Machine learning predicts neural rhythms from brainwide hemodynamics across vigilance states

Leandro P. L. Jacob (ljacob@mit.edu)
Massachusetts Institute of Technology, Cambridge, MA

Sydney M. Bailes (smbailes@mit.edu)
Massachusetts Institute of Technology, Cambridge, MA
Boston University, Boston, MA

Stephanie D. Williams (sdwilli@mit.edu)
Massachusetts Institute of Technology, Cambridge, MA
Boston University, Boston, MA

Carsen Stringer (stringerc@janelia.hhmi.org)
HHMI Janelia Research Campus, Ashburn, VA

Laura D. Lewis (ldlewis@mit.edu)
Massachusetts Institute of Technology, Cambridge, MA
Massachusetts General Hospital, Boston, MA

Abstract:
Neurons often fire in synchrony, generating rhythms that support cognition and signal distinct brain states. These rhythms have been widely studied with EEG, but EEG's low spatial resolution limits our ability to investigate the brainwide activity that underlies neural rhythms. fMRI can measure brainwide activity through hemodynamic signals, but identifying relationships between hemodynamics and electrophysiology is analytically challenging, particularly when trial averaging is not possible—such as in studies of spontaneous, naturally varying brain states. We developed a machine learning approach that predicts neural rhythms (EEG power in canonical frequency bands) from fast fMRI (<400 ms TR). Using two datasets of participants (n=21) drifting in and out of sleep, we show that neural rhythms can be predicted from brainwide fMRI dynamics in out-of-sample subjects, and that different patterns of fMRI regions predict alpha (8-12Hz) and delta (1-4Hz) EEG power. Alpha was primarily predicted by arousal-controlling subcortex and V1, while delta predictions relied on a large number of primarily cortical regions, with significant contributions from the putamen and non-gray matter components. Our results reveal the brainwide activity underlying key neural rhythms involved in cognition and arousal, and enables discovery of the large-scale dynamics linked to neural rhythms, with applications to diverse neuroscience questions.

Keywords: machine learning; neural rhythms; EEG; fMRI; sleep

Introduction

Neural rhythms play a key role in arousal and cognition, and the presence of different rhythms defines distinct brain states, such as the different stages of sleep. However, the brainwide activity that underlies the appearance of specific EEG oscillations is not well understood. Acquiring simultaneous fMRI with EEG enables measurement of the brainwide activity across all cortical and subcortical regions, but analyzing these data is challenging due to the complex relationship between hemodynamics and electrophysiology. Traditional analytical approaches convolve the EEG signal with a hemodynamic response function (HRF), but static HRFs miss many true relationships, because the hemodynamic response varies substantially across the brain (Handwerker et al., 2004). Furthermore, traditional HRF fitting methods do not involve cross-validation and can thus overfit the data, leading to non-replicable results.

We introduce a novel approach to EEG-fMRI analysis that uses machine learning to model the relationship between the two modalities. We use this method to discover the brainwide fMRI patterns underlying two canonical neural rhythms: alpha (8-12Hz), associated with diverse cognitive processes and stronger during eyes-closed wakeful rest (Clayton et al., 2018), and delta (1-4Hz), associated with memory consolidation and brain waste clearance, and strongest during non-REM sleep (Fultz et al., 2019; Huber et al., 2004).

Results

We trained two sets of machine learning models to separately predict occipital EEG power in the alpha (8-12Hz) and delta (1-4Hz) bands from simultaneous accelerated fMRI (3T, 2.5mm isotropic voxels, TR=378ms in dataset 1; TR=367ms in dataset 2) using data from 21 subjects naturally drifting in and out of sleep (Fig 1A). EEG power was calculated in 5s windows, then interpolated to match fMRI timing. The fMRI data was parcellated into 84 regions (Desikan et al., 2006), and the mean timeseries was extracted from each region. Model predictors were sliding windows of 60 TRs (~22s) from the 84 parcellated fMRI regions, trained to predict the EEG point at the center of the window. EEG and fMRI data were normalized separately within each subject. Models were iteratively trained on all subjects but one, and performance (correlation between predictions and truth) were calculated on the held-out subject. All correlation values reported here are from held-out subject data. We used two model types: a linear model trained with stochastic gradient descent and L2 regularization (Jacob et al., 2024) and a neural network with temporal convolutions (Fig 1B) based on Syeda et al. (2023).

Models were first trained under 5 input conditions: all parcellated fMRI regions ('all'), all regions but with a 2000 TR circular shift of the fMRI data to break the true relationship with EEG ('control'), only the 'cortical' regions, only the 'subcortical' regions, and only the non-gray matter regions ('non-GM') (Fig 1C, 1F). We found that distinct fMRI regions carried distinct information about each EEG band: alpha power was most strongly predicted by subcortical regions, while delta predictions benefited most from cortical regions. Delta could also be significantly predicted by non-gray matter regions, while alpha could not, likely owing to the known coupling between delta power and cerebrospinal fluid flow (Fultz et al., 2019). These patterns were consistent across the two model types. Representative predictions demonstrate that models captured both short- and long-timescale fluctuations (Fig 1D, 1G). The neural network did not generalize to held-out subjects significantly better than the linear model, implying the relationship between fMRI and alpha/delta EEG power may be less complex than expected. Future work with additional neural network architectures is needed to assess the complexity of this relationship.

To identify finer-scale information, we trained the linear model to predict EEG power from each bilateral parcellated fMRI region, and found that alpha could be significantly predicted by each region of the arousal-controlling subcortex (thalamus, the dorsal striatum separately as caudate and putamen, and pallidum), along with the cuneus (primary visual cortex) (Fig 1E). For delta predictions, the non-gray matter regions were included along with each bilateral gray matter region, as they demonstrated significant delta-predictive performance on their own (Fig 1F); this allowed us to correct for the non-neuronal fMRI components predictive of delta and identify uniquely neural information. We found that the only region that could significantly improve delta predictions over the non-gray matter regions on their own was the putamen (Fig 1H). Given that delta predictions had demonstrated superior performance when using data from all parcellated regions and from the entire cortex (Fig 1F), it is

likely that delta is diffusely represented across the brain, relying on widespread information that primarily spans the cortex with contributions from the putamen and cerebrospinal fluid flow. This broad cortical involvement may reflect the role of slow rhythms (~1-2Hz) in information transfer and memory consolidation during sleep (Helfrich et al., 2019). Similarly, the putamen is implicated in memory consolidation processes (Ribeiro et al., 2004). On the other hand, alpha was significantly predicted by several individual regions, suggesting that a more focal set of regions carry redundant information about the dynamics of alpha rhythms.

Our results reveal the brainwide dynamics underlying canonical neural rhythms and the brain states they represent, and establish the analytical groundwork for future investigations. This approach could, for instance, help determine how brainwide activity underlying neural rhythms might change as a function of aging, disease, or different cognitive processes.

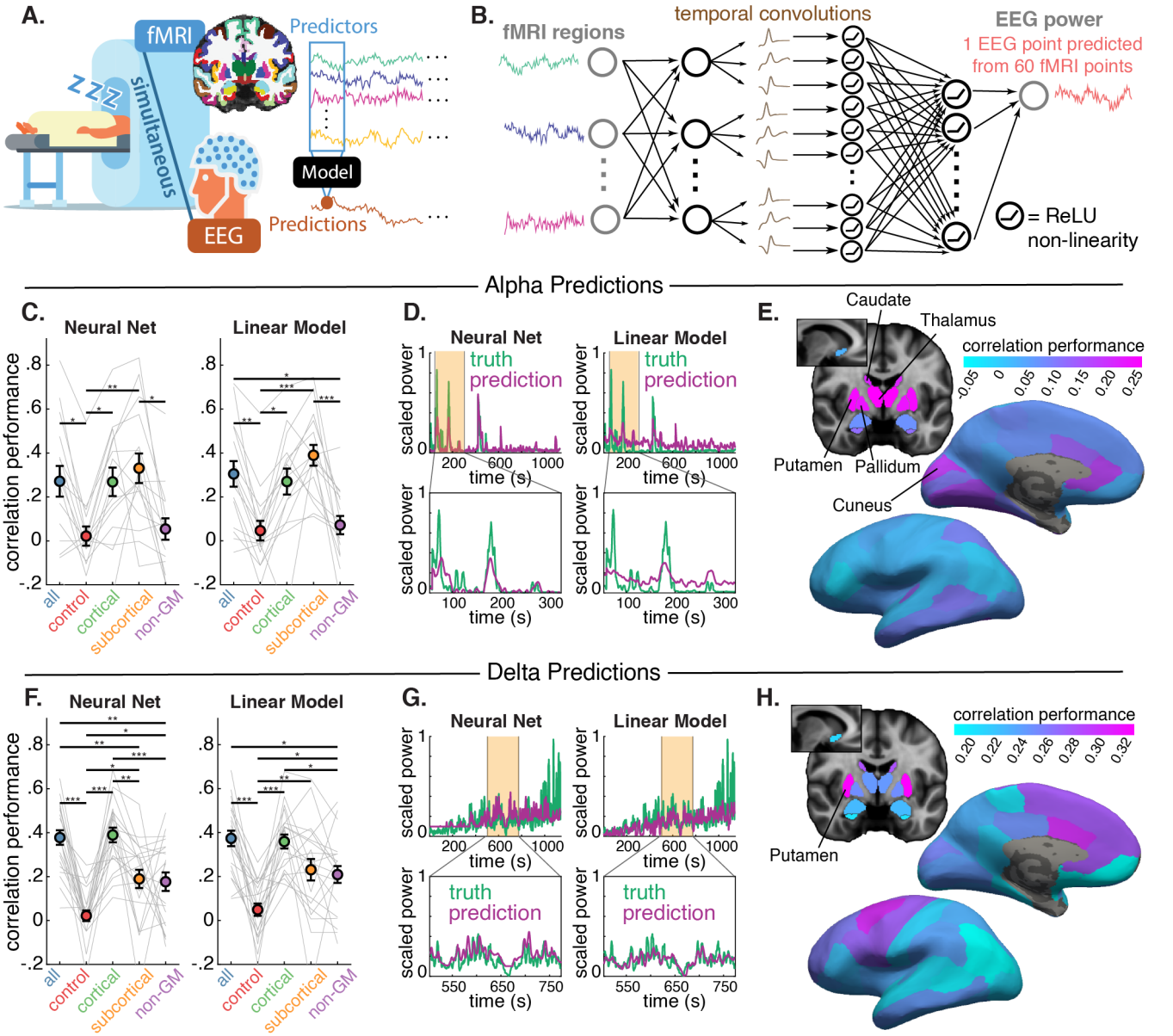

Fig 1: EEG power in the alpha and delta bands were predicted from simultaneous fMRI as subjects drifted in and out of sleep. **A.** Models were trained on sliding sequences of 60 fMRI TRs to predict each EEG point (interpolated to match fMRI TRs). **B.** Neural network structure. **C.** Correlation between predictions on held-out subjects and ground truth when models were trained under 5 input conditions (see 'Results' text for condition details). Gray lines represent subjects; colored circles show mean. Error bars are SEM. * p<0.05; ** p<0.01; *** p<0.001; Tukey's HSD. **D.** Example alpha predictions from 'subcortical' condition on a held-out subject. **E.** Performance when linear models were trained on individual bilateral regions to predict alpha. Text labels point out regions that were significantly better than control when training on single regions. **F.** Same as C, but for delta. **G.** Example delta predictions from 'all' condition on a held-out subject. **H.** Performance when linear models were trained on bilateral regions combined with non-gray matter regions (to correct for physiological components) to predict delta. Text label points out the region that was significantly better than non-gray matter regions alone.

Acknowledgments

This work was funded by National Institutes of Health grants R01-AG070135, U19-NS128613, and U19-NS123717; the Sloan Fellowship, the McKnight Scholar Award, the Pew Biomedical Scholar Award, the Simons Collaboration on Plasticity in the Aging Brain (811231), and the One Mind Rising Star Award. Resources were provided by NSF Major Research Instrumentation grant BCS-1625552.

References

Clayton, M. S., Yeung, N., & Cohen Kadosh, R. (2018). The many characters of visual alpha oscillations. *European Journal of Neuroscience*, *48*(7), 2498–2508. https://doi.org/10.1111/ejn.13747

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., Buckner, R. L., Dale, A. M., Maguire, R. P., Hyman, B. T., Albert, M. S., & Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, *31*(3), 968–980. https://doi.org/10.1016/j.neuroimage.2006.01.021

Fultz, N. E., Bonmassar, G., Setsompop, K., Stickgold, R. A., Rosen, B. R., Polimeni, J. R., & Lewis, L. D. (2019). Coupled electrophysiological, hemodynamic, and cerebrospinal fluid oscillations in human sleep. *Science*, *366*(6465), 628–631. https://doi.org/10.1126/science.aax5440

Handwerker, D. A., Ollinger, J. M., & D'Esposito, M. (2004). Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. *NeuroImage*, *21*(4), 1639–1651. https://doi.org/10.1016/j.neuroimage.2003.11.029

Helfrich, R. F., Lendner, J. D., Mander, B. A., Guillen, H., Paff, M., Mnatsakanyan, L., Vadera, S., Walker, M. P., Lin, J. J., & Knight, R. T. (2019). Bidirectional prefrontal-hippocampal dynamics organize information transfer during sleep in humans. *Nature Communications*, *10*(1), 1–16. https://doi.org/10.1038/s41467-019-11444-x

Huber, R., Ghilardi, M. F., Massimini, M., & Tononi, G. (2004). *Local sleep and learning*. *430*(July), 4–7.

Jacob, L. P. L., Bailes, S. M., Williams, S. D., Stringer, C., & Lewis, L. D. (2024). Distributed fMRI dynamics predict distinct EEG rhythms across sleep and wakefulness. *BioRxiv*. https://doi.org/10.1101/2024.01.29.577429

Ribeiro, S., Gervasoni, D., Soares, E. S., Zhou, Y., Lin, S. C., Pantoja, J., Lavine, M., & Nicolelis, M. A. L. (2004). Long-lasting novelty-induced neuronal reverberation during slow-wave sleep in multiple forebrain areas. *PLoS Biology*, *2*(1), 126–137. https://doi.org/10.1371/journal.pbio.0020024

Syeda, A., Zhong, L., Tung, R., Long, W., Pachitariu, M., & Stringer, C. (2023). Facemap: a framework for modeling neural activity based on orofacial tracking. *Nature Neuroscience*. https://doi.org/10.1038/s41593-023-01490-6