ehealth workgroup Universität Regensburg

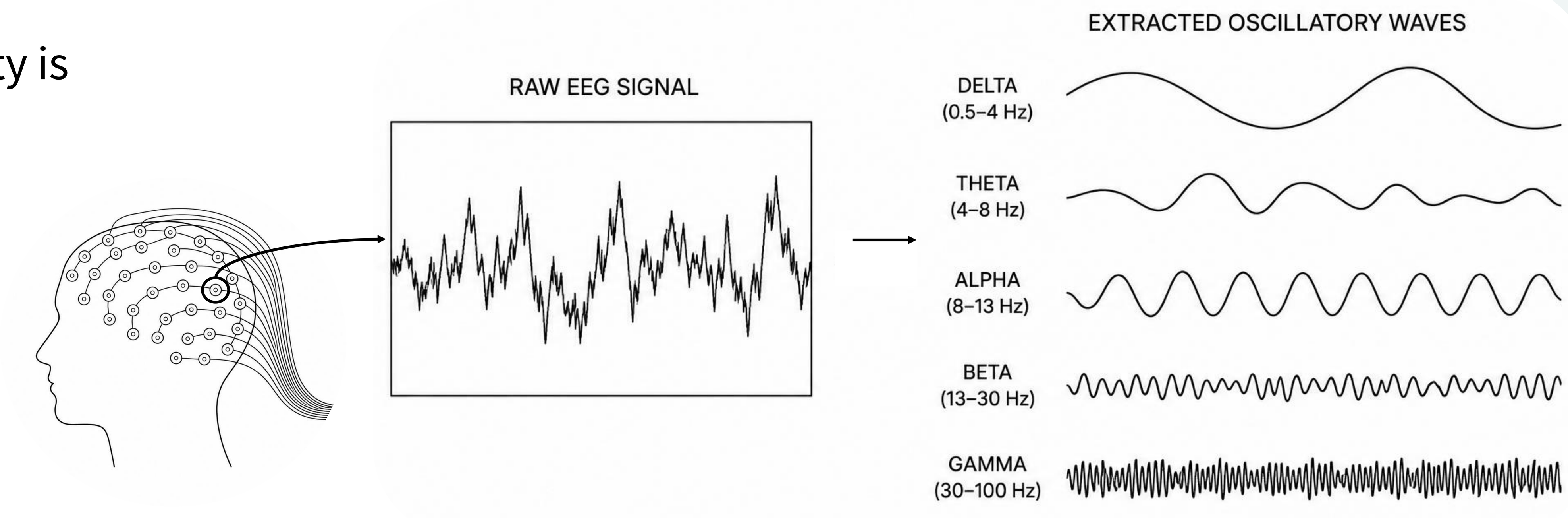
Universität Regensburg

Feature engineering in tinnitus research: data-driven definitions of oscillatory brain activity

Nina Wunder, Sven Hilbert, Winfried Schlee

Background

- electroencephalography (EEG) oscillatory activity is a key marker of interest for pathological brain states, including tinnitus[1]
- Feature extraction is based on predefined frequency band definitions that vary across studies and are often adopted without validation[2,3]



Do data-driven definitions of frequency bands differ from established ones?
What are the consequences for tinnitus–control classification performance?

Methods

Step 1: Unsupervised Machine Learning

- group frequency bins (1 Hz – 80 Hz) to bands by covariance of power spectral density (PSD) across time using **hierarchical** (hard) and **fuzzy c-means** (soft) **clustering**
- evaluate cluster quantity and quality with method-specific metrics like silhouette score

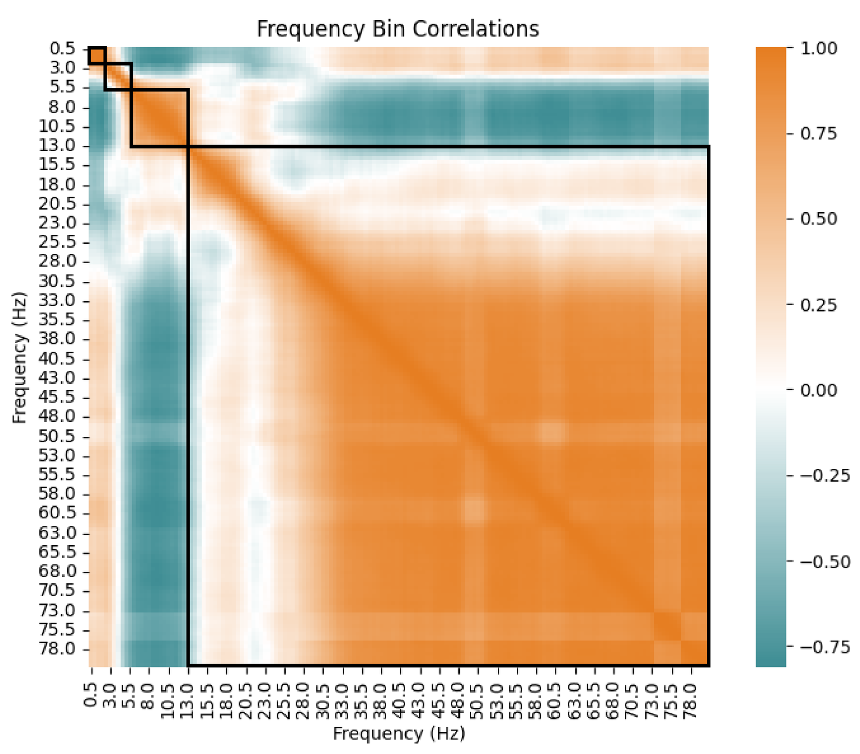
N = 562 (281/281 tinnitus/healthy controls)

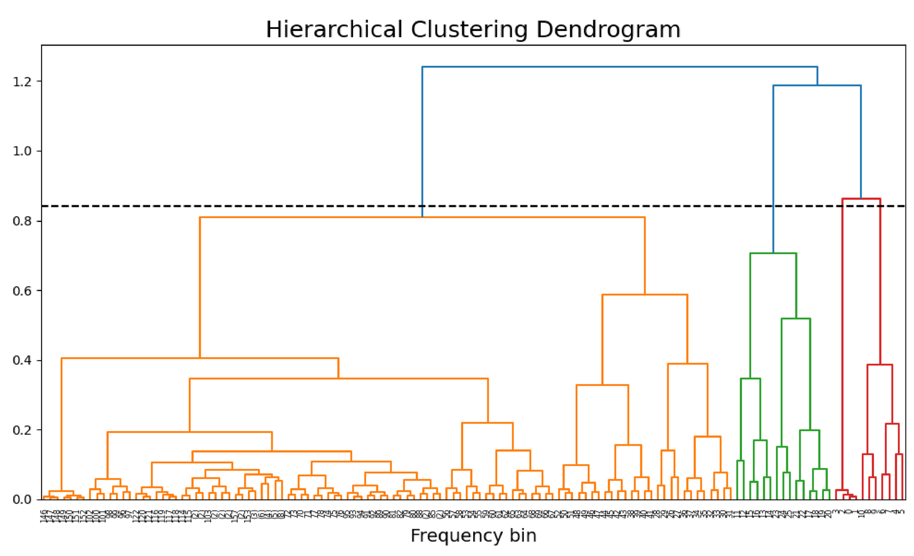
Step 2: Supervised Machine Learning

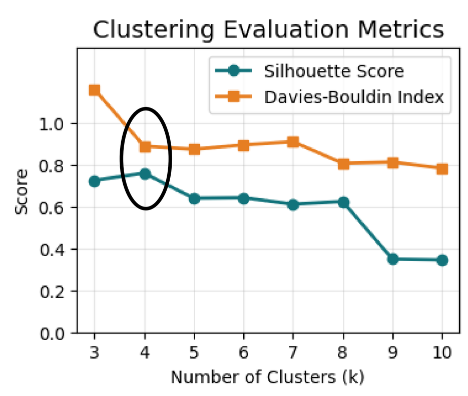
- train **binary tinnitus classifiers** with frequency band features derived from clustering versus from literature
- compare performance with ROC-AUC

Results

Hierarchical clustering identified an optimum of **four clusters**

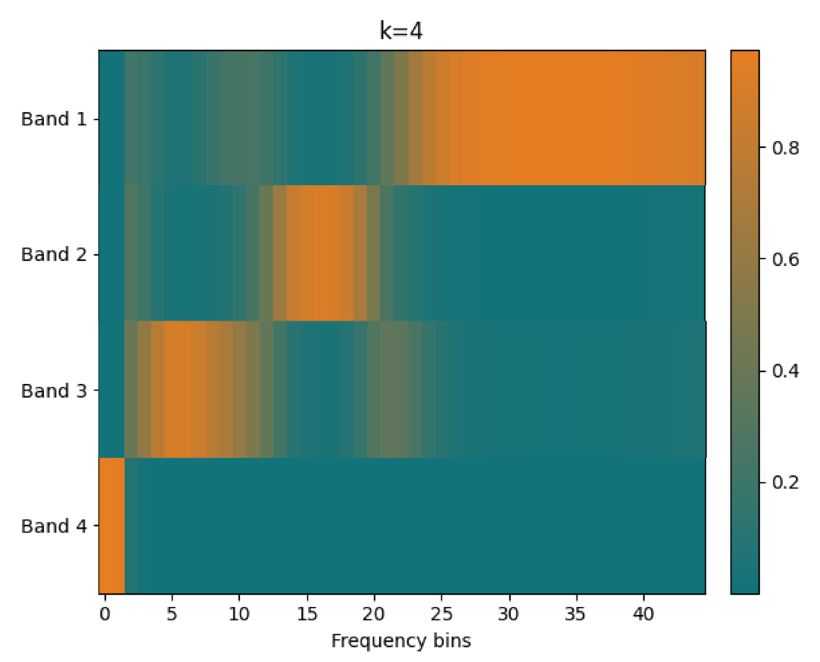


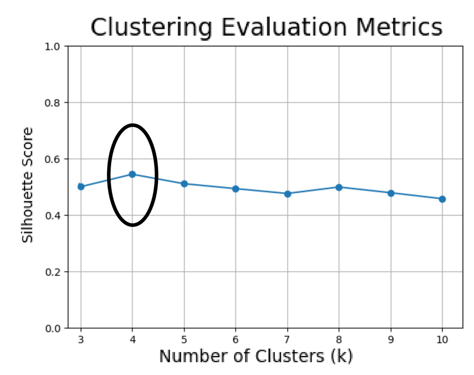


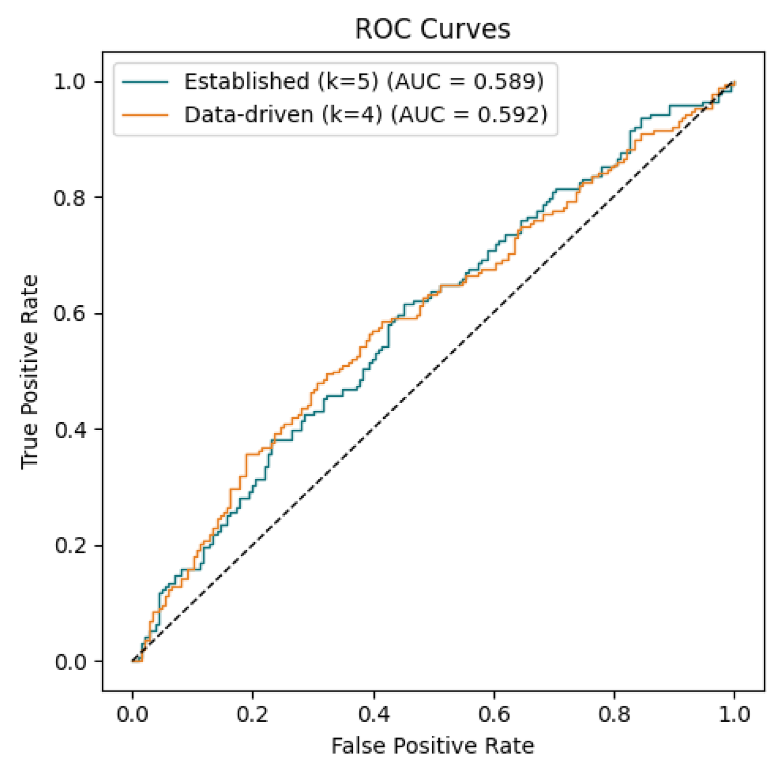


- beta and gamma are grouped together

Fuzzy c-means clustering partially supports this finding and indicates **overlap** between frequency bands

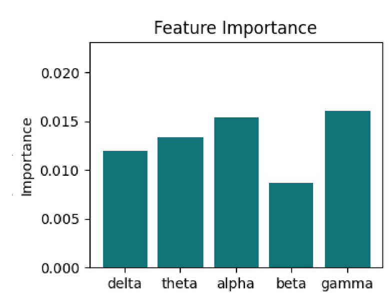


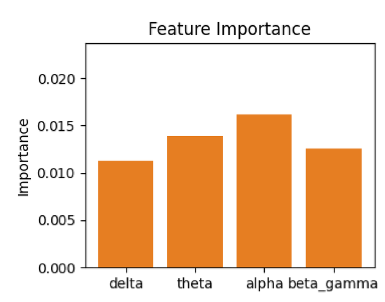




- Random Forest classification using established[4] versus data-driven* frequency band definitions achieved similarly weak discriminative performance
- Frequency band selection alone may not be sufficient to improve classification performance in this data set
- results likely reflect low signal-to-noise ratio of EEG data and inter-subject variability

**Data-driven frequency band selection: delta (0.5-2 Hz), theta (2.5-5.5 Hz), alpha (6-13 Hz), beta+gamma (13.5-80 Hz)*
Classification results are preliminary and subject to further validation.





Outlook

- compare further frequency band definitions in classification
- Compare classification algorithms & refine hyperparameter tuning
- repeat analyses using brain connectivity measures instead of PSD

Nina Wunder, M.Sc.
eHealth workgroup
Bezirksklinikum Regensburg
eMail: nina.wunder@ur.de

References
[1] De Ridder, D., Vanneste, S., Langguth, B., & Llinas, R. (2015). Thalamocortical dysrhythmia: A theoretical update in tinnitus. *Frontiers in Neurology*, *6*.
[2] Shabestari, P. S., Schoisswohl, S., Wellauer, Z., Naas, A., Kleinjung, T., Schecklmann, M., Langguth, B., & Neff, P. (2025). Prediction of acoustic tinnitus suppression using resting-state EEG via explainable AI approach. *Scientific Reports*, *15*(1), 10968.
[3] Schoisswohl, S., Schecklmann, M., Langguth, B., Schlee, W., & Neff, P. (2021). Neurophysiological correlates of residual inhibition in tinnitus: Hints for trait-like EEG power spectra. *Clinical Neurophysiology*, *132*(7), 1694–1707.
[4] Bomatter, P., Paillard, J., Garces, P., Hipp, J., & Engemann, D.-A. (2024). Machine learning of brain-specific biomarkers from EEG. *eBioMedicine*, *106*, 105259