# Dynamic biclustering for multi-subject neuroscience studies


Federica Zoe Ricci (fzricci@uci.edu)

*Joint work with E. Sudderth, J. Lee, M. Peters, M. Vannucci, M. Guindani*



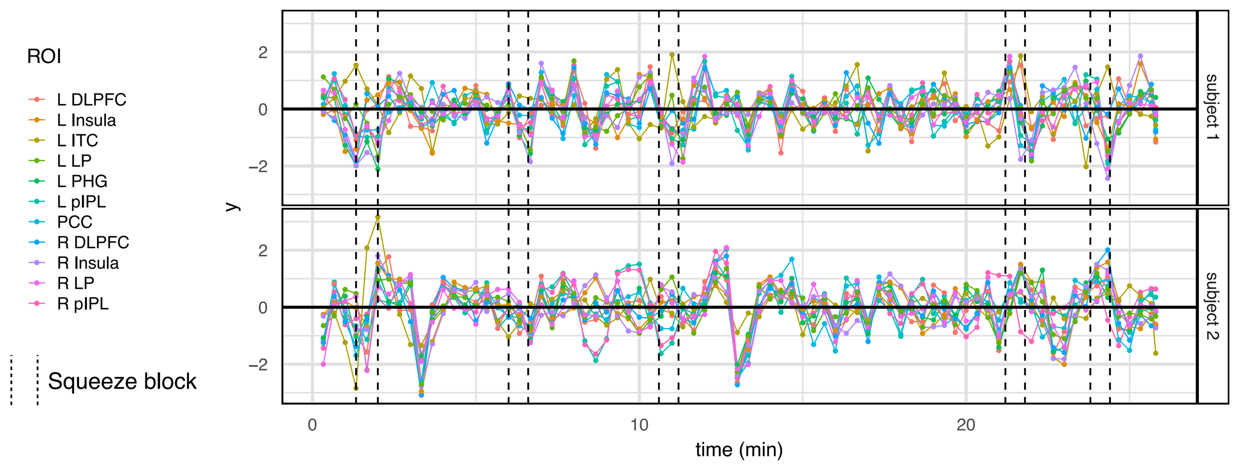
*thanks to* Hasso Plattner Institut — IT Systems Engineering | Universität Potsdam


## Motivation

Many neuroscience studies follow multiple units, or subjects, during a period of time, collecting several measurements for each unit at specific time intervals. Consider the following fMRI and EEG datasets:

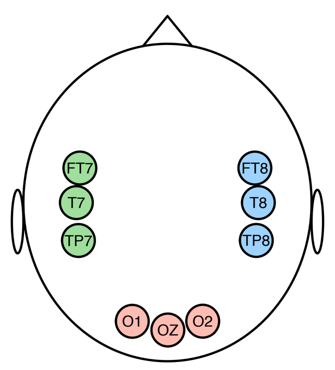
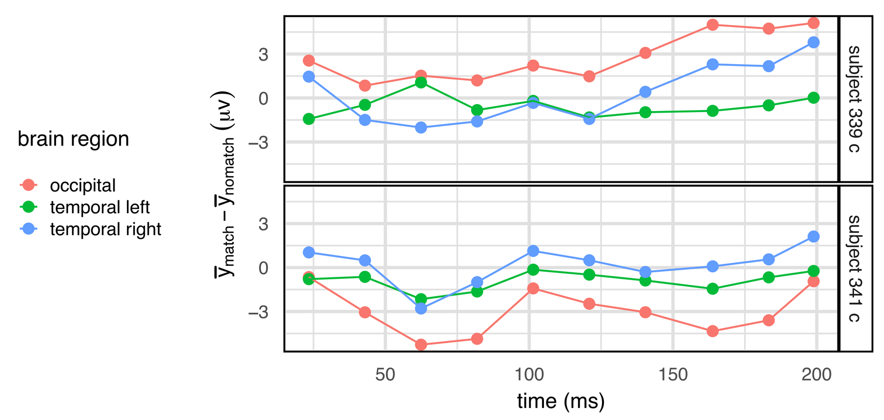
### fMRI data

Subjects = 23
ROIs = 11
Time steps = 78




### EEG data

Subjects = 121
Brain regions = 3
Time steps = 10


### Our Goal

Develop model to **simultaneously cluster time-series data along two dimensions** to identify:
- clusters of subjects who are similar throughout the experiment (***profiles***)
- groups of associated measurements (e.g. ROIs/brain regions) at each time step (***states***)

## Related work

### Nested biclusters (single time step)

Lee et al. (2013)[1] and Lin et al. (2022)[2]

Clusters along one dimension (e.g. measurement clusters) are nested within clusters along the other dimension (e.g. subject clusters).

That is, e.g., each subject cluster is identified by a specific partition of measurements

for each $i, j : s_i = s_j \rightarrow c_r^{(i)} = c_r^{(j)} \ \forall r \in R$

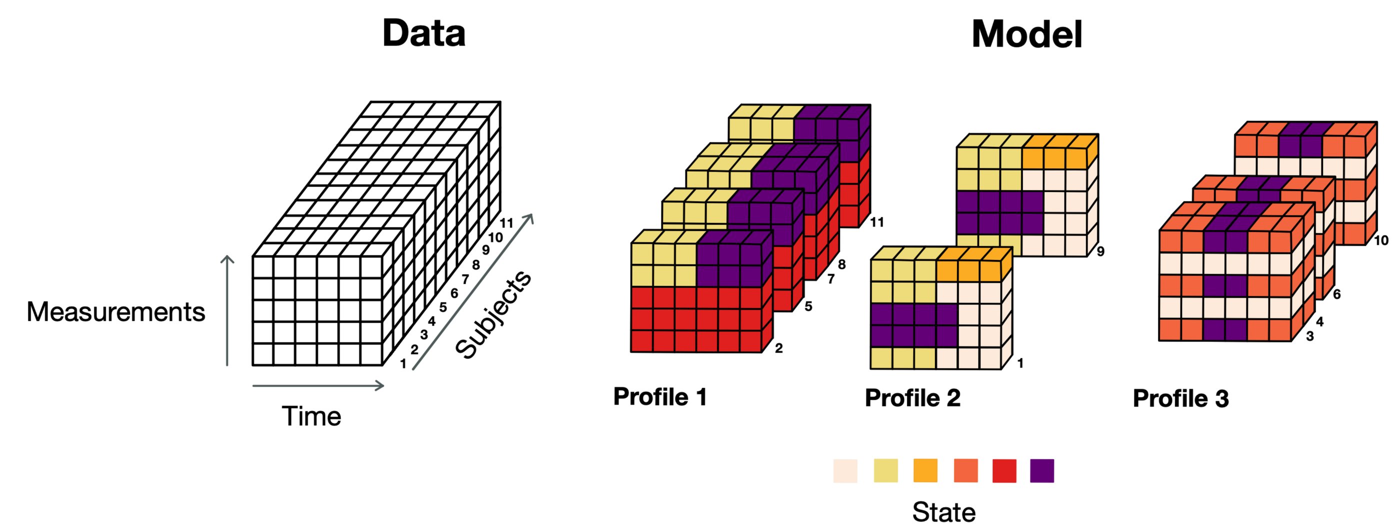
### Temporal random partitions model (tRPM) (single subject)

Page et al. (2022)[3]

Custers along one dimension (e.g. measurements) evolve over time based on a Markov model

$$c_{rt} \mid c_{r,t-1} = \begin{cases} c_{r,t-1} \text{ with prob } a_t \\ \text{new cluster } c^*_{rt} \text{ with prob } 1-a_t \end{cases}$$

## Bayesian dynamic biclustering model




### Idea

- Extend nested partition model to time-series setting
- Each subject has one of $Z$ profiles
- Each profile is identified by a time-varying partition of measurements into states
- Model evolution of measurement states similarly as tRPM

### Likelihood model

The likelihood model for an observation $Y_{i,r,t}$ of subject $i$ , measurement $r$ , at time step $t$ is determined by the state $c^{(s_i)}_{r,t} \in \{1,\ldots,K\}$, common to all subjects with the same profile $s_i$ as subject $i$

$$Y_{i,r,t} \mid c^{(s_i)}_{r,t} = k, \theta^*_k \overset{iid}{\sim} F_{\theta^*_k}$$

In our experiments we let $\theta = \{\mu_k, \sigma_k\}$ be the parameters of a location-scale t-distribution, interpretable but diffuse enough to allow for some heterogeneity between observations assigned to the same state.

### Evolution of measurement states

Account for temporal dependences by encouraging measurements to persist in the same state over consecutive time-steps, while allowing for states to change and for learning the number and position of changepoints from the data:

$$c^{(z)}_{r,t} \mid \omega^{(z)}, \gamma^{(z)}_{r,t} \overset{ind}{\sim} \gamma^{(z)}_{r,t}\, \delta_{c^{(z)}_{r,t-1}} + \left(1-\gamma^{(z)}_{r,t}\right) \text{Categorical}(\omega^{(z)}_1, \ldots, \omega^{(z)}_K)$$

where $\gamma^{(z)}_{r,t} \mid a^{(z)}_t \overset{ind}{\sim} \text{Bernoulli}\left(a^{(z)}_t\right)$ and $a^{(z)}_t \overset{iid}{\sim} \text{Beta}(\alpha, \beta)$.

### Profile and state assignments

Finite approximation to Dirichlet process (DP) to learn number of profiles from data:

$$s_i \mid \pi \overset{iid}{\mid} \text{Categorical}\,(\pi_1, \ldots, \pi_Z),$$
$$\pi \mid \zeta \sim \text{Dirichlet}(\zeta, \ldots, \zeta),$$

where we let $Z = N$, $\zeta = \dfrac{\varepsilon}{Z}$ and $\varepsilon \sim \text{Gamma}(b_1, b_2)$ (Malsiner-Walli et al. 2016)[4]

Similar model for the states (approximation of hierarchical DP). States are shared across profiles but state probabilities are profile-specific:

$$\omega^{(z)} \mid \omega_0 \overset{iid}{\sim} \text{Dirichlet}(\phi\,\omega_{01}, \ldots, \phi\,\omega_{0K}), \quad z = 1, \ldots, Z$$

$$\omega_0 \mid \eta \sim \text{Dirichlet}\left(\frac{\eta}{K}, \ldots, \frac{\eta}{K}\right)$$

## Statistical inference

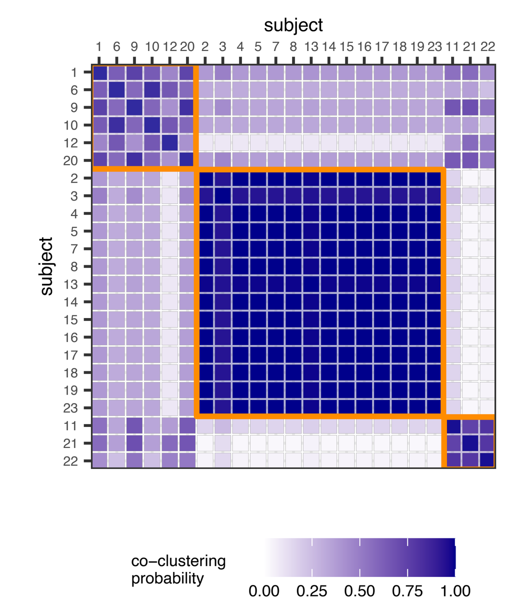
Perform posterior inference via Markov Chain Monte Carlo.

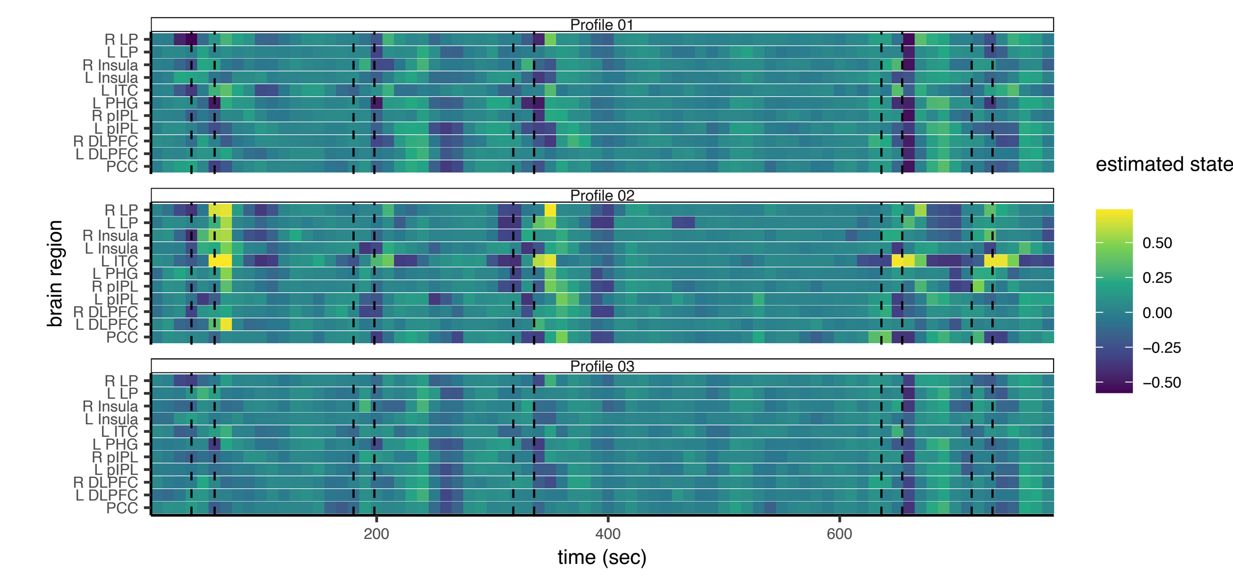
1. **Profile variables:** Update profile probabilities $\boldsymbol{\pi}$ conditional on subject assignments, resample their concentration hyperparameter $\zeta$, update subjects' assignment to profiles $(s_1, \ldots, s_N)$ conditional on $\boldsymbol{\pi}$.
2. **State variables:** Update vector of global state probabilities $\boldsymbol{\omega}_0$, and its concentration hyperparameter $\eta$. For each profile $z$, sample the profile-specific vector of state probabilities $\boldsymbol{\omega}_z$ conditional on $\boldsymbol{\omega}_0$. Update the state persistence indicator $\gamma^{(z)}_{r,t}$ and the state assignment $c^{(z)}_{r,t}$ for each profile, measurement and time step. Update the probability of state persistence $a^{(z)}_t$ conditional on all state persistence indicators, for each profile and time step.
3. **Likelihood parameters**: For each state, sample its associated likelihood parameters $\boldsymbol{\theta}_k$ conditional on state assignment sequences for all observations.

### Profiles

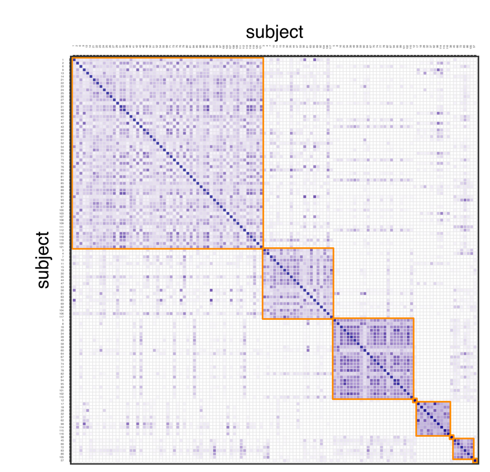
### State assignments

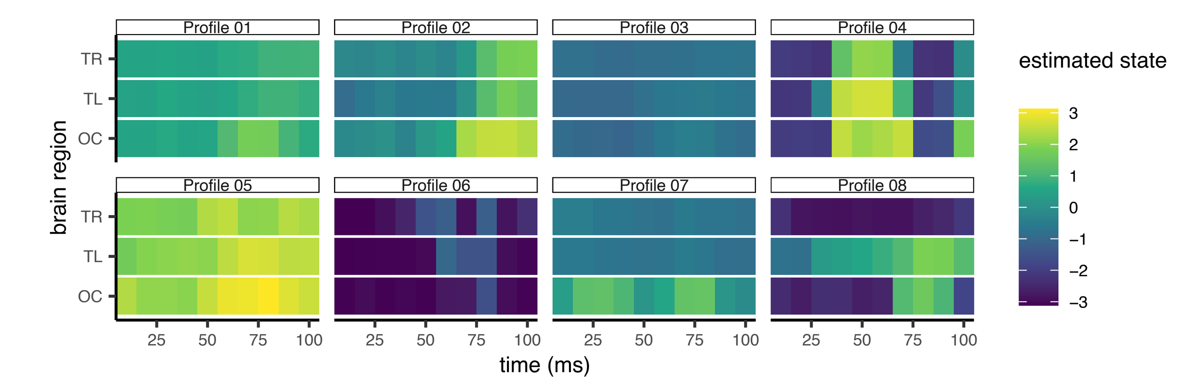
fMRI data




EEG data




## Future directions

- Incorporate time-invariant covariates —> study how clustering depends on covariates
- Relax assumption that subjects with the same profile must share the same temporal partition for all measurements —> can include more measurements and automatically select relevant ones
- Allow partition of subjects to vary over time —> analyze longer time periods or experiments with multiple heterogeneous tasks

### References


1. Lee, Juhee, Peter Müller, Yitan Zhu, and Yuan Ji. "A nonparametric Bayesian model for local clustering with application to proteomics." *Journal of the American Statistical Association* 108.503 (2013): 775-788.
2. Lin, Qiaohui, Giovanni Rebaudo, and Peter Mueller. "Separate exchangeability as modeling principle in Bayesian nonparametrics." arXiv preprint arXiv:2112.07755 (2021).
3. Page, Garritt L., Fernando A. Quintana, and David B. Dahl. "Dependent modeling of temporal sequences of random partitions." *Journal of Computational and Graphical Statistics* 31.2 (2022): 614-627.
4. Malsiner-Walli, Gertraud, Sylvia Frühwirth-Schnatter, and Bettina Grün. "Model-based clustering based on sparse finite Gaussian mixtures." *Statistics and Computing* 26.1-2 (2016): 303-324.