

Pretrained zero-shot (PRE-ZS) provides a baseline for out-of-the-box generalization by applying the general population model directly to all trials from a new subject without any subject-specific calibration or finetuning. Pretrained calibrated (PRE-CAL) isolates the benefit of the initial calibration phase by evaluating model performance on remaining trials after calibration on the first 100 trials, without further online adaptation. Pretrained finetuned (PRE-FT) represents the final performance of the complete framework, measuring overall ROC-AUC across all predictions made during continuous, trial-by-trial online adaptation following calibration.

For each stage, we report performance on two data subsets to provide comprehensive evaluation: across all available test trials ("All") and specifically for trials with extreme TEP amplitudes (highest and lowest 25 % of responses). Extreme trials represent the most clinically relevant scenarios where clear distinctions between high and low excitability states are crucial for therapeutic applications. The final reported scores for each condition are averaged across all subjects and test folds.

3.6. Experiments

Our experimental evaluation was designed to: 1) systematically assess the performance of the full PRIME framework against ablated and alternative conditions; 2) benchmark it against established deep learning architectures; 3) analyze its sensitivity to key hyperparameters; and 4) test the generalizability of its learned representations to different physiological targets. To determine if our proposed framework

significantly outperformed the alternatives, we compared the performance of PRIME against each other condition using a paired, one-sided t -test. P-values were corrected for multiple comparisons across the datasets using the Benjamini-Hochberg procedure (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$)

3.7. Ablation and comparison conditions

To systematically evaluate the contribution of each framework stage, we compared the full PRIME system against a hierarchy of ablated and alternative conditions. Each condition isolates the impact of population-level pretraining, subject-specific calibration, and continual finetuning through controlled comparisons.

Subject-specific calibration only (SS-CAL) establishes a baseline for performance using only a small, subject-specific dataset. The model initializes with random weights without pretraining and trains exclusively on the 100-trial calibration set. Model weights are then frozen and performance is evaluated on remaining trials. **Subject-specific finetuning (SS-FT)** represents a "from-scratch" learning approach that omits population-level pretraining. The model initializes with random weights and follows the same two-phase training protocol as the full framework: initial training on the 100-trial calibration set followed by trial-by-trial continual finetuning on all remaining data.

Pretrained zero-shot (PRE-ZS) serves as the primary baseline for generalization from the population model by applying the pretrained model directly to new subject data without calibration or finetuning, with weights remaining frozen during evaluation. **Pretrained + calibrated (PRE-CAL)** isolates the benefit of the calibration step by updating the population-pretrained model using the 100-trial subject-specific calibration set, then freezing weights and evaluating performance on remaining trials without continual finetuning.

Pretrained + finetuned (PRE-FT) represents the complete proposed PRIME framework, leveraging all three stages: model initialization from population-level pretraining, specialization during subject-specific calibration, and continuous updating through trial-by-trial finetuning on all subsequent data.

3.8. Benchmark architectures

To contextualize the performance of PRIME, we benchmarked it against four established deep neural network architectures from the EEG decoding literature. Given the short input sequences used in our task (50 time points, corresponding to 50 ms of data), we appropriately scaled down the kernel sizes and pooling layers of the original benchmark architectures to prevent excessive temporal dimension reduction and ensure fair comparison. **ShallowConvNet** provides a compact architecture inspired by the filter bank common spatial pattern (FBCSP) pipeline (Schirrmester et al. 2017), employing temporal convolution followed by spatial filtering to extract spatiotemporal features. **DeepConvNet** implements a deeper, multi-layer architecture composed of stacked convolutional blocks (Schirrmester et al. 2017), designed to learn feature hierarchies from low-level signal characteristics to more abstract and discriminative representations. **EEGNet** offers a highly compact and efficient architecture that utilizes depthwise and separable convolutions to learn temporal and spatial filters independently (Lawhern et al. 2018). **ATCNet** presents a hybrid architecture that augments a convolutional frontend with multi-head self-attention and a temporal convolutional network (TCN) (Altaheri et al. 2023), enabling the model to capture long-range temporal dependencies and explicitly weight the importance of learned features.

3.9. Architectural ablations

To isolate the contribution of the Structured State Space (S4) component within the PRIME architecture, we designed and evaluated two ablated variants that systematically remove or replace this core

component.

PRIME without S4 removes the entire S4 temporal modeling block, leaving the model to consist solely of the spatiotemporal convolutional frontend feeding directly into global average pooling and the final linear classifier. This ablation measures performance attributable to the convolutional feature extractor alone. **PRIME with convolutional core** replaces the S4 block with a standard temporal feature extractor comprising a stack of conventional 1D convolutional layers of equivalent depth, enabling direct comparison between S4's ability to model long-range dependencies and traditional CNN-based approaches.

3.10. Hyperparameter sensitivity analysis

We investigated the sensitivity of PRIME to three key hyperparameters using the full PRE-FT configuration as baseline. We varied the number of trials used for subject-specific calibration, testing the default 100 trials against 200, 300 and 400 trials to assess calibration requirements. We evaluated the impact of pre-stimulus EEG window size by testing windows of 50 ms (default), 100 ms, 200 ms, 300 ms, 400 ms, and 500 ms to determine optimal temporal resolution. We examined the importance of window temporal location by testing offsets ending 10 ms (default), 20 ms, 30 ms, 40 ms and 50 ms before the TMS pulse to identify the most predictive temporal window placement.

3.11. Generalizability to different physiological targets

To assess whether features learned by PRIME generalize beyond predicting the N45 component, we conducted two additional experiments using the best-performing PRE-FT configuration. We tested in-domain generalizability by retraining the framework to predict the later P60 TEP component, examining whether similar cortical excitability mechanisms underlie different TEP responses. We evaluated cross-domain generalizability by repurposing the framework to predict MEPs as a measure of corticospinal excitability derived from EMG signals, to determine whether EEG features predictive of cortical responsiveness extend to downstream motor pathway excitability.

3.12. Real-time latency analysis

To assess the real-time feasibility of PRIME, we quantified the computational latency of its core online components under realistic hardware constraints. All benchmarks were conducted in a PyTorch environment (Paszke et al. 2019) using an 8-core CPU (AMD EPYC 7302) and a consumer-grade GPU (NVIDIA GeForce RTX 2080Ti with 11 GB VRAM) to reflect standard laboratory computing setups.

We measured wall-clock time in milliseconds for four distinct computational stages across all subjects, grouping operations by their occurrence relative to the TMS pulse. For GPU benchmarks, we ensured accurate timings by synchronizing the device before and after each measurement to account for asynchronous operation completion. Prediction and model adaptation measurements were averaged over 100 repetitions, preceded by 10 warm-up runs to stabilize system state and eliminate initialization overhead.

3.12.1. Pre-stimulus operations

These operations must complete within the interval between EEG recording and TMS pulse delivery to maintain real-time performance. EEG preprocessing encompasses the time required to process incoming pre-stimulus EEG segments, including filtering, average-referencing, and threshold-based trial rejection. TEP prediction measures the time for a single forward pass of the PRIME model to generate predictions from preprocessed EEG data.

3.12.2. Post-stimulus operations

These operations occur during the inter-stimulus interval after TMS pulse delivery, allowing more relaxed timing constraints. Post-stimulus

processing and TEP extraction quantifies the time required for all post-pulse data handling, encompassing initial processing of post-stimulus EEG data and subsequent extraction of the relevant TEP component (e.g., N45) that serves as the prediction target for model adaptation. Model adaptation measures the time to perform one complete supervised finetuning step, including forward pass, loss calculation, back-propagation, and optimizer weight update using a batch of the 50 most recent trials, consistent with the online learning protocol.

3.13. Neurophysiological feature interpretability

To identify the neurophysiological features underlying the predictive capabilities of PRIME, we conducted occlusion-based sensitivity analysis on final, subject-specific finetuned models. This approach systematically removes distinct spatial or spectral components from input EEG signals and quantifies the resulting degradation in model performance, measured as the change in ROC-AUC (Δ ROC-AUC). Larger performance drops indicate higher importance of the occluded feature. We assessed statistical significance of performance drops using one-sided Wilcoxon signed-rank tests, with p-values corrected for multiple comparisons using the Benjamini-Hochberg procedure. Three complementary analyses created a comprehensive interpretability map of the decision-making process of PRIME.

3.13.1. Spatial importance mapping

To identify spatial origins of predictive information, we occluded broadband EEG signals from each electrode and its eight nearest neighbors. The resulting performance drop for each occlusion was mapped to the corresponding electrode location, creating topographical maps of spatial importance across the scalp. This analysis reveals which cortical regions contribute most critically to excitability predictions.

3.13.2. Frequency band importance

To determine the contribution of distinct neural oscillations, we selectively removed individual frequency bands across all electrodes using band-stop filters (4th order Butterworth). We evaluated four canonical bands: theta (4–8 Hz), alpha (8–13 Hz), beta (13–25 Hz), and gamma (25–47 Hz). The aggregate drop in ROC-AUC for each occluded band reflects its overall predictive value for cortical excitability assessment.

3.13.3. Spatially resolved frequency analysis

To create detailed spatio-spectral profiles, we combined spatial and frequency approaches by systematically occluding each frequency band at specific electrode locations (including eight nearest neighbors). This procedure mapped the precise cortical locations from which predictive oscillatory activity originates, providing spatially and spectrally resolved feature importance maps that reveal the neurophysiological basis of the excitability predictions of PRIME.

3.14. Data availability

Data are available upon reasonable request from the corresponding authors. Compliance with EU data protection law (Regulation (EU) 2016/679) is required. An "End User Licence Agreement" (EULA) will be required before data will be released. Data will be shared on a project specific basis.

De-identified data is available from the corresponding authors upon reasonable request, subject to ethical approval and data agreements.

3.15. Code availability

Analysis code along with relevant documentation is publicly available on GitHub at <https://github.com/mackelab/PRIME/>

3.16. Ethics declarations

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.

4. Results

4.1. PRIME accurately predicts cortical excitability in real-time

We developed PRIME to address the limitations of static, open-loop neuromodulation by predicting single-trial cortical excitability directly from pre-stimulus EEG recordings. PRIME uses TMS-evoked potential (TEP) amplitudes as a direct truth measure of cortical responsiveness, providing a computational foundation for personalized closed-loop stimulation. The model's architecture is designed to process the raw EEG signal from end-to-end, meaning it learns to identify predictive features automatically without depending on predefined biomarkers like alpha- or beta-band power. To achieve this, the model first analyzes local patterns across nearby electrodes and short time windows, then captures how these patterns evolve over longer timescales to learn the complex temporal dependencies within the signal (Fig. 2a).

To prepare PRIME for practical use and test its performance, we designed a multi-stage training and adaptation process. Using a 2-fold cross-validation scheme across 50 subjects (954 ± 149 valid trials per subject), we first trained a general "population model" on data from half the subjects. For each new individual in the other half (the test set), we then calibrated this population model by finetuning it on their first 100 trials, allowing it to adapt to their unique neural characteristics. Finally, during the remainder of the session, the model continually adapted its predictions on a trial-by-trial basis.

We compared the full PRIME system (PRE-FT), which incorporates population-level pretraining, subject-specific calibration, and continual adaptation, against four control conditions to systematically evaluate the contribution of each training stage. To isolate the effects of personalization and ongoing learning, we tested a generic "zero-shot" model using only population-level pretraining (PRE-ZS) and a model that included initial calibration but omitted continual adaptation (PRE-CAL). Finally, to assess the value of population-level pretraining itself, we evaluated two models trained exclusively on single-subject data: one using only the initial calibration trials (SS-CAL) and another that also incorporated continual adaptation (SS-FT). In all conditions, TEPs and preprocessing parameters (e.g., ICA filters, channel interpolation matrices, artifact rejection thresholds) were individualized during the calibration phase of each subject and then held fixed for the remainder of the session, ensuring that performance differences across models reflected only the learning strategy rather than variability in preprocessing or target definition.

We assessed performance using ROC-AUC across all trials and, additionally, for trials with extreme TEP amplitudes (highest and lowest 25 % of responses) (Fig. 2c, Table S1). The complete PRIME system (PRE-FT) achieved superior predictive accuracy compared to all other conditions, with median ROC-AUC values of 0.68 for all trials and 0.77 for extreme trials. Statistical analysis confirmed that PRIME significantly outperformed all comparison conditions ($p < 0.01$, paired one-sided t -tests with Benjamini-Hochberg correction). Population-level pretraining provided substantial benefits over training from scratch, with PRE-CAL achieving 0.66 ROC-AUC compared to 0.56 for SS-CAL. Analysis of temporal learning dynamics using a sliding window of 50 trials revealed that PRIME consistently maintained superior performance throughout experimental sessions, substantially outperforming a random control model using shuffled TEP labels (Fig. 2d).

PRIME meets computational latency requirements essential for real-time implementation across all subjects on both GPU and CPU hardware. Latency analysis using an 8-core CPU and NVIDIA 2080Ti GPU revealed mean processing times of 5.6 ms for pre-stimulus EEG preprocessing

(CPU; 99th percentile: 7.1 ms) and 1.6 ms for GPU-based TEP prediction (99th percentile: 1.7 ms) or 2.1 ms for CPU-based prediction (99th percentile: 2.4 ms). The combined pre-pulse computation time is 7.2 ms (GPU) or 7.7 ms (CPU)—both, well within the 10 ms constraint required to maintain the reported predictive performance. Post-pulse computations, including TEP extraction (CPU; 82.2 ms, 99th percentile: 101.8 ms) and model adaptation (8.3 ms on GPU, 99th percentile: 9.0 ms; 35.3 ms on CPU, 99th percentile: 40.1 ms), operate comfortably within the 2.5 s inter-stimulus interval typical of single-pulse TMS protocols (Fig. 2e and Fig. S1).

4.2. Structured state-space modeling enables superior temporal feature learning

The PRIME architecture is designed to extract predictive features from multi-channel EEG time series through a combination of spatial and temporal processing components (Fig. 3a). The model begins with temporal convolution to learn time-domain features, followed by spatial convolution across all EEG channels to create spatiotemporal representations. These features are processed through normalization, pooling, and dropout layers before entering the core component: a bidirectional structured state-space (S4) block (Gu et al. 2022). The S4 block is optimized for capturing long-range temporal dependencies in neural signals and includes residual connections to facilitate stable training. A linear classifier with global average pooling generates the final TEP amplitude prediction.

To validate the architectural design of PRIME, we compared performance against four established EEG decoding architectures: ShallowConvNet (Schirrneister et al. 2017), DeepConvNet (Schirrneister et al. 2017), EEGNet (Lawhern et al. 2018), and ATCNet (Altaheri et al. 2023). We also tested two ablated variants to isolate the contribution of the S4 component: PRIME without the S4 block (PRIME w/o S4), which removes the temporal modeling layer entirely and uses only the convolutional frontend, and PRIME with conventional convolution (PRIME w/ Conv), which replaces the S4 block with standard 1D convolutional layers of equivalent depth.

Across all 50 datasets, the complete PRIME model demonstrated significantly higher predictive accuracy than all comparison models and ablated variants (Fig. 3b, Table S2; $p < 0.001$, paired one-sided t -tests, Benjamini-Hochberg corrected). The performance degradation in ablated models revealed the critical importance of structured temporal modeling for cortical excitability prediction. Removing the S4 block entirely resulted in a statistically significant performance drop (ROC-AUC 0.62 vs. 0.68 for full PRIME), while replacing S4 with conventional convolution also reduced performance (ROC-AUC 0.66), demonstrating that the superior performance of PRIME stems from the S4 layer's ability to capture temporal patterns in neural activity that conventional EEG decoding architectures cannot effectively learn.

4.3. Predictive features for cortical excitability are temporally localized

To establish an effective real-time adaptation strategy for PRIME, we compared the performance of several methods (Fig. 4a-d and Tables S2-S5). We first established a benchmark using full continual finetuning (CFT), where all model parameters are updated after each trial to allow for full adaptation to the subject's ongoing brain activity. This method achieved the highest overall predictive accuracy, significantly outperforming all other strategies ($p < 0.01$ to $p < 0.0001$, Table S3).

We then evaluated whether more targeted, computationally simpler updates could match this performance. One alternative involved updating only the final classifier layer of the model (Dec-CFT), while an even more minimal approach adapted only the decision threshold (Thr-CFT). Both of these strategies resulted in significantly lower accuracy than full CFT, demonstrating that adapting the entire network, including its deep feature-extraction layers, is critical for achieving high

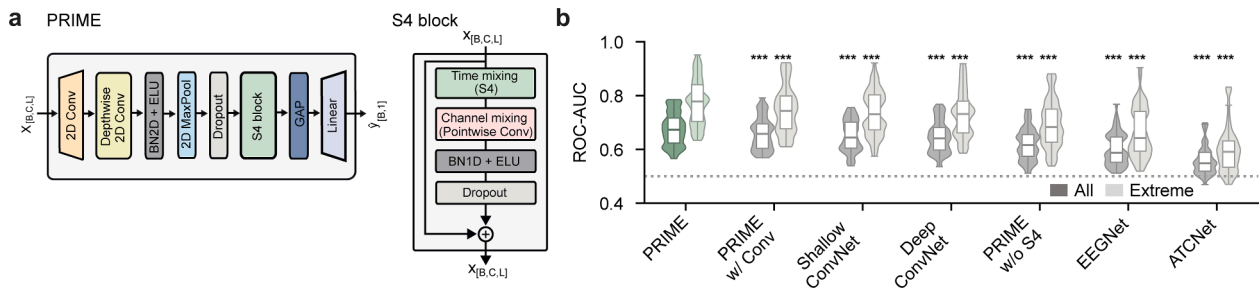


Fig. 3. The PRIME structured state-space architecture outperforms conventional EEG decoding models. a. PRIME processes input EEG data (B, batch; C, channels; L, time points) through temporal convolution followed by spatial convolution to extract spatiotemporal features. Features pass through normalization, pooling, and dropout layers before entering a structured state-space (S4) block for temporal modeling. A linear classifier with global average pooling generates the final TEP amplitude prediction. b. Performance comparison of PRIME against established EEG decoding models (EEGNet, ATCNet, Shallow ConvNet, Deep ConvNet) and ablated variants (PRIME w/o S4; PRIME w/ Conv, i.e., S4 blocks replaced with standard convolution). PRIME significantly outperforms all alternatives for both all trials (All) and extreme amplitude trials (Extreme, highest and lowest 25 %; $*** p < 0.001$, paired one-sided t -tests, Benjamini-Hochberg corrected). ROC-AUC distributions shown as violin plots across 50 datasets (approximately 955 trials each).

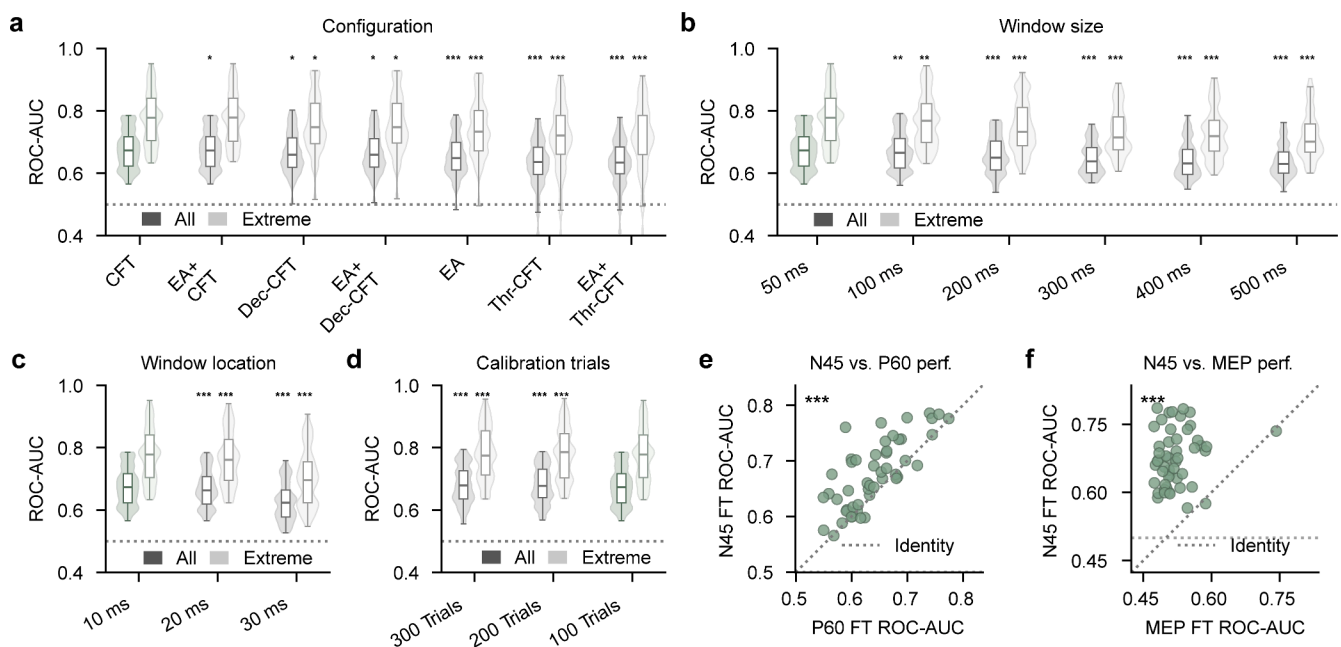


Fig. 4. Optimization of PRIME components reveals critical timing constraints and TEP-specific prediction capabilities. a. Performance comparison across PRIME configuration variants: continual finetuning (CFT), Euclidean alignment + CFT (EA+CFT), decision-level continual finetuning (Dec-CFT), Euclidean alignment with decision-level continual finetuning (EA+Dec-CFT), Euclidean alignment only (EA), and thresholded CFT variants (Thr-CFT, EA+Thr-CFT). Standard CFT significantly outperformed most variants for both all trials and extreme trials (highest and lowest 25 %; $*** p < 0.001$, $* p < 0.05$, paired one-sided t -tests, Benjamini-Hochberg corrected). b. Performance is best when using a 50 ms EEG input window. Using longer windows leads to a significant degradation in performance, particularly for extreme trials ($** p < 0.01$, $*** p < 0.001$, paired one-sided t -tests, Benjamini-Hochberg corrected). c. Placing the 50-ms window to end 10 ms before the TMS pulse onset provides significantly better performance than earlier placements ($*** p < 0.001$). d. A calibration set of 100 trials per subject is sufficient, larger sets provide only minimal further benefit ($** p < 0.01$, $*** p < 0.001$, paired one-sided t -tests, Benjamini-Hochberg corrected). e. PRIME predicts N45 amplitudes significantly more accurately than P60 amplitudes, with strong cross-subject correlation ($*** p < 0.001$, paired one-sided t -tests, Benjamini-Hochberg corrected). f. PRIME predicts N45 TEP amplitudes significantly more accurately than MEP amplitudes, with no significant correlation between prediction accuracies ($*** p < 0.001$, paired one-sided t -tests, Benjamini-Hochberg corrected).

performance.

Finally, we tested if regularization could improve performance by preventing the model from drifting too far from its pretrained state. To do this, we incorporated Euclidean Alignment (EA), a technique that aligns the statistical distribution of the incoming data with the population data it was trained on (He and Wu 2020). However, adding EA provided no benefit for predicting extreme trials and significantly lowered performance across all trials. Based on these results, we selected full continual finetuning (CFT) as the most suitable adaptation method for PRIME.

To determine the most effective pre-stimulus EEG window, we

investigated the influence of its size and temporal placement on predictive accuracy. We found that performance depends on critical temporal constraints, achieving the highest accuracy with a 50-ms window. Performance declined significantly as we increased the window size from 100 ms to 500 ms ($p < 0.001$ for all trials; $p < 0.05$ – 0.001 for extreme trials; Fig. 4b and Table S4). The placement of this window was equally important; we maximized predictive accuracy by positioning the 50-ms window to end 10 ms before the TMS pulse, as compared to earlier placements ($p < 0.001$ for all trials and extreme trials; Fig. 4c and Table S5). To test robustness to realistic timing variability, we further introduced trial-wise random jitter shifting the window earlier (away

from the TMS pulse) by up to 30 ms. Performance remained robust even at maximum jitter, with the window ending 10–40 ms pre-pulse (median AUC = 0.60–0.68 for all trials, 0.66–0.77 for extreme trials; Fig. S4 and Table S6). Finally, when we evaluated calibration requirements, we found that while using 200 or 300 trials provided a statistically significant improvement over a 100-trial baseline, the absolute accuracy gains were minimal (Fig. 4d and Table S7), suggesting that 100 trials provide sufficient subject-specific calibration for practical implementation.

Next, we assessed the ability of PRIME to generalize across different physiological readouts. We found that the model could successfully predict the later P60 TEP component, although its performance was significantly lower than for the N45 component (Fig. 4e and Table S8; paired t -test, $p < 0.001$). Importantly, we observed a strong correlation between the predictive accuracies for the two TEP components across subjects (Pearson's $r = 0.75$, $p < 0.0001$). In contrast, when we tested whether the model could generalize from cortical to corticospinal excitability (MEPs), we found that its predictive performance for MEPs was substantially lower and approached chance levels (paired t -test, $p < 0.001$). Consequently, we found no significant correlation between prediction accuracies for the N45 TEP component and MEPs across subjects (Fig. 4f and Table S8; Pearson's $r = 0.17$, $p = 0.23$).

4.4. Theta and alpha oscillations in fronto-central regions drive predictive performance

To identify the neurophysiological features underlying the predictive capabilities of PRIME, we performed occlusion-based sensitivity analysis on subject-specific finetuned models (Ieracitano et al. 2022). This approach systematically removes distinct spatial or spectral components from input EEG and measures the resulting decrease in model performance (Δ ROC-AUC), thereby mapping the features most relevant for predicting cortical excitability (Fig. 5). Our analysis revealed that the prediction success of PRIME depends on specific oscillatory patterns localized to discrete cortical regions rather than distributed global brain activity.

Spatially, the most predictive broadband EEG signals originated from a localized cluster of electrodes over the left frontal and central cortex (Fig. 5a). Occluding individual electrodes and their nearest neighbors from these regions significantly degraded model performance, whereas electrodes over the right hemisphere contributed minimally to prediction accuracy.

While all frequency bands contained predictive information, the model was most sensitive to low-frequency oscillations (Fig. 5b).

Removing theta (4–8 Hz) and alpha (8–13 Hz) bands caused the largest drops in accuracy, followed by beta (13–25 Hz) and gamma (25–47 Hz) bands, respectively.

The predictive power of these oscillations originated from similar cortical areas identified in the broadband analysis (Fig. 5c). The high importance of theta and alpha activity was predominantly localized to left fronto-central regions, with beta band activity showing a similar but weaker spatial pattern. Gamma activity offered negligible predictive value from any specific region.

5. Discussion

We introduced PRIME, a deep learning framework designed to predict cortical excitability from real-time EEG data using TMS-evoked potential (TEP) amplitudes as a measure of cortical excitability. Unlike current approaches that rely on static biomarkers, PRIME adapts in real-time to dynamic and individual-specific neural states. This approach addresses critical shortcomings of current EEG-guided TMS: it automatically learns personalized biomarkers instead of using fixed features, continuously adapts to evolving brain states, and targets cortical rather than corticospinal excitability. This TEP-based approach provides a methodological foundation that could enable closed-loop stimulation across diverse cortical targets. However, our current validation is limited to single-pulse stimulation of M1 in healthy adults, and extension to other brain regions, repetitive stimulation protocols, and clinical populations requires systematic validation.

PRIME achieved strong predictive performance with a median ROC-AUC of 0.68 across all TEP amplitudes and 0.77 when distinguishing extreme excitability states (highest and lowest 25 % of trials). This performance was achieved through the general-to-individual training strategy of PRIME, involving population-level pretraining, subject-specific calibration, and continual online adaptation—which was essential for navigating significant inter-subject variability (Ziemann and Siebner 2015; Ridging and Ziemann 2010; Haxel et al. 2025; Ahola, et al. 2025). By integrating structured state-space (S4) modeling, PRIME captured EEG dynamics relevant to cortical responsiveness, proving superior to established EEG-based decoding networks (Fig. 2b and Table 4). (Pezzopane et al. 2023; Vetter et al. 2023).

PRIME predicts the brain's immediate electrophysiological response to a single pulse of TMS as an instantaneous read-out of cortical excitability. Instantaneous response amplitude has long served as a surrogate for membrane excitability (Rossini et al. 1994; Ziemann and Siebner 2015). Crucially, a more excitable, or receptive, state is not only

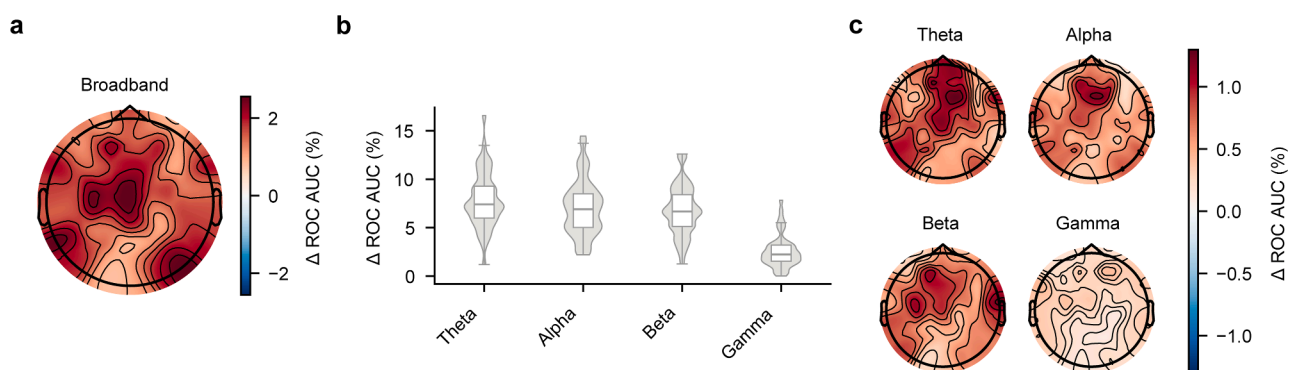


Fig. 5. Theta and alpha oscillations in fronto-central regions provide the primary predictive features for cortical excitability. a. Spatial distribution of electrode importance determined by occlusion analysis. The magnitude of the decrease in prediction accuracy (positive Δ ROC-AUC values, %) indicates the importance of each electrode and their 8-nearest-neighbors for broadband EEG-based prediction. b. Contribution of the theta, alpha, beta, and gamma frequency bands to prediction performance, determined by occluding each band across all electrodes. Positive Δ ROC-AUC (%) indicates the importance of each frequency band for accurate prediction. c. Spatial distribution of predictive importance for each frequency band, obtained by selectively occluding frequency bands at individual electrodes and their 8-nearest-neighbors. Positive Δ ROC-AUC (%) indicates the importance of specific frequency bands at different cortical locations. All analyses used subject-specific finetuned PRIME models, with statistical significance determined using Wilcoxon signed-rank tests with FDR correction across test subjects (all displayed results significant at $p < 0.05$).

expected to produce a stronger immediate response but is also considered important for inducing therapeutic long-term changes in synaptic strength (i.e., plasticity) (Silvanto and Pascual-Leone 2008; Ziemann and Siebner 2015). Repetitive stimulation (rTMS) is thought to induce these plastic changes through molecular cascades similar to long-term potentiation (LTP) (Hoogendam et al. 2010). Cortical excitability itself fluctuates rhythmically, most prominently in theta (4–8 Hz) and alpha (8–13 Hz) bands. These rhythms reflect cyclical changes in neuronal excitability. During a high-excitability phase, neurons are closer to their firing threshold and thus more susceptible to external stimulation, creating conditions favorable for inducing LTP (Buzsáki 2006). Indeed, delivering TMS at EEG-derived high-excitability phases in humans enhances both the immediate response amplitude and the longer-term, LTP-like after-effects of stimulation (Zrenner et al. 2018; Bergmann et al. 2019; Baur et al. 2020).

PRIME leverages this state-dependence by predicting effective stimulation moments from real-time EEG data. Rather than delivering stimulation randomly with respect to brain state, PRIME ensures that more pulses arrive when neurons are neurophysiologically primed for plasticity induction. This transforms TMS from a probabilistic intervention into a more deterministic therapeutic tool. Potential implications become evident when considering practical scenarios. In a typical 800-trial session, PRIME correctly predicts cortical excitability in 544 trials overall (68 % accuracy) and 308 out of 400 extreme excitability trials (77 % accuracy), compared to 400 and 200 expected by chance, respectively. This enrichment represents a 36 % improvement in overall accuracy (54 % for extreme states), which, when accumulated across hundreds of pulses, could plausibly translate into more efficient stimulation protocols. Importantly, prior EEG-phase-targeted TMS studies have shown that even modest increases in the proportion of high-excitability stimulations yield measurable gains in both immediate response amplitudes and LTP- or LTD-like aftereffects (Zrenner et al. 2018; Bergmann et al. 2019; Baur et al. 2020). PRIME extends this principle by inferring individualized excitability states rather than relying on fixed oscillatory phases. However, whether predictive enrichment of the magnitude reported here translates into therapeutic efficacy remains an open question that requires direct prospective validation in closed-loop trials.

The neurophysiological signatures underlying the predictive success of PRIME provide further insight into potential cortical excitability mechanisms. Theta (4–8 Hz) and alpha (8–13 Hz) oscillations emerged as the most predictive features (Fig. 4b–c), consistent with their established roles in motor cortical excitability and sensorimotor integration. While our EEG-based analysis cannot establish causal mechanisms, prior work shows that oscillations shape cyclical fluctuations in neuronal membrane potential, with certain phases corresponding to periods of heightened or reduced excitability (Buzsáki, 2006). Experimental studies using phase-targeted TMS support this principle. During wakefulness, stimulation delivered at specific phases of the mu-alpha rhythm has been shown to facilitate corticospinal responses more strongly than stimulation at the opposite phases, an effect described as asymmetric “pulsed facilitation” (Bergmann et al., 2019). In contrast, spindle oscillations during non-REM sleep transiently suppress corticospinal responsiveness, producing an asymmetric “pulsed inhibition” profile (Hassan et al., 2025). Together, these findings suggest that oscillatory rhythms can act as temporal gates, either opening or closing windows for effective stimulation depending on frequency band and phase. Within this broader framework, theta oscillations are thought to coordinate excitatory-inhibitory cycles that support selective information exchange through phase-coupling (Fries 2015), as demonstrated in both cognitive (Gevins 1997; Onton et al. 2005; Klimesch et al. 2007) and sensorimotor tasks (Cruikshank et al. 2012; Watanabe et al. 2021). Alpha oscillations, by contrast, have been consistently linked to voluntary motor control, including imagery, preparation, execution, and somatosensory performance (Pfurtscheller and Aranibar 1979; Pfurtscheller and Lopes Da Silva 1999; Pfurtscheller et al. 2006; Jones

et al. 2010; Baumgarten et al. 2015). In our study, predictive information was primarily concentrated in fronto-central regions (channels Cz, CPz, CP3, C3, F1), aligning with previous research (Ahola et al. 2025) and potentially reflecting integrative mechanisms linking sensorimotor coordination with cognitive states including attention, alertness, and emotional processes (Noreika et al. 2020; Chen and Huang 2018; Conte et al. 2007; Baumert et al. 2011). The temporal specificity of predictive features—with optimal performance achieved using a 50-ms window ending 10 ms before stimulation (Fig. 3bc and Table 6,7)—indicates that cortical excitability states are determined by very recent neural activity patterns. Supplementary analyses show that PRIME’s prediction accuracy declines gradually rather than abruptly when windows are shifted further from the pulse. For example, windows positioned 20–50 ms pre-pulse still achieve performance well above chance (Fig. 4c and Table S5). In addition to systematic window shifts, PRIME retained robust accuracy under random timing jitter up to 30 ms away from the TMS pulse (Fig. S4 and Table S6). Thus, PRIME does not require millisecond-precise alignment: while performance peaks near stimulation, earlier windows remain predictive and, by providing extra lead time, relax real-time latency demands. This robustness suggests that modest jitter or delay in clinical settings is unlikely to diminish enrichment over random stimulation.

A critical finding of our study is the stark contrast between the ability of PRIME to predict TEP amplitudes versus MEP responses. While achieving robust performance for TEP prediction, the model’s accuracy for MEPs approached chance levels (0.52 ROC-AUC), with no significant correlation between TEP and MEP prediction performance across subjects. This difference suggests fundamental differences between cortical and corticospinal excitability mechanisms. MEP amplitudes likely reflect the integrated output of multiple neural processing stages: cortical excitability, corticospinal tract integrity, spinal motor neuron pool excitability, and neuromuscular transmission (Goetz et al. 2014, 2022). In contrast, TEPs provide a direct measure of cortical responsiveness by capturing post-synaptic potentials evoked by TMS without spinal confounds (Darmani and Ziemann 2019). The success of PRIME in predicting both N45 and P60 TEP components, with similar spatial and frequency patterns of predictive features (Fig. S2), would be consistent with cortical excitability mechanisms underlying both responses. The slightly lower prediction performance for P60 compared to N45 (0.64 vs. 0.68 ROC-AUC for all trials; 0.72 vs. 0.77 for extreme trials) potentially reflects increased contamination from sensory activity at longer latencies (Gordon et al. 2021) rather than fundamentally different predictive mechanisms. This finding suggests that PRIME is successfully isolating and decoding specific cortical dynamics that are upstream of, and distinct from, factors that modulate spinal and muscular excitability. The clear dissociation indicates that the EEG features PRIME uses capture instantaneous cortical states rather than non-specific markers of global arousal or muscle artifacts.

From a practical implementation perspective, PRIME meets stringent latency requirements for real-time TMS applications on both GPU and CPU hardware, making the approach feasible for standard clinical computing environments. The framework processes EEG data in 7.2 ms on GPU or 7.7 ms on CPU—both well below the 10 ms constraint required to maintain predictive performance without interfering with stimulation timing. While GPU hardware accelerates model updates (8.3 ms vs 35.3 ms), both platforms complete adaptation following TEP extraction (82.2 ms) well within the typical 2.5-second inter-stimulus interval, enabling continuous model refinement without disrupting established TMS protocols.

Several challenges must be addressed before clinical deployment. First, our evaluation used retrospective data to simulate real-time prediction performance rather than prospective closed-loop control. While we demonstrate computational feasibility within realistic latency constraints, true prospective validation, where PRIME actively controls stimulation timing, is essential to confirm that predicted high-excitability states translate to enhanced therapeutic outcomes in

clinical practice. Second, PRIME was developed and validated exclusively using single-pulse TMS applied to M1 in healthy young adults. While TEPs can theoretically be recorded from any cortical target, our predictive framework has not been tested in association cortices, where neural dynamics and optimal predictive features may differ substantially from motor cortex. The generalizability of theta and alpha oscillations as predictive biomarkers across different cortical regions remains an open empirical question. Furthermore, translation to rTMS protocols, where the clinical utility of PRIME would be greatest, introduces additional complexities. During rTMS, ongoing stimulation itself modulates brain state, creating a dynamic interaction between prediction and intervention that our current offline validation cannot capture. Third, the performance of PRIME must be validated in real-world clinical environments where EEG data are often contaminated by movement artifacts, muscle activity (EMG), and fluctuations in patient vigilance and alertness (van Diessen et al. 2015). The ability to generalize across heterogeneous patient populations varying in age, medication status, and comorbidity needs to be established, as the neurophysiological signatures of excitability may differ significantly in older adults or patients with psychiatric comorbidities. Fourth, while grounded in evidence (Zrenner et al. 2018; Bergmann et al. 2019; Baur et al. 2020), the assumption that consistently stimulating during high- or low-excitability states leads to greater therapeutic LTP or LTD requires direct clinical confirmation. The relationship between excitability and plasticity is likely non-linear, with a potential optimal window rather than a simple “more is better” function (Karabanov et al. 2015). Future clinical trials are needed to measure both symptom improvement and neurophysiological data confirming the engagement and enhancement of LTP-like plasticity. A further limitation concerns the reliability of online preprocessing. Although our pipeline combines established community-validated methods—ICA with ICLabel for ocular artifacts (Pion-Tonachini et al., 2019), SOUND for denoising (Mutanen et al., 2018), and SSP-SIR for suppressing TMS-induced decay and muscle artifacts (Mutanen et al., 2016)—and meets stringent real-time latency requirements, no method can fully eliminate TMS-related contamination. To mitigate this, all preprocessing parameters are estimated once during a calibration stage and then applied identically during online operation, ensuring consistent processing across trials. We also focused on later TEP components (N45, P60), which are less affected by early-latency artifacts even after correction (Ahola et al., 2025). While residual noise may still occur, our framework handles such events conservatively by excluding corrupted trials from model adaptation rather than allowing them to degrade predictive stability. Future work should directly quantify preprocessing reliability across different hardware setups and clinical environments and develop automated contingency strategies to further increase robustness in routine practice. Finally, while PRIME outperformed established EEG decoding architectures and ablated variants, we did not perform an exhaustive search over alternative architectural designs or hyperparameter configurations. Our choices were guided by real-time feasibility and prior work in EEG decoding, but additional exploration of model depth, hidden dimensions, and alternative temporal modeling strategies may further improve performance.

PRIME establishes a computational framework for personalized neuromodulation by shifting from static biomarkers to dynamically inferred EEG-derived brain states. Its ability to predict cortical excitability with 36–54 % improvement over chance levels provides a foundation for more precise, individualized closed-loop TMS interventions. Following validation in association cortices, clinical populations, and repetitive stimulation protocols, future clinical applications could potentially leverage the continuous, graded predictions of PRIME to reduce treatment session requirements and enhance therapeutic effectiveness across neurological and psychiatric disorders. The next critical steps involve validating these promising results to translate this technical proof-of-concept into measurable clinical benefits. The real-time adaptability of our framework opens possibilities

for advanced closed-loop systems that could dynamically modulate both stimulation timing and intensity to achieve optimal biological effects across fluctuating brain states.

Data and code availability statement

Data are available upon reasonable request from the corresponding authors. Compliance with EU data protection law (Regulation (EU) 2016/679) is required. An “End User Licence Agreement” (EULA) will be required before data will be released. Data will be shared on a project specific basis.

Code for our preprocessing and decoding pipelines are available on Github (<https://github.com/mackelab/PRIME>).

CRediT authorship contribution statement

Lisa Haxel: Writing – review & editing, Writing – original draft, Visualization, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Oskari Ahola:** Methodology, Software, Data curation, Writing – review & editing. **Jaivardhan Kapoor:** Methodology, Writing – review & editing. **Ulf Ziemann:** Conceptualization, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Jakob H. Macke:** Conceptualization, Writing – review & editing, Supervision, Project administration, Funding acquisition.

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.

Acknowledgments

This work was supported by the Else Kröner Medical Scientists Kolleg Clinbrain: Artificial Intelligence for Clinical Brain Research, the Machine Learning Cluster of Excellence, EXC number 2064/1–390727645, the German Federal Ministry of Education and Research (BMBF): Tübingen AI Center, FKZ: 01IS18039A and the European Research Council (ERC Synergy) under the European Union’s Horizon 2020 research and innovation program (ConnectToBrain, grant number 810377)

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.neuroimage.2025.121547](https://doi.org/10.1016/j.neuroimage.2025.121547).

References

- Ahola, O., Haxel, L., Ermolova, M., Humaidan, D., Mutanen, T.P., Laine, M., Makkonen, M., Ukharova, E., Roine, T., Lioumis, P., Guidotti, R., Ilmoniemi, R.J., Ziemann, U., 2025. Predictive modeling of TMS-evoked responses: unraveling instantaneous excitability states. *bioRxiv*. <https://doi.org/10.1101/2025.06.18.660274>.
- Altaheri, H., Muhammad, G., Alsulaiman, M., 2023. Physics-informed attention temporal convolutional network for Eeg-based motor imagery classification. *IEEe Trans. Industr. Inform.* 19 (2), 2249–2258. <https://doi.org/10.1109/TII.2022.3197419>.
- Bai, Z., Zhang, J., Fong, K.N.K., 2022. Effects of transcranial magnetic stimulation in modulating cortical excitability in patients with stroke: a systematic review and meta-analysis. *J. Neuroeng. Rehabil.* 19 (1), 24. <https://doi.org/10.1186/s12984-022-00999-4>.
- Baumert, A., Sinclair, C., MacLeod, C., Hammond, G., 2011. Negative emotional processing induced by spoken scenarios modulates corticospinal excitability. *Cognitive, Affect., Behav. Neurosc.* 11 (3), 404–412. <https://doi.org/10.3758/s13415-011-0044-z>.
- Baumgarten, T.J., Schnitzler, A., Lange, J., 2015. Beta oscillations define discrete perceptual cycles in the somatosensory domain. *Proceed. National Acad. Sci.* 112 (39), 12187–12192. <https://doi.org/10.1073/pnas.1501438112>.
- Baur, D., Galevska, D., Hussain, S., Cohen, L.G., Ziemann, U., Zrenner, C., 2020. Induction of LTD-like corticospinal plasticity by low-frequency rTMS depends on

- pre-stimulus phase of sensorimotor μ -rhythm. *Brain Stimul.* 13 (6), 1580–1587. <https://doi.org/10.1016/j.brs.2020.09.005>.
- Bell, A.J., Sejnowski, T.J., 1995. An information-maximization approach to blind separation and blind deconvolution. *Neural Comput.* 7 (6), 1129–1159. <https://doi.org/10.1162/neco.1995.7.6.1129>.
- Bembek, J.P., Kurczyk, K., Klysz, B., Cudna, A., Antczak, J., Członkowska, A., 2020. Prediction of recovery and outcome using motor evoked potentials and brain derived neurotrophic factor in subacute stroke. *J. Stroke Cerebrovasc. Dis.* 29 (11), 105202. <https://doi.org/10.1016/j.jstrokecerebrovasdis.2020.105202>.
- Bergmann, T.O., 2018. Brain state-dependent brain stimulation. *Front. Psychol.* 9. <https://doi.org/10.3389/fpsyg.2018.02108>.
- Bergmann, T.O., Lieb, A., Zrenner, C., Ziemann, U., 2019. Pulsed facilitation of corticospinal excitability by the sensorimotor μ -alpha rhythm. *J. Neurosci.* 39 (50), 10034–10043. <https://doi.org/10.1523/JNEUROSCI.1730-19.2019>.
- Buzsáki, G., 2006. *Rhythms of the Brain*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195301069.001.0001>.
- Chen, K.-H., Huang, Y.-Z., 2018. The change of motor cortical excitability between eyes open and closed conditions. *Neuroreport* 29 (3), 214–218. <https://doi.org/10.1097/WNR.0000000000000955>.
- Conte, A., Gilio, F., Iezzi, E., Frasca, V., Inghilleri, M., Berardelli, A., 2007. Attention influences the excitability of cortical motor areas in healthy humans. *Exp. Brain Res.* 182 (1), 109–117. <https://doi.org/10.1007/s00221-007-0975-3>.
- Cruikshank, L.C., Singhal, A., Hueppelsheuser, M., Caplan, J.B., 2012. Theta oscillations reflect a putative neural mechanism for human sensorimotor integration. *J. Neurophysiol.* 107 (1), 65–77. <https://doi.org/10.1152/jn.00893.2010>.
- Darmani, G., Ziemann, U., 2019. Pharmacophysiology of TMS-evoked EEG potentials: a mini-review. *Brain Stimul.* 12 (3), 829–831. <https://doi.org/10.1016/j.brs.2019.02.021>.
- Ermolova, M., Metsomaa, J., Belardinelli, P., Zrenner, C., Ziemann, U., 2024. Blindly separated spontaneous network-level oscillations predict corticospinal excitability. *J. Neural Eng.* 21 (3), 036041. <https://doi.org/10.1088/1741-2552/ad5404>.
- Fischl, B., Sereno, M.I., Tootell, R.B.H., Dale, A.M., 1999. High-resolution intersubject averaging and a coordinate system for the cortical surface. *Hum. Brain Mapp.* 8 (4), 272–284. [https://doi.org/10.1002/\(SICI\)1097-0193\(1999\)8:4<272::AID-BM10>3.0.CO;2-4](https://doi.org/10.1002/(SICI)1097-0193(1999)8:4<272::AID-BM10>3.0.CO;2-4).
- Fries, P., 2015. Rhythms for Cognition: communication through Coherence. *Neuron* 88 (1), 220–235. <https://doi.org/10.1016/j.neuron.2015.09.034>.
- Gevins, A., 1997. High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice. *Cerebral Cortex* 7 (4), 374–385. <https://doi.org/10.1093/cecror/7.4.374>.
- Gordon, P.C., Jovellar, D.B., Song, Y., Zrenner, C., Belardinelli, P., Siebner, H.R., Ziemann, U., 2021. Recording brain responses to TMS of primary motor cortex by EEG - utility of an optimized sham procedure. *Neuroimage* 245, 118708. <https://doi.org/10.1016/j.neuroimage.2021.118708>.
- Goetz, S.M., Luber, B., Lisanby, S.H., Peterchev, A.V., 2014. A novel model incorporating two variability sources for describing motor evoked potentials. *Brain Stimul.* 7 (4), 541–552. <https://doi.org/10.1016/j.brs.2014.02.002>.
- Goetz, S.M., Howell, B., Wang, B., Li, Z., Sommer, M.A., Peterchev, A.V., Grill, W.M., 2022. Isolating two sources of variability of subcortical stimulation to quantify fluctuations of corticospinal tract excitability. *Clin. Neurophysiol.* 138, 134–142. <https://doi.org/10.1016/j.clinph.2022.02.009>.
- Gu, A., Goel, K., and Ré, C. Efficiently Modeling Long Sequences with Structured State Spaces. 2022. <https://doi.org/10.48550/arXiv.2111.00396>.
- Hassan, U., Okyere, P., Masouleh, M.A., Zrenner, C., Ziemann, U., Bergmann, T.O., 2025. Pulsed inhibition of corticospinal excitability by the thalamocortical sleep spindle. *Brain Stimul.* 18 (2), 265–275. <https://doi.org/10.1016/j.brs.2024.12.003>.
- Haxel, L., Ahola, O., Belardinelli, P., Ermolova, M., Humaidan, D., Macke, J.H., Ziemann, U., 2025a. Decoding Motor Excitability in TMS Using EEG-Features: an Exploratory Machine Learning Approach. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 33, 103–112. <https://doi.org/10.1109/TNSRE.2024.3516393>.
- Haxel, L., Kapoor, J., Ziemann, U., and Macke, J.H. EDAPT-towards calibration-free BCIs with continual online adaptation. *arXiv preprint arXiv:2508.10474* (2025). <https://doi.org/10.48550/arXiv.2508.10474>.
- He, H., Wu, D., 2020. Transfer learning for brain-computer interfaces: a euclidean space data alignment approach. *IEEE Trans. Biomed. Eng.* 67 (2), 399–410. <https://doi.org/10.1109/TBME.2019.2913914>.
- Henderson, C.J., Butler, S.R., Glass, A., 1975. The localization of equivalent dipoles of EEG sources by the application of electrical field theory. *Electroencephalogr. Clin. Neurophysiol.* 39 (2), 117–130. [https://doi.org/10.1016/0013-4694\(75\)90002-4](https://doi.org/10.1016/0013-4694(75)90002-4).
- Hoogendam, J.M., Ramakers, G.M.J., Di Lazzaro, V., 2010. Physiology of repetitive transcranial magnetic stimulation of the human brain. *Brain Stimul.* 3 (2), 95–118. <https://doi.org/10.1016/j.brs.2009.10.005>.
- Humaidan, D., Xu, J., Kirchoff, M., Romani, G.L., Ilmoniemi, R.J., Ziemann, U., 2024. Towards real-time EEG-TMS modulation of brain state in a closed-loop approach. *Clinical Neurophysiology* 158, 212–217. <https://doi.org/10.1016/j.clinph.2023.12.006>.
- Humaidan, D., Xu, J., Zrenner, C., Marzetti, L., Belardinelli, P., Ilmoniemi, R., Romani, G. L., Ziemann, U., 2025. Personalized Closed-Loop rTMS Using Reinforcement Learning to Optimize Phase-Targeted Stimulation of the SMA-M1 Network. *en. Brain Stimulation* 18 (1), 543. <https://doi.org/10.1016/j.brs.2024.12.956>.
- Hussain, S.J., Quentin, R., 2022. Decoding personalized motor cortical excitability states from human electroencephalography. *Sci. Rep.* 12 (1), 6323. <https://doi.org/10.1038/s41598-022-10239-3>.
- Ieracitano, C., Mammone, N., Hussain, A., Morabito, F.C., 2022. A novel explainable machine learning approach for eeg-based brain-computer interface systems. *Neural Computing and Applications* 34 (14), 11347–11360.
- John, Y.J., Sawyer, K.S., Srinivasan, K., Müller, E.J., Munn, B.R., Shine, J.M., 2022. It's about time: linking dynamical systems with human neuroimaging to understand the brain. *Network Neurosci.* 6 (4), 960–979. https://doi.org/10.1162/netn_a_00230.
- Jones, S.R., Kerr, C.E., Wan, Q., Pritchett, D.L., Hämmäläinen, M., Moore, C.L., 2010. Cued spatial attention drives functionally relevant modulation of the mu rhythm in primary somatosensory cortex. *J. Neurosci.* 30 (41), 13760–13765. <https://doi.org/10.1523/JNEUROSCI.2969-10.2010>.
- Kapoor, J., Schulz, A., Vetter, J., Pei, F., Gao, R., Macke, J.H., 2024. Latent diffusion for neural spiking data. *Adv. Neural Inf. Process. Syst.* 37, 118119–118154.
- Karabanov, A., Ziemann, U., Hamada, M., George, M.S., Quartarone, A., Classen, J., Massimini, M., Rothwell, J., Siebner, H.R., 2015. Consensus paper: probing homeostatic plasticity of human cortex with non-invasive transcranial brain stimulation. *Brain Stimul.* 8 (3), 442–454. <https://doi.org/10.1016/j.brs.2015.01.404>.
- Khatry, U.U., Pulliani, K., Manesiya, M., Cortez, M.V., Millán, J.d.R., Hussain, S.J., 2025. Personalized whole-brain activity patterns predict human corticospinal tract activation in real-time. *Brain Stimul.* 18 (1), 64–76. <https://doi.org/10.1016/j.brs.2024.12.1193>.
- Klimesch, W., Sauseng, P., Hanslmayr, S., 2007. EEG alpha oscillations: the inhibition-timing hypothesis. *Brain Res. Rev.* 53 (1), 63–88. <https://doi.org/10.1016/j.brainresrev.2006.06.003>.
- Kricheldorff, J., Göke, K., Kiebs, M., Kasten, F.H., Herrmann, C.S., Witt, K., Hurlmann, R., 2022. Evidence of Neuroplastic Changes after Transcranial Magnetic, Electric, and Deep Brain Stimulation. *Brain Sci.* 12 (7), 929. <https://doi.org/10.3390/brainsci12070929>.
- Lawhern, V.J., Solon, A.J., Waytowich, N.R., Gordon, S.M., Hung, C.P., Lance, B.J., 2018. EEGNet: a compact convolutional neural network for EEG-based brain-computer interfaces. *J. Neural Eng.* 15 (5), 056013. <https://doi.org/10.1088/1741-2552/ace8c>.
- Lefaucher, J.-P., Aleman, A., Baeken, C., Benninger, D.H., Brunelin, J., Di Lazzaro, V., Filipović, S.R., Grefkes, C., Hasan, A., Hummel, F.C., Jääskeläinen, S.K., Langguth, B., Leocani, L., Londero, A., Nardone, R., Nguyen, J.-P., Nyffeler, T., Oliveira-Maia, A.J., Oliviero, A., Padberg, F., Palm, U., Paulus, W., Poulet, E., Quartarone, A., Rachid, F., Rektorová, I., Rossi, S., Sahlsten, H., Schecklmann, M., Szekely, D., Ziemann, U., 2020. Evidence-based guidelines on the therapeutic use of repetitive transcranial magnetic stimulation (rTMS): an update (2014–2018). *Clin. Neurophysiol.* 131 (2), 474–528. <https://doi.org/10.1016/j.clinph.2019.11.002>.
- Lee, T.-W., Girolami, M., Sejnowski, T.J., 1999. Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. *Neural Comput.* 11 (2), 417–441. <https://doi.org/10.1162/08997669900016719>.
- Metsomaa, J., Belardinelli, P., Ermolova, M., Ziemann, U., Zrenner, C., 2021. Causal decoding of individual cortical excitability states. *en. NeuroImage* 245, 118652. <https://doi.org/10.1016/j.neuroimage.2021.118652>.
- Noreika, V., Kamke, M.R., Canales-Johnson, A., Chennu, S., Bekinschtein, T.A., Mattingley, J.B., 2020. Alertness fluctuations when performing a task modulate cortical evoked responses to transcranial magnetic stimulation. *Neuroimage* 223, 117305. <https://doi.org/10.1016/j.neuroimage.2020.117305>.
- Onton, J., Delorme, A., Makeig, S., 2005. Frontal midline EEG dynamics during working memory. *Neuroimage* 27 (2), 341–356. <https://doi.org/10.1016/j.neuroimage.2005.04.014>.
- Mutanen, T.P., Kukkonen, M., Nieminen, J.O., Stenroos, M., Sarvas, J., Ilmoniemi, R.J., 2016. Recovering TMS-evoked EEG responses masked by muscle artifacts. *Neuroimage* 139, 157–166. <https://doi.org/10.1016/j.neuroimage.2016.05.028>.
- Mutanen, T.P., Metsomaa, J., Liljander, S., Ilmoniemi, R.J., 2018. Automatic and robust noise suppression in EEG and MEG: the SOUND algorithm. *Neuroimage* 166, 135–151. <https://doi.org/10.1016/j.neuroimage.2017.10.021>.
- Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., et al., 2019. PyTorch: an imperative style, high-performance deep learning library. *Adv. Neural Inf. Process. Syst.* 8024–8035. <https://arxiv.org/abs/1912.01703>.
- Pellicciari, M.C., Bonni, S., Ponzio, V., Cinnera, A.M., Mancini, M., Casula, E.P., Sallustio, F., Paolucci, S., Caltagirone, C., Koch, G., 2018. Dynamic reorganization of TMS-evoked activity in subcortical stroke patients. *NeuroImage* 175, 365–378. <https://doi.org/10.1016/j.neuroimage.2018.04.011>.
- Pezzopane, V., D'Acunzio, A., Casula, E.P., Spampinato, D., Koch, G., Fadiga, L., 2023. Neuromodulation using cross-frequency coupling ACS: preliminary data. *Brain Stimul.* 16 (1), 329. <https://doi.org/10.1016/j.brs.2023.01.616>.
- Pfurtscheller, G., Aranibar, A., 1979. Evaluation of event-related desynchronization (ERD) preceding and following voluntary self-paced movement. *Electroencephalogr. Clin. Neurophysiol.* 46 (2), 138–146. [https://doi.org/10.1016/0013-4694\(79\)90063-4](https://doi.org/10.1016/0013-4694(79)90063-4).
- Pfurtscheller, G., Lopes Da Silva, F.H., 1999. Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clin. Neurophysiol.* 110 (11), 1842–1857. [https://doi.org/10.1016/S1388-2457\(99\)00141-8](https://doi.org/10.1016/S1388-2457(99)00141-8).
- Pfurtscheller, G., Brunner, C., Schlögl, A., Lopes da Silva, F.H., 2006. Mu rhythm (de) synchronization and EEG single-trial classification of different motor imagery tasks. *Neuroimage* 31 (1), 153–159. <https://doi.org/10.1016/j.neuroimage.2005.12.003>.
- Pion-Tonachini, L., Kreutz-Delgado, K., Makeig, S., 2019. ICLabel: an automated electroencephalographic independent component classifier, dataset, and website. *Neuroimage* 198, 181–197. <https://doi.org/10.1016/j.neuroimage.2019.05.026>.
- Ridding, M.C., Ziemann, U., 2010. Determinants of the induction of cortical plasticity by non-invasive brain stimulation in healthy subjects. *J. Physiol. (Lond.)* 588 (Pt 13), 2291–2304. <https://doi.org/10.1113/jphysiol.2010.190314>.

- Rossini, P.M., Barker, A.T., Berardelli, A., Caramia, M.D., Caruso, G., Cracco, R.Q., et al., 1994. Non-invasive electrical and magnetic stimulation of the brain, spinal cord and roots: basic principles and procedures for routine clinical application. Report of an IFCN committee. *Electroencephalogr. Clin. Neurophysiol.* 91 (2), 79–92. [https://doi.org/10.1016/0013-4694\(94\)90029-9](https://doi.org/10.1016/0013-4694(94)90029-9).
- Sarvas, J., 1987. Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. *Phys. Med. Biol.* 32 (1), 11–22. <https://doi.org/10.1088/0031-9155/32/1/004>.
- Schirrmeyer, R.T., Springenberg, J.T., Fiederer, L.D.J., Glasstetter, M., Eggenberger, K., Tangermann, M., Hutter, F., Burgard, W., Ball, T., 2017. Deep learning with convolutional neural networks for EEG decoding and visualization. *Hum. Brain Mapp.* 38 (11), 5391–5420. <https://doi.org/10.1002/hbm.23730>.
- Strafella, R., Momi, D., Zomorodi, R., Lissemore, J., Noda, Y., Chen, R., Rajji, T.K., Griffiths, J.D., Vila-Rodriguez, F., Downar, J., Daskalakis, Z.J., Blumberger, D.M., Voineskos, D., 2023. Identifying neurophysiological markers of intermittent theta burst stimulation in treatment-resistant depression using transcranial magnetic stimulation-electroencephalography. *Biol. Psychiatry* 94 (6), 454–465. <https://doi.org/10.1016/j.biopsych.2023.04.011>.
- Silvanto, J., Pascual-Leone, A., 2008. State-dependency of transcranial magnetic stimulation. *Brain Topogr.* 21 (1), 1–10. <https://doi.org/10.1007/s10548-008-0067-0>.
- van Diessen, E., Numan, T., van Dellen, E., van der Kooij, A.W., Boersma, M., Hofman, D., et al., 2015. Opportunities and methodological challenges in EEG and MEG resting state functional brain network research. *Clin. Neurophysiol.* 126 (8), 1468–1481. <https://doi.org/10.1016/j.clinph.2014.11.018>.
- Vetter, D.E., Zrenner, C., Belardinelli, P., Mutanen, T.P., Kozák, G., Marzetti, L., Ziemann, U., 2023. Targeting motor cortex high-excitability states defined by functional connectivity with real-time EEG–TMS. *Neuroimage* 284, 120427. <https://doi.org/10.1016/j.neuroimage.2023.120427>.
- Vida, R.G., Sághy, E., Bella, R., Kovács, S., Erdősi, D., Józwiak-Hagymásy, J., Zemplényi, A., Tényi, T., Osváth, P., Voros, V., 2023. Efficacy of repetitive transcranial magnetic stimulation (rTMS) adjunctive therapy for major depressive disorder (MDD) after two antidepressant treatment failures: meta-analysis of randomized sham-controlled trials. *BMC. Psychiatry* 23 (1), 545. <https://doi.org/10.1186/s12888-023-05033-y>.
- Voineskos, D., Blumberger, D.M., Zomorodi, R., Rogasch, N.C., Farzan, F., Foussias, G., Rajji, T.K., Daskalakis, Z.J., 2019. Altered transcranial magnetic stimulation-electroencephalographic markers of inhibition and excitation in the dorsolateral prefrontal cortex in major depressive disorder. *Biol. Psychiatry* 85 (6), 477–486. <https://doi.org/10.1016/j.biopsych.2018.09.032>.
- Watanabe, T., Mima, T., Shibata, S., Kirimoto, H., 2021. Midfrontal theta as moderator between beta oscillations and precision control. *Neuroimage* 235, 118022. <https://doi.org/10.1016/j.neuroimage.2021.118022>.
- Ziemann, U., Siebner, H.R., 2015. Inter-subject and inter-session variability of plasticity induction by non-invasive brain stimulation: boon or bane? *Brain Stimul.* 8 (3), 662–663. <https://doi.org/10.1016/j.brs.2015.01.409>.
- Zrenner, C., Desideri, D., Belardinelli, P., Ziemann, U., 2018. Real-time EEG-defined excitability states determine efficacy of TMS-induced plasticity in human motor cortex. *Brain Stimul.* 11 (2), 374–389. <https://doi.org/10.1016/j.brs.2017.11.016>.