School of Mathematical Sciences | Lancaster University

Generalised mixed effects models for changepoint analysis of biomedical time series

Rebecca Killick(r.killick@lancs.ac.uk)
Joint work with Mark Fiecas and Kathryn Cullen (University of Minnesota)
Paris, July 2025

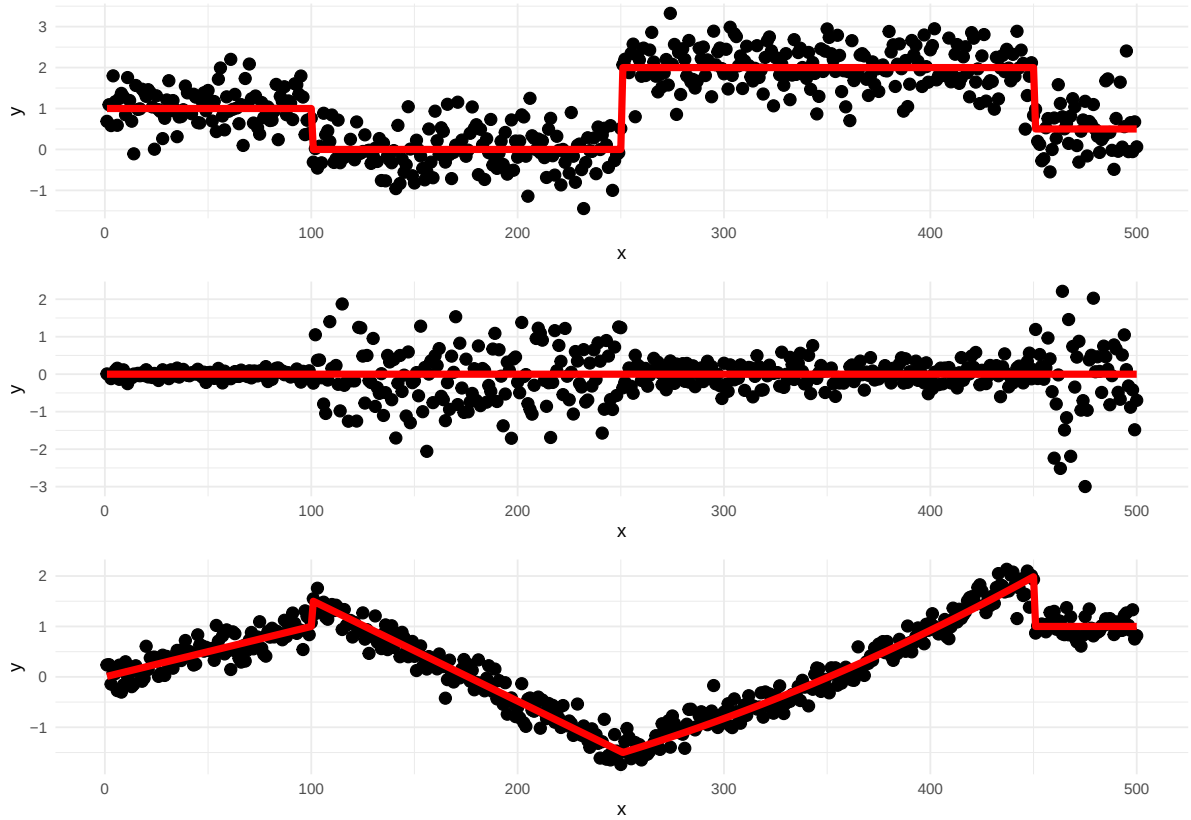
Outline

- Putting us all on the same page
- Motivation from Biomedical Time Series
- A generalised mixed effects model with changepoints
- Simulations + Data Applications
- Summary

What are changepoints?

For data $y_1, \ldots, y_n$, if a changepoint exists at τ, then $y_1, \ldots, y_\tau$ differ from $y_{\tau+1}, \ldots, y_n$ in some way.

There are many different types of change.



Research Areas

- Online vs offline
- Single change vs multiple changes
  - Multiple changepoint search
- Univariate and multivariate
- Parametric and non-parametric
- Frequentist and Bayesian
- Types of change
- Inference
  - Detection vs localisation
  - Confidence sets
  - Downstream analysis
  - Theoretical and application driven
- Others I have missed!

This talk

- Online vs offline
- Single change vs multiple changes
  - Multiple changepoint search
- Univariate and multivariate
- Parametric and non-parametric
- Frequentist and Bayesian
- Types of change - first order and second order random effect models
- Inference
  - Detection vs localisation
  - Confidence sets - ask me later if interested
  - Downstream analysis
  - Theoretical and application-driven
- Others I have missed!

FMRI - BRIDGES

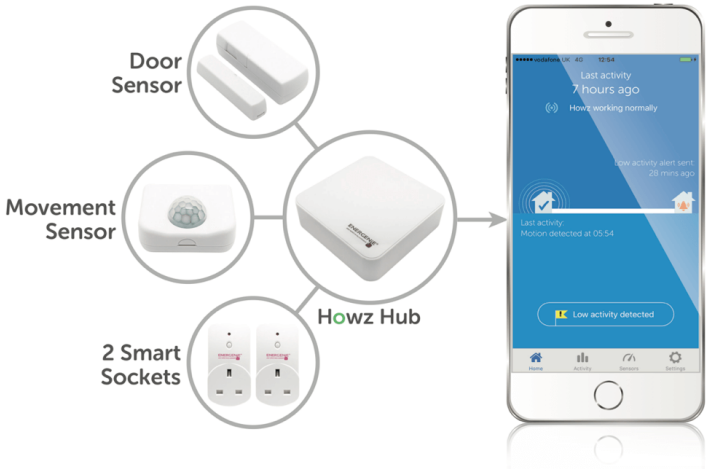
Brain Imaging Development of Girls' Emotion and Self (BRIDGES) Study

- Resting state data from 137 girls, no NSSI to severe NSSI.
- 56 spatial locations, 4 regions of interest
- 904 observations from each individual

At home monitoring

The Howz Smart System:

- Discreet passive sensors.
- Measure activity in the home.
- Learn about your daily routine.



Objective:

- Identify abnormal behaviour:
- Indicate a decline in health or well-being in an older person.

Data:

- 4 people
- Binary yes/no activity in 15 mins
- 8 weeks (56 days) of activity data

Commonalities

- Mutiple people
- Repeated observations per person
- Common structure with random variation across people
- Potential changes over time

Commonly used model = random effects model

$$\mathbb{E}(y_{i,j}) = \mu + \mu_i + \epsilon_{i,j}$$

GLMER

Generalized Linear Mixed Effects Regression (GLMER) model.

$$y_i|b \sim \text{Distr}\left(\mu_i, \sigma^2\right) \tag{1}$$

$$g(\mu) = X\beta + Zb, \tag{2}$$

- X - external regressors with coefficients β
- Distr - conditional distribution
- Z - random effects with coefficients b

Standard GLMER estimators work well.

$$y_i|b \sim \text{Distr}\left(0, \sigma_i^2\right) \tag{3}$$

(4)

We want to write the covariance in a similar random effects form.

For Uniform-block (UB) structure we get

$$\Sigma = \widetilde{\Sigma}_\epsilon \circ \mathbf{Id}(\mathbf{k}) + \Sigma_\mu \circ \mathbf{J}(\mathbf{k}). \tag{5}$$

where:

- $\widetilde{\Sigma}_\epsilon = \text{diag}(\sigma^2_{\epsilon,1}, \ldots, \sigma^2_{\epsilon,K})$
- $\mathbf{Id}(\mathbf{k}) = \text{Bdiag}(\text{Id}_{k_1}, \ldots, \text{Id}_{k_K})$, where Bdiag$(\cdot)$ constructs a block-diagonal matrix
- $\Sigma_\mu = (b_{ij})$ be a $K \times K$ covariance matrix
- $\mathbf{J}(\mathbf{k}) = (\mathbf{1}_{k_u \times k_v})$
- $\circ$ be the block Hadamard product

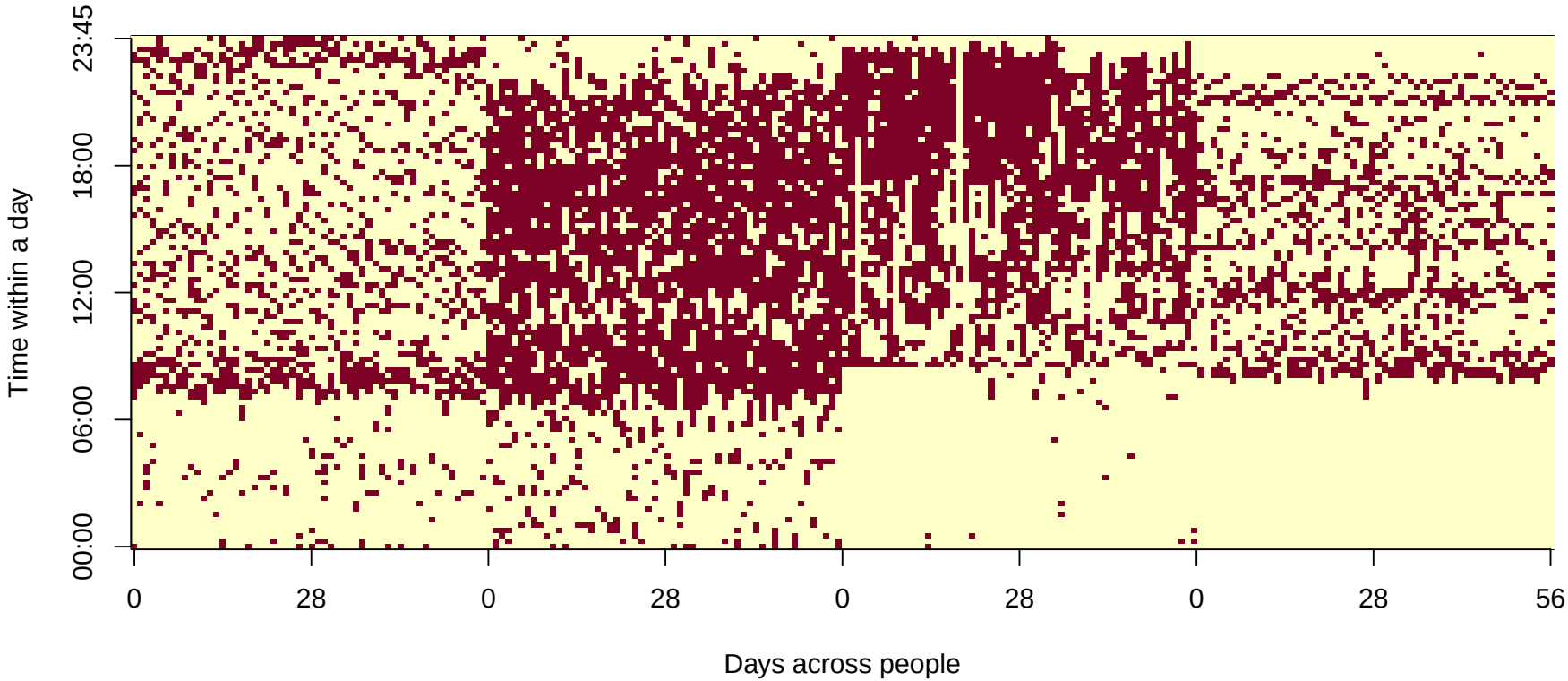
For Heterogeneous-block (HB) structure we get

$$\Sigma = \widetilde{\Sigma}_{\epsilon} + \Sigma_{\mu} \circ \mathbf{J}(\mathbf{k}). \tag{6}$$

- May be prohibitive for high dimensional (n relative to K)

Estimate parameters by block-wise method of moments (see ArXiV paper).

What now?

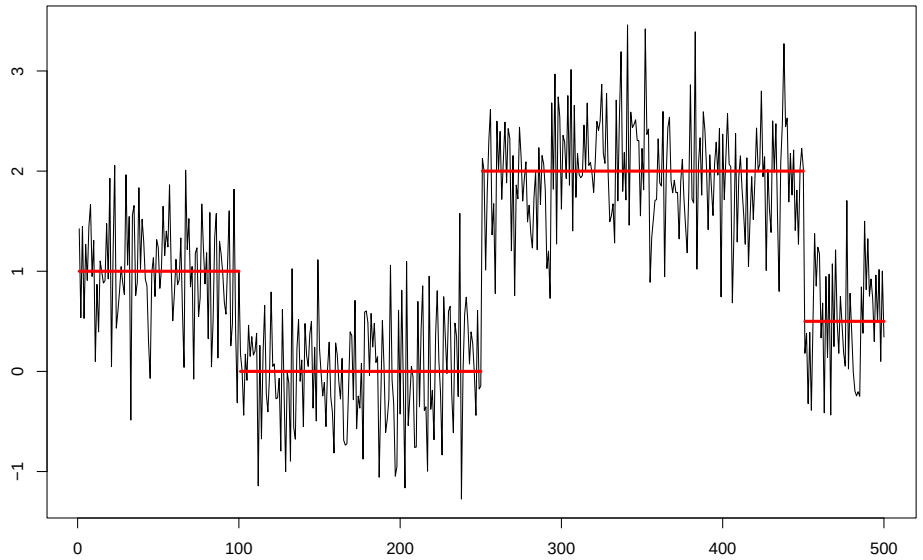

Time within a day
00:00
06:00
12:00
18:00
23:45
0
28
0
28
0
28
0
28
56
Days across people

Segway

Assume we have time-series data where

$$Y_t|\theta_t \sim \mathrm{N}(\theta_t, 1),$$

but where the means, θ_t, are piecewise constant through time.



Inferring Changepoints

We want to infer the number and position of the points at which the mean changes. One approach:

Likelihood Ratio Test

To detect a single changepoint we can use the likelihood ratio test statistic:

$$LR = \max_{\tau}\{\ell(y_{1:\tau}) + \ell(y_{\tau+1:n}) - \ell(y_{1:n})\}.$$

We infer a changepoint if $LR > \beta$ for some (suitably chosen) β. If we infer a changepoint its position is estimated as

$$\tau = \arg\max\{\ell(y_{1:\tau}) + \ell(y_{\tau+1:n}) - \ell(y_{1:n})\}.$$

This can test can be repeatedly applied to new segments to find multiple changepoints.

Likelihood Ratio Tests

Define m to be the number of changepoints, with positions $\boldsymbol{\tau} = (\tau_0, \tau_1, \ldots, \tau_{m+1})$ where $\tau_0 = 0$ and $\tau_{m+1} = n$.

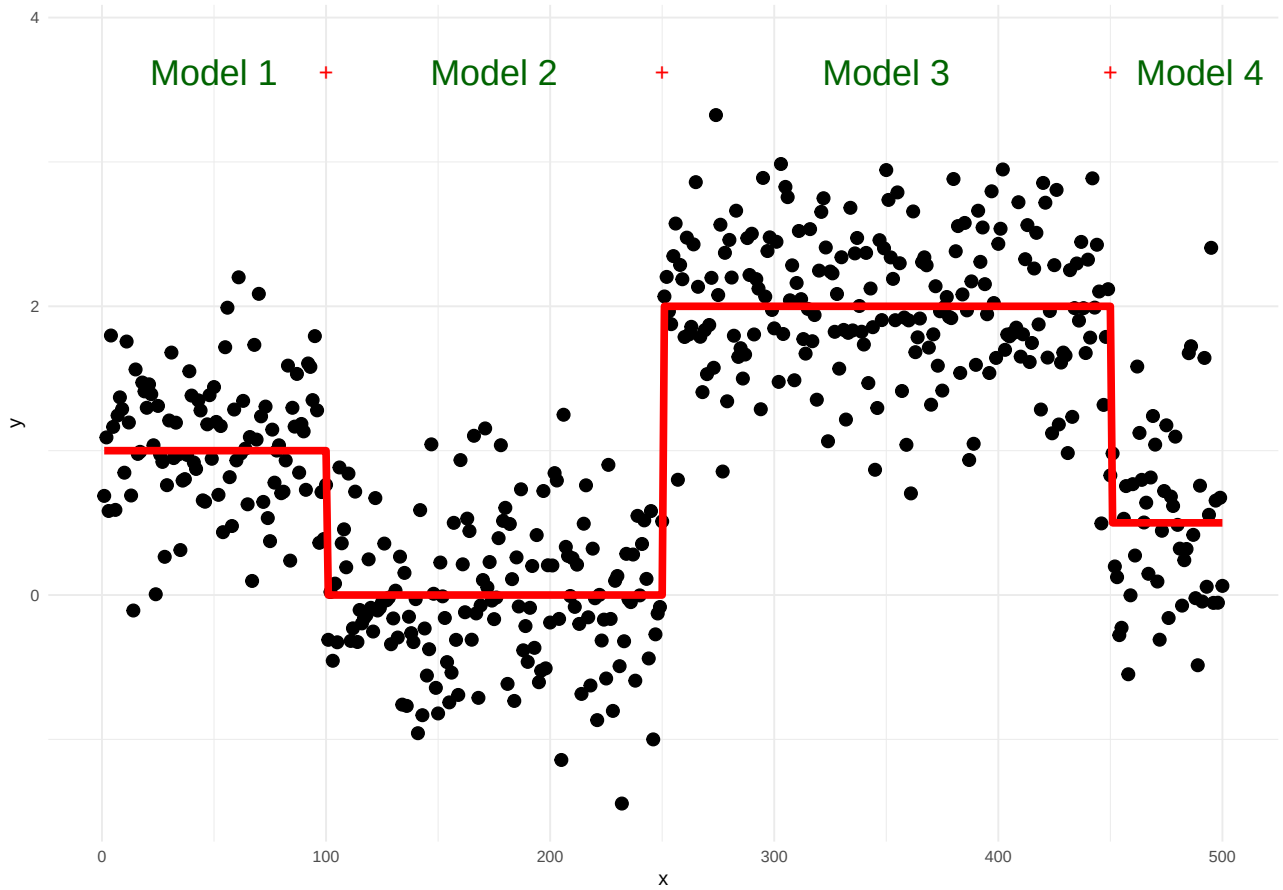
Then one application of the Likelihood ratio test can be viewed as

$$\min_{m \in \{0,1\}, \boldsymbol{\tau}} \left\{ \sum_{i=1}^{m+1} \left[-\ell(y_{\tau_{i-1}:\tau_i}) \right] + \beta m \right\}$$

Repeated application is thus aiming to minimise

$$\min_{m, \boldsymbol{\tau}} \left\{ \sum_{i=1}^{m+1} \left[-\ell(y_{\tau_{i-1}:\tau_i}) \right] + \beta m \right\}$$

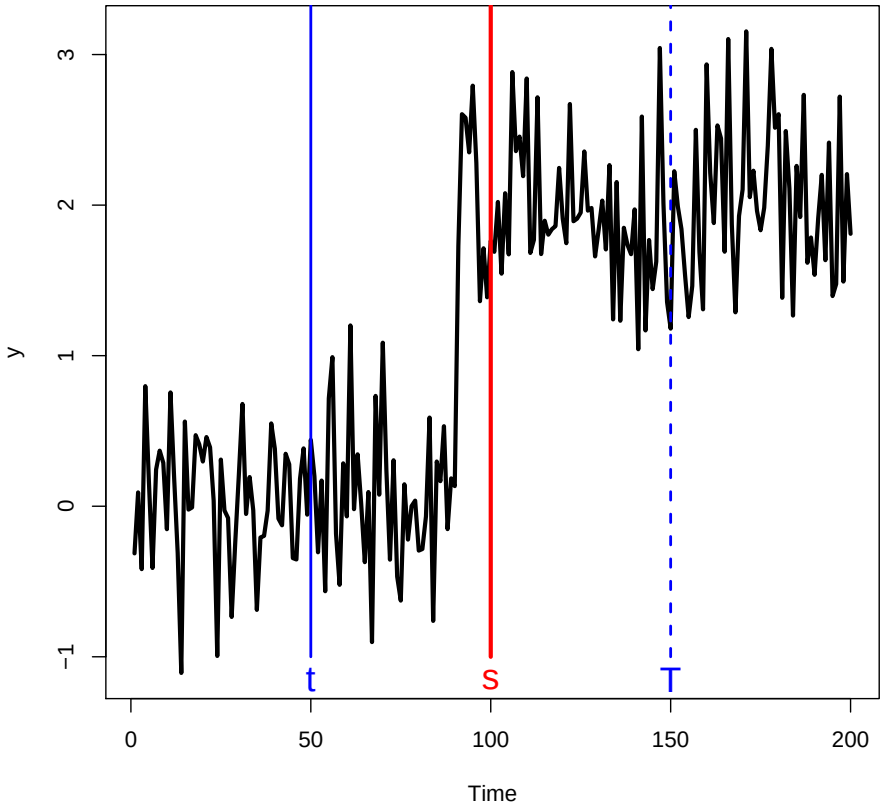
Problem



- How many changes?
- Where are the changes? 2^{n-1} possible solutions!

PELT in a nutshell

- Dynamic programming allows us to only worry about the location of the *last* change.
- Pruning means that as we go through the data we are smart about which locations are potential last change locations.

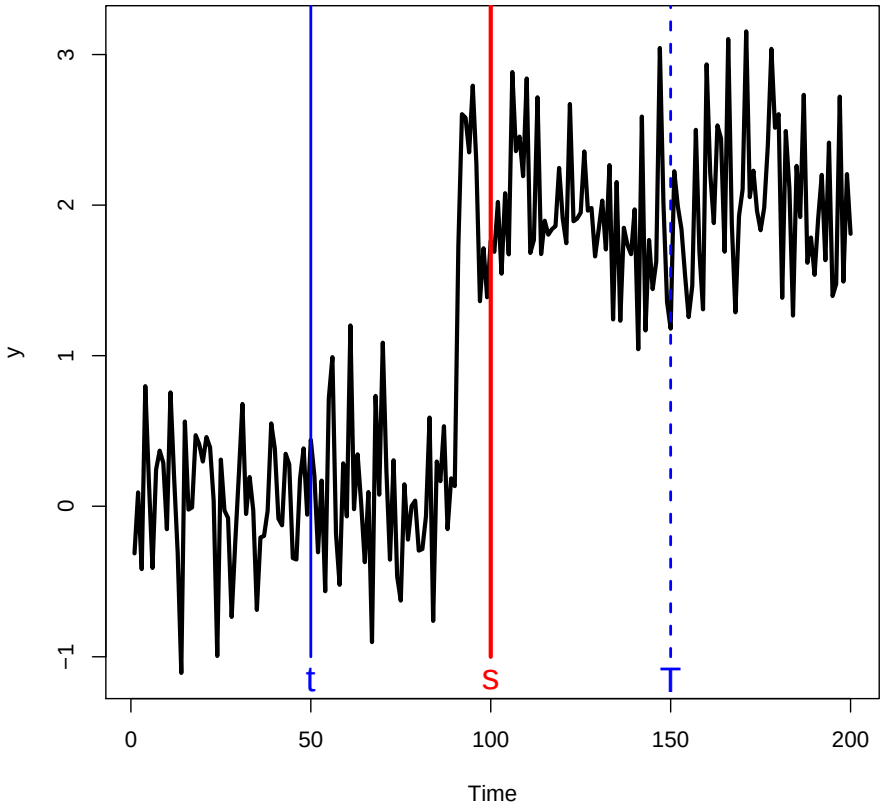


PELT: Pruning

Let $0 < t < s < T$, if

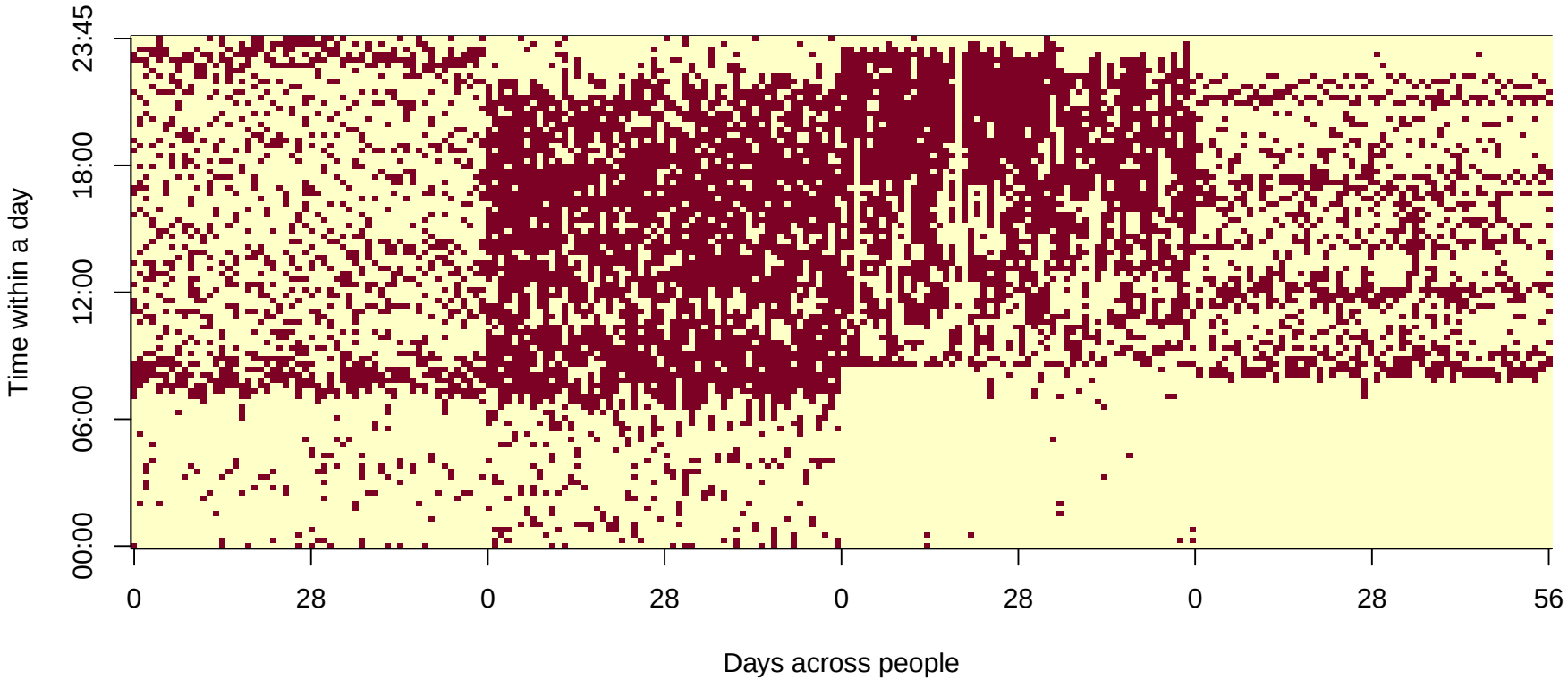
$$F(t) + \mathcal{C}(y_{(t+1):s}) < F(s)$$

then at any future time $T > s$, t can never be the optimal last changepoint prior to T.
We can prove that, under certain regularity conditions, the expected computational complexity will be $\mathcal{O}(n)$.



Final approach

- Use the GLMER or LMEC likelihood as the cost function for PELT
- Other multiple changepoint search options are available :-)



Practical considerations

- What can change?
- Minimum segment length - theoretical minimum? Covariance estimation is hard!
- Penalty for changepoint estimation - BIC
- Assume known group structure

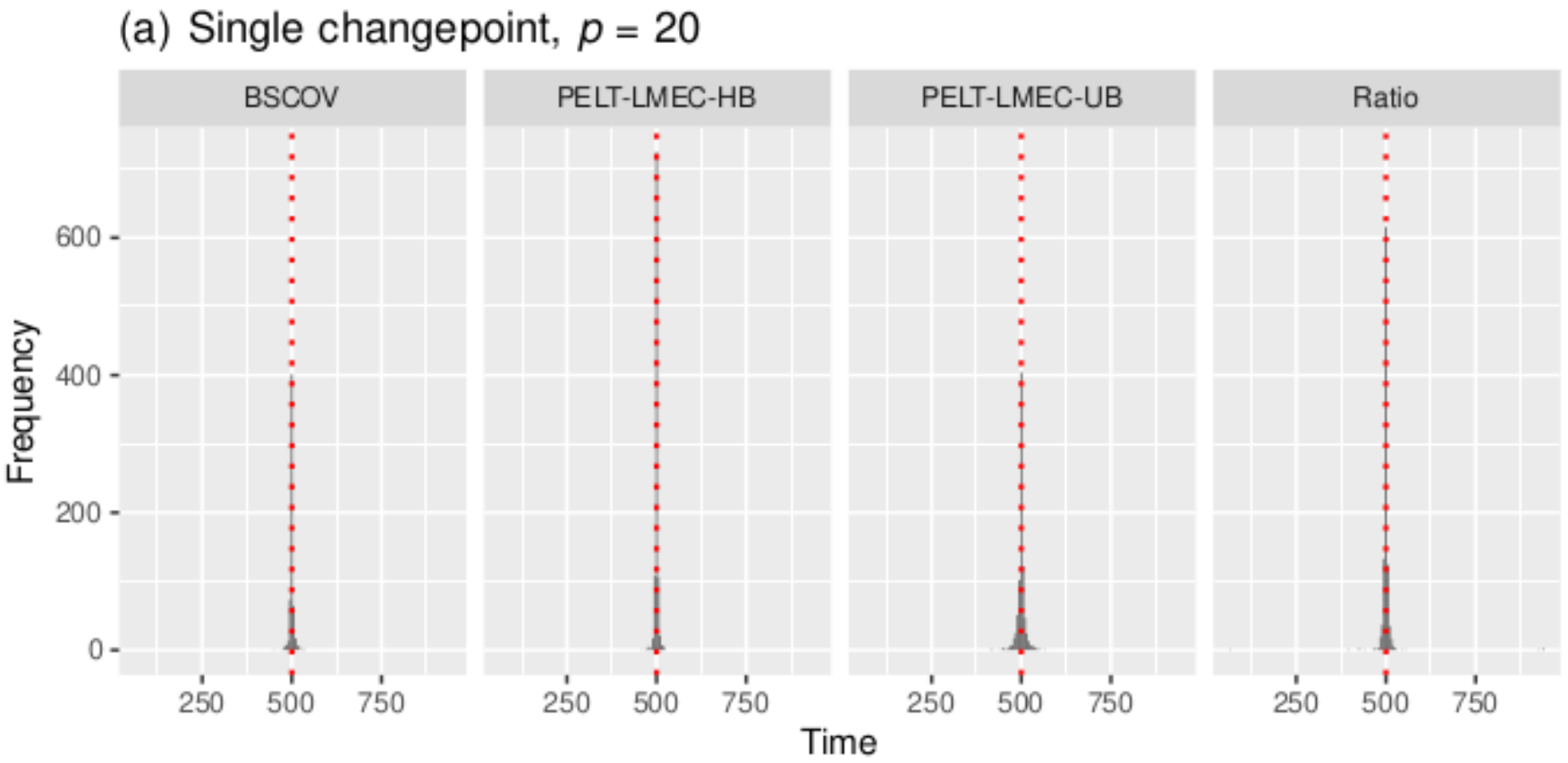
Simulations

- n=1000
- K=4 groups, 5 series per group
- Simulate covariances from Wishart distribution and manipulate

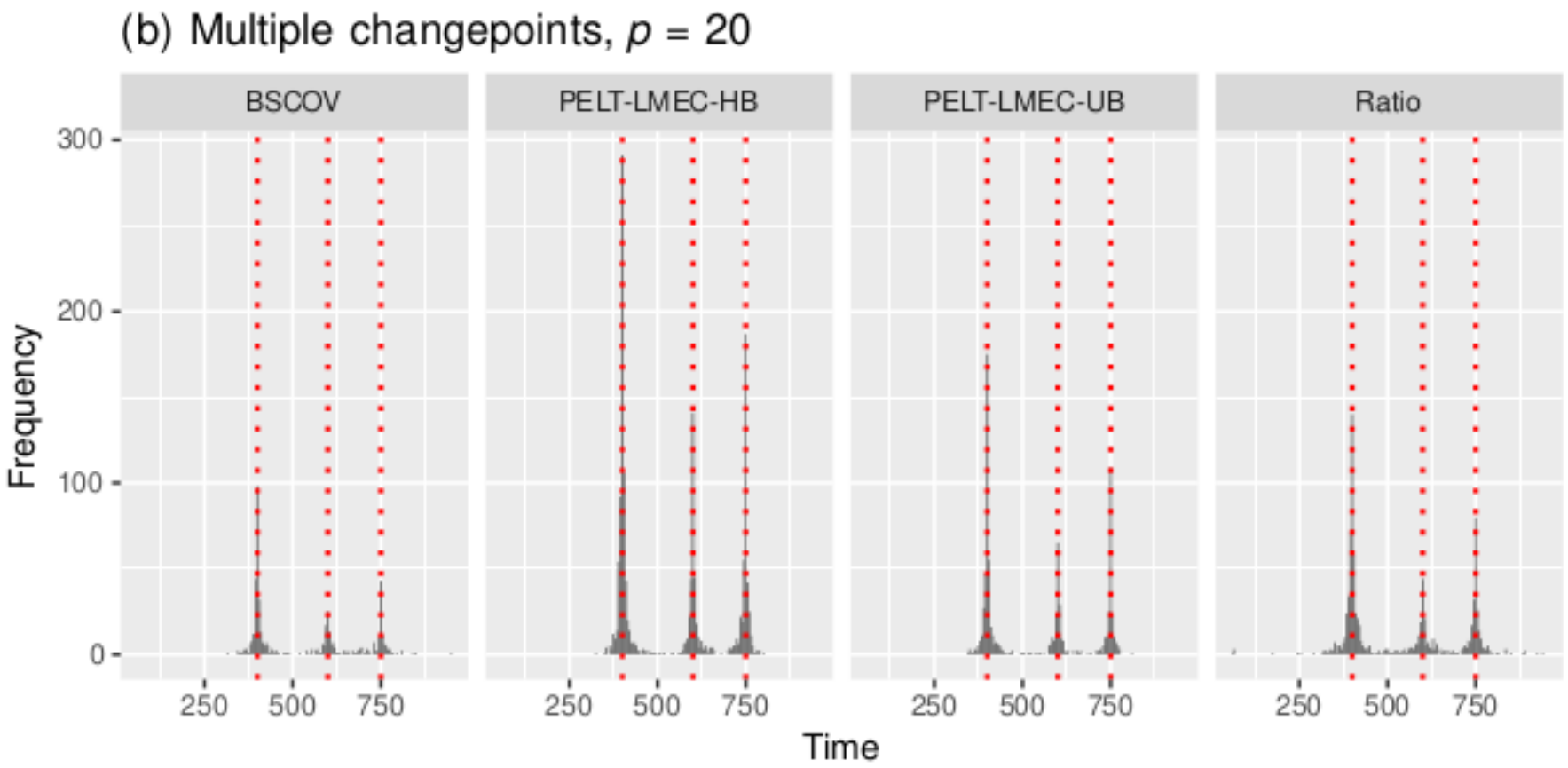
Compare to

- Ratio method (Ryan and Killick (2023)), no group structure
- Wild Sparsified Binary Segmentation (Li et al.(2023)), factor model

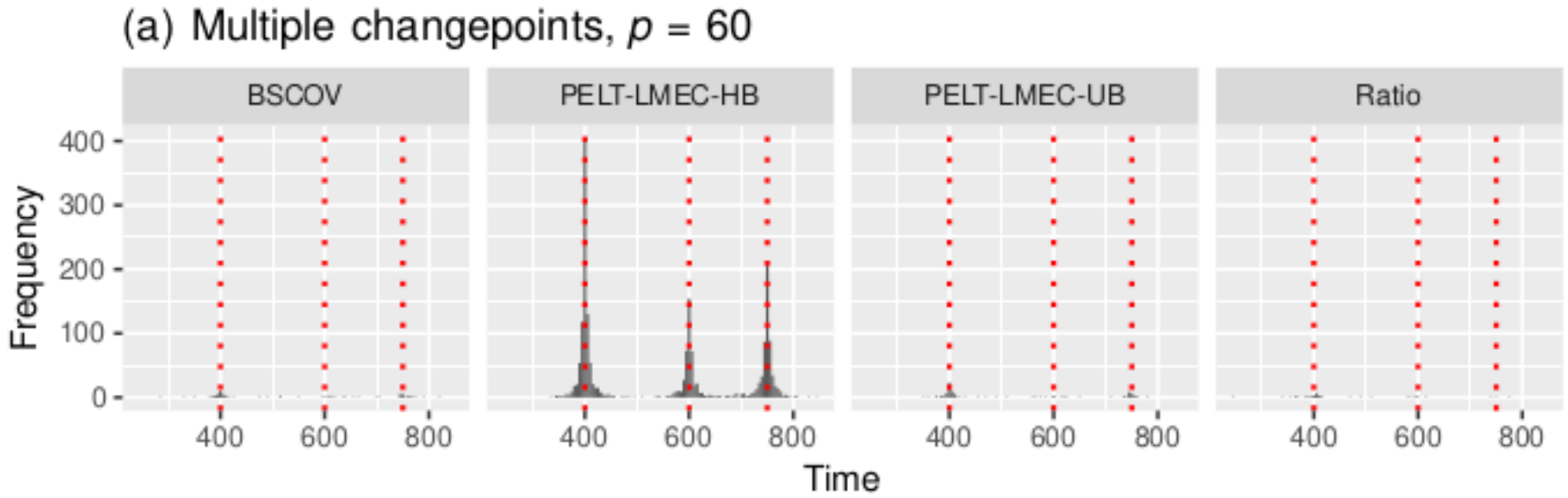
Single cpt p=20



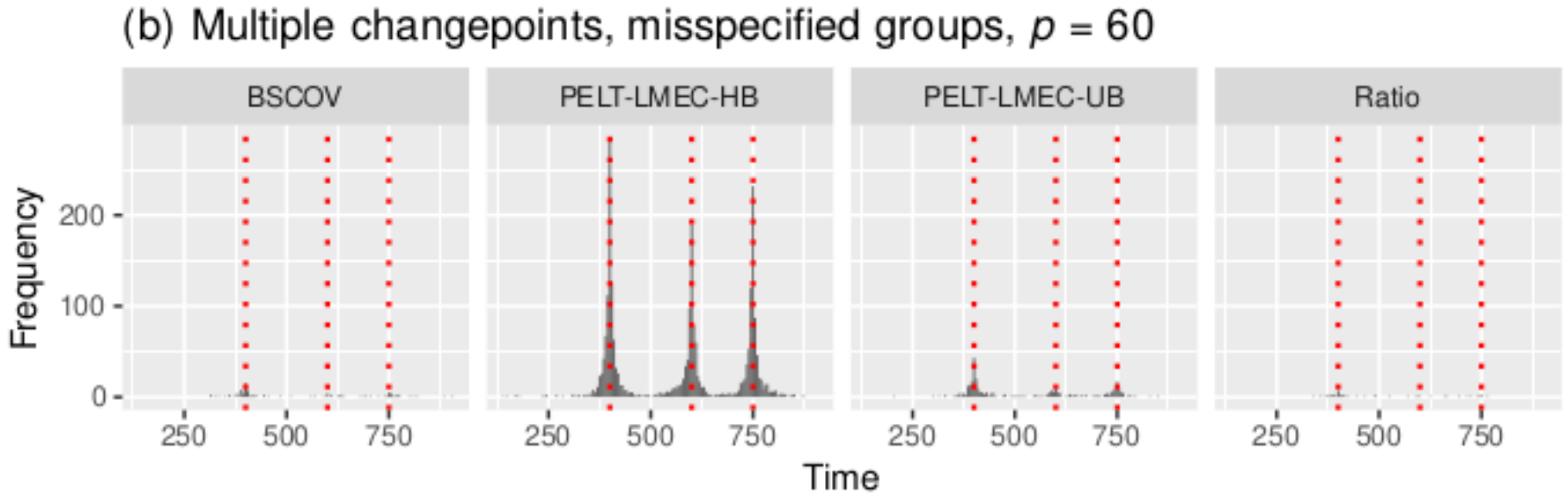
3 cpts p=20


(b) Multiple changepoints, p = 20
BSCOV
PELT-LMEC-HB
PELT-LMEC-UB
Ratio
Frequency
300
200
100
0
250
500
750
Time

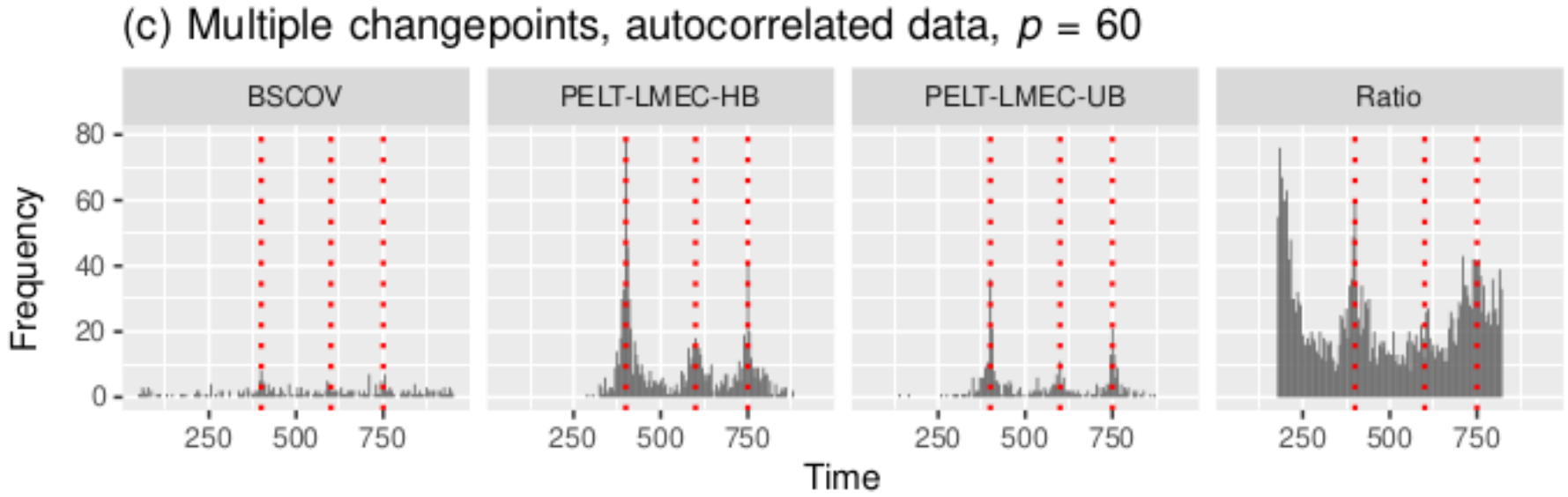
3 cpts p=60


(a) Multiple changepoints, p = 60
BSCOV
PELT-LMEC-HB
PELT-LMEC-UB
Ratio
Frequency
400
300
200
100
0
400
600
800
Time

Misspecified groups

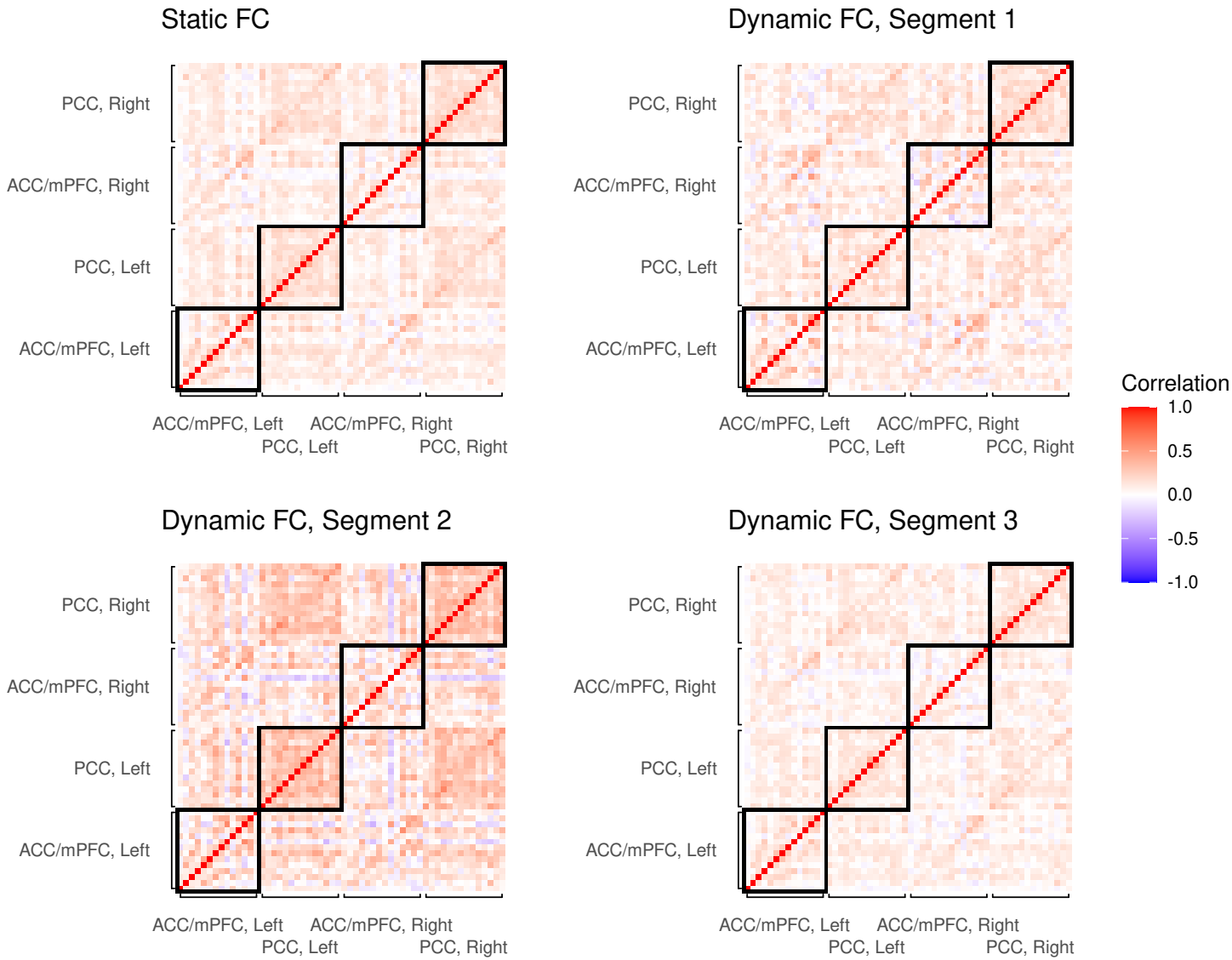


Autocorrelation

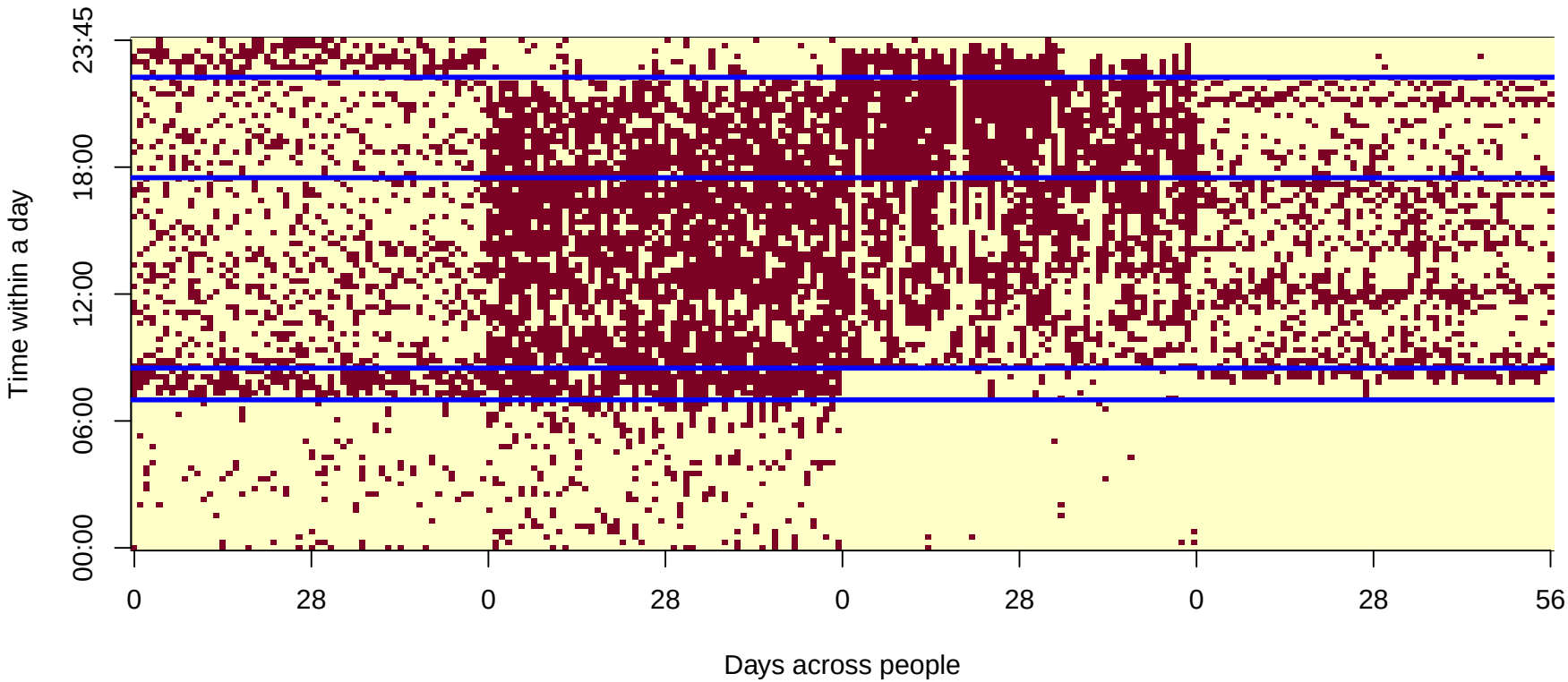


Summary of sims

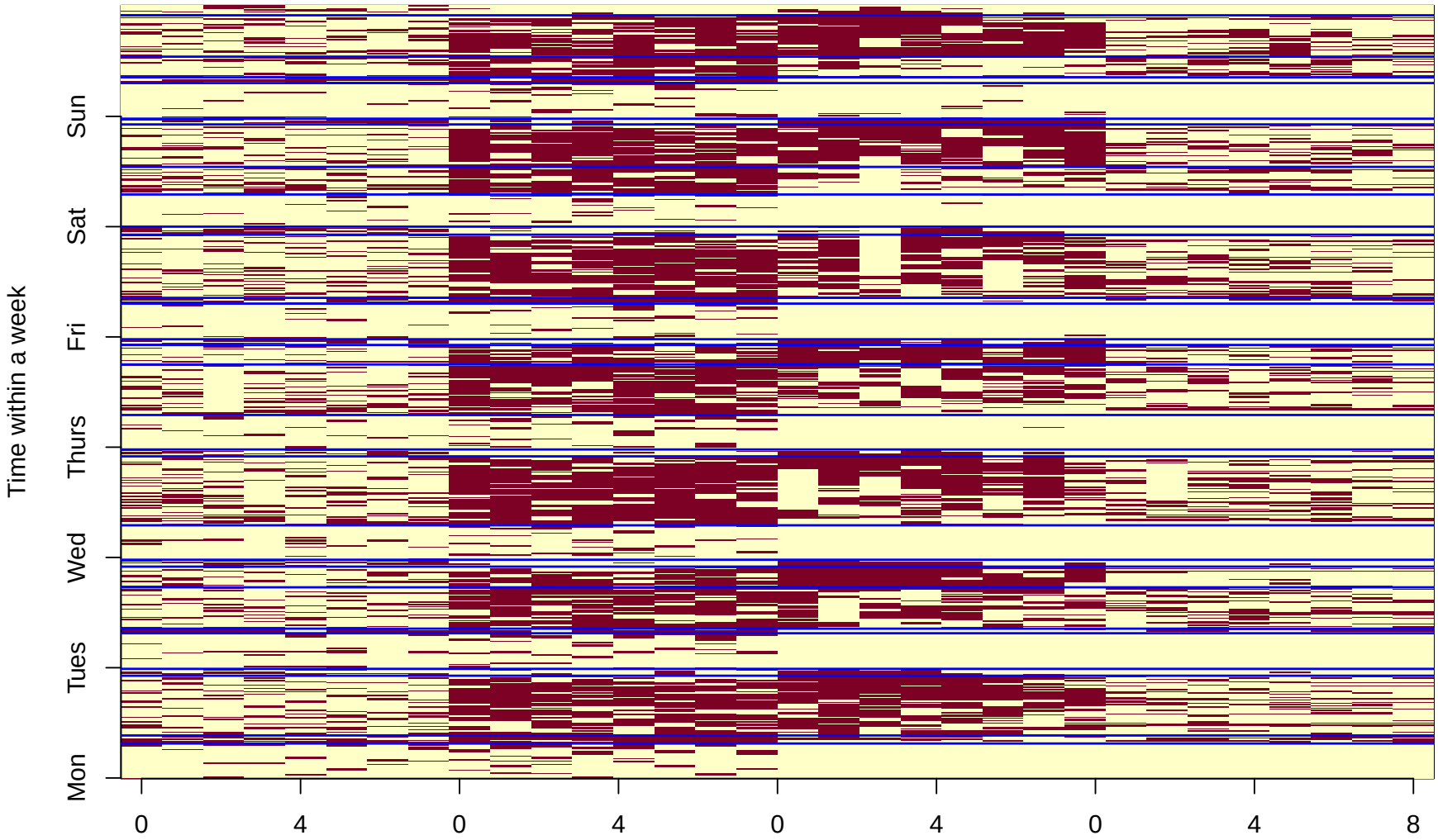
- LMEC-HB best overall
- Mild misspecification doesn't affect cpt detection
- No accounting for autocorrelation is bad


Static FC
Dynamic FC, Segment 1
Dynamic FC, Segment 2
Dynamic FC, Segment 3
PCC, Right
ACC/mPFC, Right
PCC, Left
ACC/mPFC, Left
ACC/mPFC, Left
ACC/mPFC, Right
PCC, Left
PCC, Right
Correlation
1.0
0.5
0.0
-0.5
-1.0

Howz


Time within a day
00:00
06:00
12:00
18:00
23:45
0
28
0
28
0
28
0
28
56
Days across people

Howz


Time within a week
Sun
Sat
Fri
Thurs
Wed
Tues
Mon
0
4
0
4
0
4
0
4
8
Weeks across people

Summary

- We all love changepoint research but in potentially different ways
- I'm excited to hear about the wide range of activity in Paris
- Presented a generalised random effects model with changepoints
- Showed its utility in two very different data settings
- I enjoy working with messy real life data . . .
- . . . as often it sparks my next research challenge.