Probabilistic modelling and functional data analysis of sleep structure

Zuzana Rošťáková

Institute of Measurement Science, Slovak Academy of Sciences

July 12, 2018

supervisor: Dr. Roman Rosipal

Sleep and daily behaviour

Sleep data

- 146 subjects, European sleep project SIESTA [Klösch et al., 2001]
- EEG recording of two consecutive nights spent in the sleep laboratory
- results of measures representing daily behaviour
  - age, gender
  - subjectively scored sleep and awakening quality
  - level of mood and drowsiness after awakening
  - physiological factors (blood pressure, pulse rate)
  - cognitive tests for working–memory, concentration, fine–motor activity
  - ...

Sleep structure and daily measures

Sleep structure

- total sleep time
- sleep efficiency
- sleep latency
- time spent awake
- time spent in each sleep stage
- . . .

$\Longleftrightarrow$

correlation coefficient,

linear regression,

prediction,

$\vdots$

Daily measures

- subjectively scored sleep quality
- level of drowsiness, mood
- fine–motor activity test
- working–memory test
- pulse rate
- blood pressure
- . . .

$\rightarrow$ the dynamics of the sleep process is not taken into account

Probabilistic sleep model (PSM)

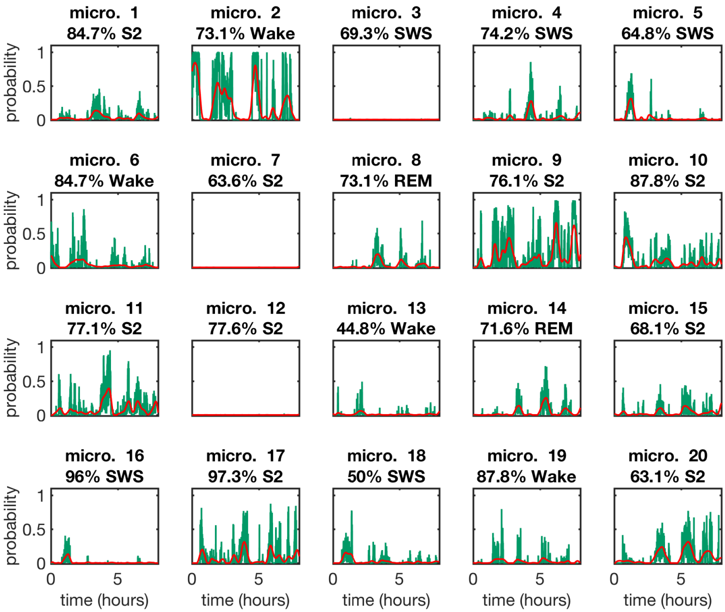
- [Lewandowski et al., 2012]
- 20 sleep microstates
- three–second long time segments of the EEG signal
- time segment $X_t \longrightarrow$ AR(10) representation: $$X_t = \sum_{i=1}^{10} a_i X_{t-i} + e_t$$
- Gaussian mixture model with latent variable Z representing membership to a sleep microstate $z \in \{1, \dots, 20\}$

$$p(a) = \sum_{z=1}^{20} p(z)p(a|z) = \sum_{z=1}^{20} \pi_z \mathcal{N}(a|\mu_z, \Sigma_z)$$

- sleep representation = posterior probability values

$$p(z|a) = \frac{\widehat{p}(z)\widehat{p}(a|z)}{\sum_{k=1}^{20} \widehat{p}(k)\widehat{p}(a|k)}$$

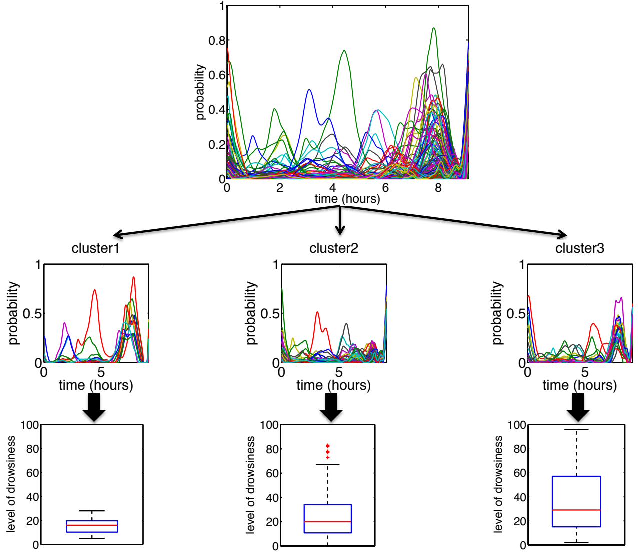
Probabilistic sleep model



Methodology

- detection of specific sleep profiles correlated with daily measures
    - cluster analysis of sleep probabilistic curves
    - time alignment of sleep probabilistic curves
- modelling individuality in the sleep pattern
    - Multilevel functional principal component analysis [Di et al., 2009]

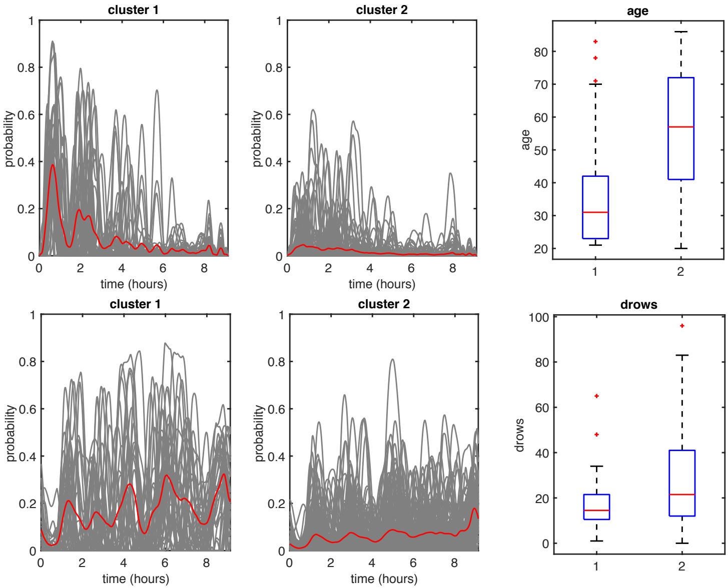
I. Cluster analysis of sleep probabilistic curves



Results

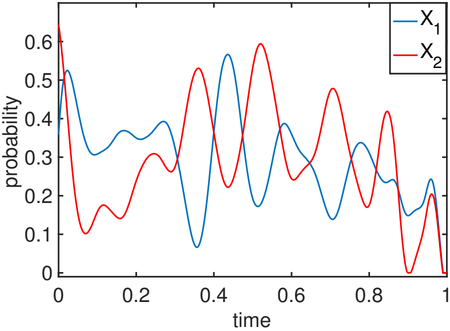
microstate 16
96% *SWS*

microstate 14
72% *REM*



Curves misalignment problem

- let X_1, X_2 to be two curves defined on the time interval $T = [0, 1]$



- the curves are **similar in shape**
- but their important features are **shifted in time**

$\rightarrow$ **the curves are misaligned in time**

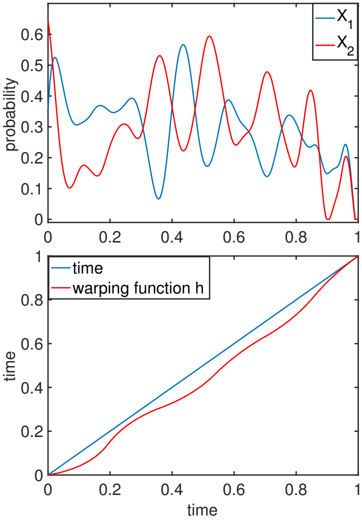
Curves alignment (registration, synchronisation)

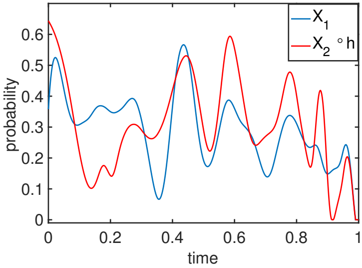
- two curves X_1, X_2 are defined over the time interval $T = [0, 1]$
- to align a pair of curves = to find a **warping function** $h^\star : T \to T$

$$h^\star \in \mathrm{argmin}_h \int_T (X_1(t) - (X_2 \circ h)(t))^2 \, dt$$

  - $h^\star$ is strictly increasing,
  - $h^\star(0) = 0, \quad h^\star(1) = 1,$
  - $\int_T (h^\star(t) - t)^2 dt < \lambda,$
  - ...

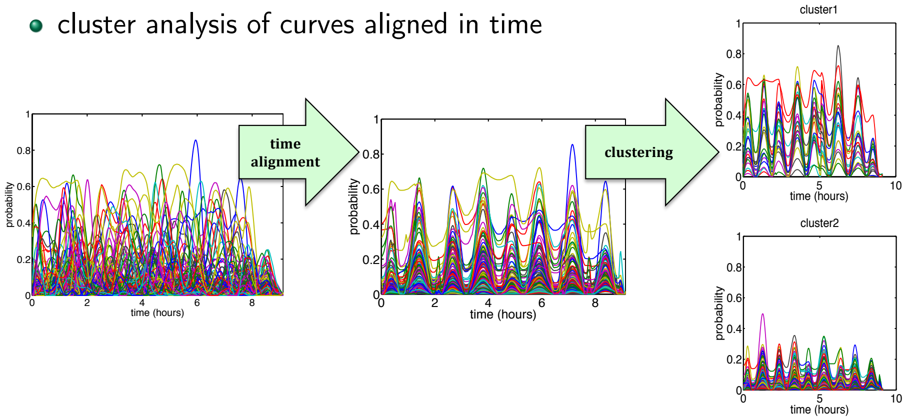
Example of the curves alignment



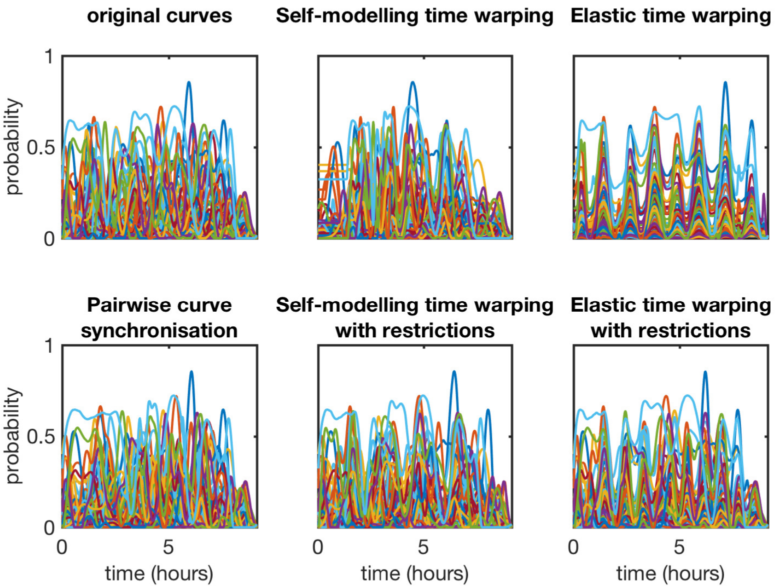


Curves alignment preceding the clustering step

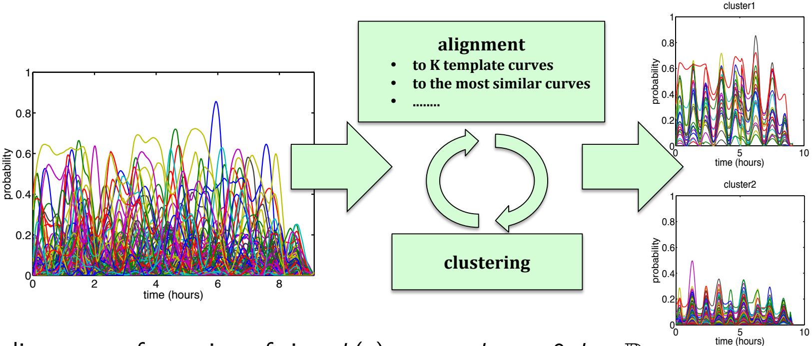
- curves are aligned as the whole dataset
  - Self–modelling time warping (SMTW) [Gervini and Gasser, 2004],
  - Pairwise curve synchronisation (PCS) [Müller and Tang, 2008],
  - Elastic time warping (ETW) [Tucker et al., 2013]
- cluster analysis of curves aligned in time



Curves alignment preceding the clustering step



Simultaneous curves alignment and clustering



- linear transformation of time $h(t) = at + b, a > 0, b \in \mathbb{R}$ [Gaffney and Smyth, 2005, Sangalli et al., 2010]

  $\Rightarrow$ **the assumption of common start and end point is violated**

- truncated Pairwise curve synchronisation [Müller and Tang, 2008]

  $\Rightarrow$ **not satisfactory alignment**

2step approach

- initial clustering
  - distance matrix M based on the Dynamic time warping algorithm [Wang and Gasser, 1997, Montero and Vilar, 2014]
  - k–medoids algorithm applied to the matrix M

- k^{th} step: align curves in each cluster separately
- $(k+1)^{th}$ step: recompute matrix M for aligned curves, cluster aligned curves . . .

- stopping criteria:
  - maximal number of iterations; changes in cluster membership
  - average dissimilarity L within clusters is small

$$L = \frac{1}{N}\int_0^1 \left(X_i(h_i(t)) - \mu_{C_i}(t)\right)^2 dt, \qquad \mu_{C_i}(t) = \frac{1}{|C_i|}\sum_{j:X_j\in C_i}^{N} X_i(h_i(t))$$

Application to the sleep probabilistic curves

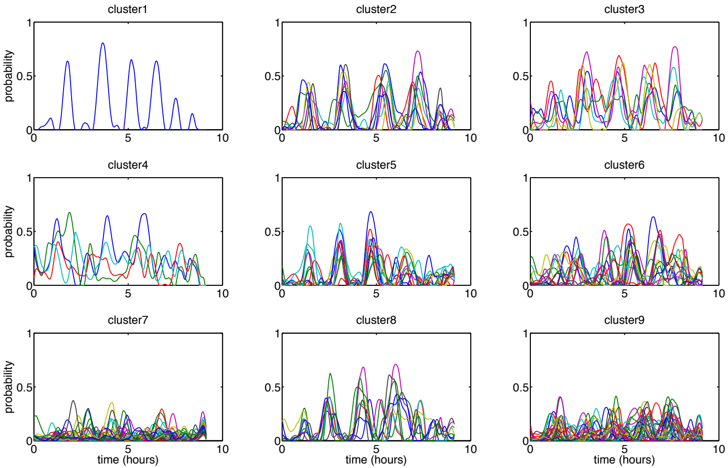


Figure: Microstate 8 (73% *REM*). Clustering of 146 in time misaligned sleep probabilistic curves into 9 clusters by the *k*–means algorithm.

Application to the sleep probabilistic curves

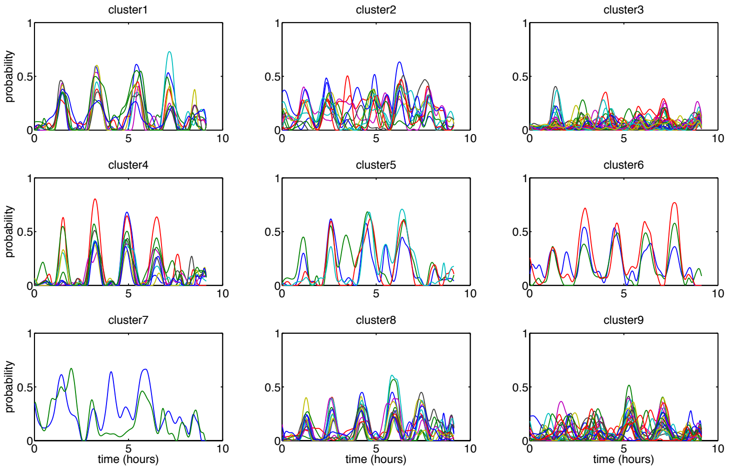


Figure: Microstate 8 (73% *REM*). Clustering of 146 sleep probabilistic curves into 9 clusters by the 2step approach with the SMTW algorithm used in the alignment step.

Application to the sleep probabilistic curves

Average silhouette	number of clust.	k–means	tPCS	2DTW– –SMTW	2DTW– –PCS	2DTW– –ETW
Microstate 16	8	0.56	0.57	**0.64**	0.61	0.47
Microstate 8	9	0.33	0.34	**0.48**	0.44	0.46
Microstate 14	3	0.53	0.54	**0.60**	0.10	**0.60**
Microstate 1	2	0.77	0.79	0.79	0.64	**0.80**
Microstate 6	3	0.62	0.63	**0.73**	0.71	0.53

L-criterion	number of clust.	k–means	tPCS	2DTW– –SMTW	2DTW– –PCS	2DTW– –ETW
Microstate 16	8	0.40	0.39	0.26	0.33	**0.19**
Microstate 8	9	0.79	0.77	0.54	0.59	**0.34**
Microstate 14	3	4.55	4.51	2.71	5.01	**1.77**
Microstate 1	2	4.07	4.05	3.37	4.34	**2.84**
Microstate 6	3	1.02	1.01	**0.78**	0.85	**0.73**

Results

- $\uparrow$ **Microstate 16** (96% *SWS*) $\Rightarrow \uparrow$ drowsiness
- $\uparrow$ **Microstate 13** (44% *Wake*) $\Rightarrow \downarrow$ cognitive tests

- $\uparrow$ micro. similar to **Wake**, **S1** $\Rightarrow$ people above 60 years of age
- $\uparrow$ micro. similar to **S2**, **SWS** $\Rightarrow$ people under 40 years of age

- $\uparrow$ micro. similar to **REM** $\Rightarrow \uparrow$ mood, $\uparrow$ drive, $\downarrow$ drowsiness
  $\Rightarrow \uparrow$ subjectively scored sleep quality

- $\uparrow$ micro. similar to **Wake** $\Rightarrow \downarrow$ mood, $\downarrow$ drive, $\uparrow$ drowsiness
  $\Rightarrow \downarrow$ subjectively scored sleep quality
  - **Microstate 19** (88% *Wake*) $\Rightarrow$ visible after the alignment
  - **Microstate 6** (85% *Wake*) $\Rightarrow$ the effect is **lost** after the alignment

- ...

Subject–specific sleep profiles – open problem

- individual pattern in the subjects' EEG signal
  - [Finelli et al., 2001, De Gennaro et al., 2005, De Gennaro et al., 2008, Lewandowski et al., 2013],...
  - PSM is based on the EEG signal
    $\rightarrow$ the individual pattern is inherited also in the sleep probabilistic curves
- estimation of the subject–specific profiles
  - need for multiple observations (> 2)
  - Multilevel functional principal component analysis [Di et al., 2009]

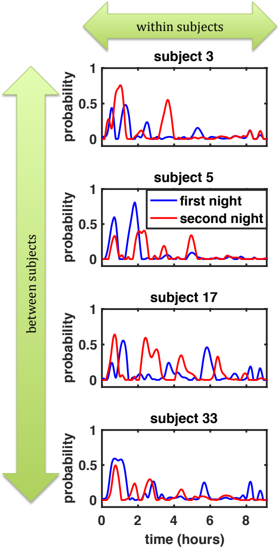
Multilevel functional principal component analysis

- MFPCA, [Di et al., 2009]

- functional data observed at multiple visits
- assumptions
  - **balanced design:** the same number of observation for each subject
  - **ordered visits:** observations within subjects have natural order

  - our database
    $\rightarrow$ 146 subjects with 2 visits
    $\rightarrow$ first and second night

- two types of variability
  - between subjects
  - within subjects



Multilevel functional principal component analysis

$$X_{ij}(t) = \mu(t) + \eta_j(t) + Z_i(t) + W_{ij}(t)$$

- X_{ij} is the j^{th} sleep curve for the i^{th} subject, i=1,...,146; j = 1,2
- fixed effects
  - **overall mean** μ

$$\widehat{\mu}(t) = \bar{X}_{..}(t) = \frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J} X_{ij}(t), \qquad t \in T$$

  - the **visit–specific deviation** from the overall mean $\boldsymbol{\eta_j}$

$$\sum_{j=1}^{J} \eta_j(t) = 0, \quad t \in T$$

$$\widehat{\eta}_j(t) = \bar{X}_{.j}(t) - \bar{X}_{..}(t) = \frac{1}{I}\sum_{i=1}^{I} X_{ij}(t) - \bar{X}_{..}(t), \qquad j = 1, \ldots, J$$

Multilevel functional principal component analysis

$$X_{ij}(t) = \mu(t) + \eta_j(t) + Z_i(t) + W_{ij}(t)$$

- random effects
  - the subject–specific deviation from the visit–specific profile $\mathbf{Z_i}$
  - residual deviation from the subject– and visit–specific profile $\mathbf{W_{ij}}$
    - $\rightarrow$ uncorrelated stochastic processes
    - $\rightarrow$ $\mathrm{E}\,(Z_i(t)) = \mathrm{E}\,(W_{ij}(t)) = 0, \qquad t \in T$
    - $\rightarrow$ covariance functions $R_Z : T \times T \to \mathbb{R}, R_W : T \times T \to \mathbb{R}$

$$Z_i(t) = \sum_{l=1}^{\infty} \alpha_{il} \phi_l^{(1)}(t) \qquad W_{ij}(t) = \sum_{k=1}^{\infty} \beta_{ijk} \phi_k^{(2)}(t)$$

- level 1 functional principal components $\phi_k^{(1)}, k = 1, 2, \ldots$
  $\longrightarrow$ eigenfunctions of R_Z
- level 2 functional principal components $\phi_l^{(2)}, l = 1, 2, \ldots$
  $\longrightarrow$ eigenfunctions of R_W

Covariance function estimators

- covariance functions

$$\begin{aligned} R_T(s,t) &= Cov\left(X_{ij}(s), X_{ij}(t)\right) \\ R_Z(s,t) &= Cov\left(X_{ij}(s), X_{il}(t)\right), \qquad && j \neq l \\ R_W(s,t) &= R_T(s,t) - R_Z(s,t), && s,t \in T \end{aligned}$$

- estimators by the method of moments
  - $\widehat{R}_T(s,t) = \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{\left(X_{ij}(s) - \widehat{\mu}(s) - \widehat{\eta}_j(s)\right)\left(X_{ij}(t) - \widehat{\mu}(t) - \widehat{\eta}_j(t)\right)}{IJ}$
  - $\widehat{R}_Z(s,t) = \sum_{i=1}^{I} \sum_{j \neq l}^{J} \frac{\left(X_{ij}(s) - \widehat{\mu}(s) - \widehat{\eta}_j(s)\right)\left(X_{il}(t) - \widehat{\mu}(t) - \widehat{\eta}_l(t)\right)}{IJ(J-1)}$
  - $\widehat{R}_W(s,t) = \widehat{R}_T(s,t) - \widehat{R}_Z(s,t)$

Algorithm

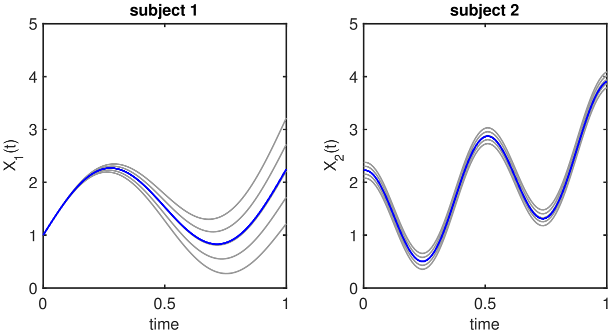
1. compute $\widehat{\mu}$ and $\widehat{\eta}_j$
2. compute $\widehat{R}_T$ and $\widehat{R}_Z$, $\widehat{R}_W = \widehat{R}_T - \widehat{R}_Z$
3. using $\widehat{R}_Z$ estimate the level 1 functional principal components
4. using $\widehat{R}_W$ estimate the level 2 functional principal components
5. appropriate number of principal components on both levels
6. estimate level 1 and level 2 principal component scores
   $\rightarrow$ BLUP, MCMC,...

Example

- 2 subjects, 10 curves

$$X_{1j}(t) = e^t + \sin(2\pi t) + \left(1 - \frac{j}{2}\right) t^2, \qquad j = 1, \ldots, 5; \qquad t \in [0, 1],$$

$$X_{2k}(t) = e^t + \cos(4\pi t) + 0.15\frac{k}{2}, \qquad k = 1, \ldots, 5.$$

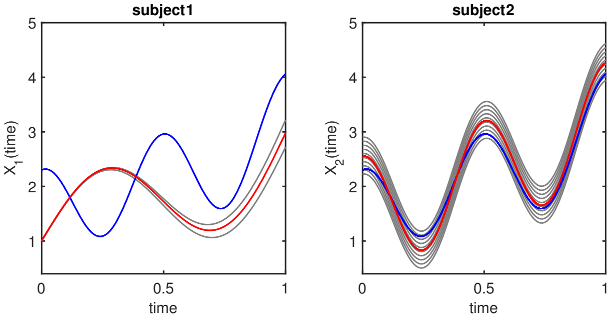


Example

- 2 subjects, 12 curves

$$X_{1j}(t) = e^t + \sin(2\pi t) + \left(1 - \frac{j}{2}\right) t^2, \qquad j = 1, 2; \qquad t \in [0, 1],$$

$$X_{2k}(t) = e^t + \cos(4\pi t) + 0.15\frac{k}{2}, \qquad k = 1, \ldots, 10.$$



- the assumptions of MFPCA are violated
  $\Rightarrow$ the method is not able to properly estimate subject–specific profiles

MFPCA for unbalanced data with unordered visits

- unordered visits $\rightarrow \eta_j \equiv 0$
- estimator for $\mu \rightarrow$ unweighted mean

$$\widehat{\mu}(t) = \frac{1}{I}\sum_{i=1}^{I}\frac{1}{J_i}\sum_{j=1}^{J_i} X_{ij}(t) = \frac{1}{I}\sum_{i=1}^{I}\overline{X_{i.}}(t)$$

- modified estimator $\widehat{R}_W$

$$\widehat{R}_W(s,t) = \frac{1}{2}\sum_{i=1}^{I}\sum_{j=1}^{J_i}\sum_{l\neq j}^{J_i} w_i\left(X_{ij}(s) - X_{il}(s)\right)\left(X_{ij}(t) - X_{il}(t)\right),$$

$$\sum_{i=1}^{I} w_i J_i(J_i - 1) = 1 \quad \text{and} \quad w_i \geq 0, \quad i = 1,\ldots,I$$

and $\widehat{R}_W$ minimises

$$\mathrm{E}\left(\|\widehat{R}_W - R_W\|^2\right)$$

MFPCA for unbalanced data with unordered visits

- covariance function estimators

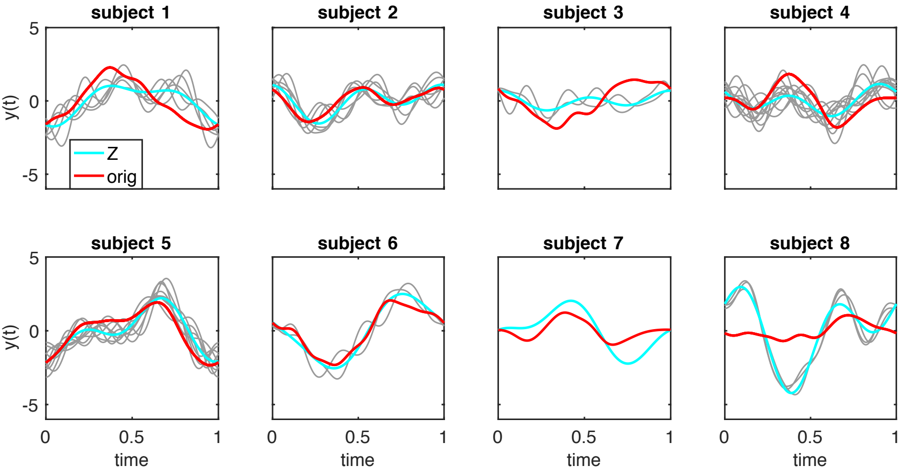
$$\widehat{R}_W(s,t) = \frac{1}{2}\sum_{i=1}^{I}\sum_{j=1}^{J_i}\sum_{l\neq j}^{J_i}\frac{1}{(N_1-I)J_i}\left(X_{ij}(s)-X_{il}(s)\right)\left(X_{ij}(t)-X_{il}(t)\right)$$

$$\widehat{R}_T(s,t) = \frac{1}{N_1}\sum_{i=1}^{I}\sum_{j=1}^{J_i}\left(X_{ij}(s)-\widehat{\mu}_{uw}(s)\right)\left(X_{ij}(t)-\widehat{\mu}_{uw}(t)\right)$$

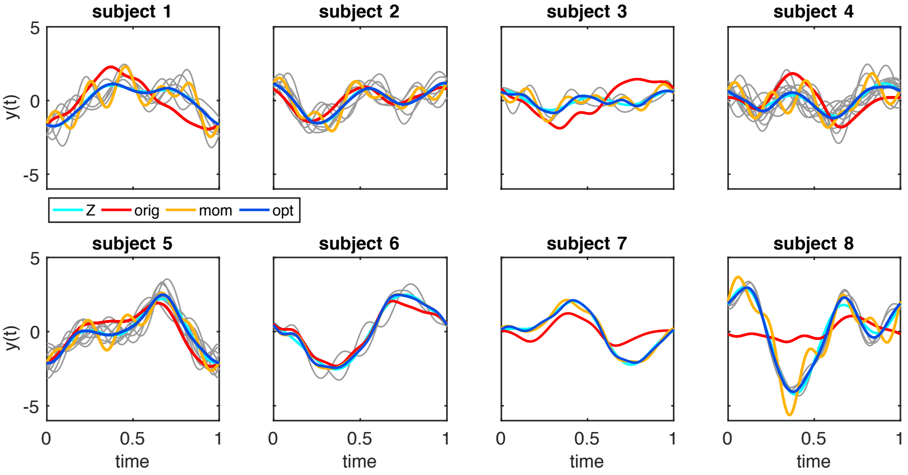
$$\widehat{R}_B(s,t) = \frac{I}{I-1}\left(\widehat{R}_T(s,t) - \left(1-\frac{2}{N_1}+\frac{1}{I^2}\sum_{i=1}^{I}\frac{1}{J_i}-\frac{L}{N_1}\frac{I-2}{I}\right)\widehat{R}_W(s,t)\right)$$

- $N_1 = \sum_{i=1}^{I} J_i$
- L = number of subjects with only one observation ($J_i = 1$)

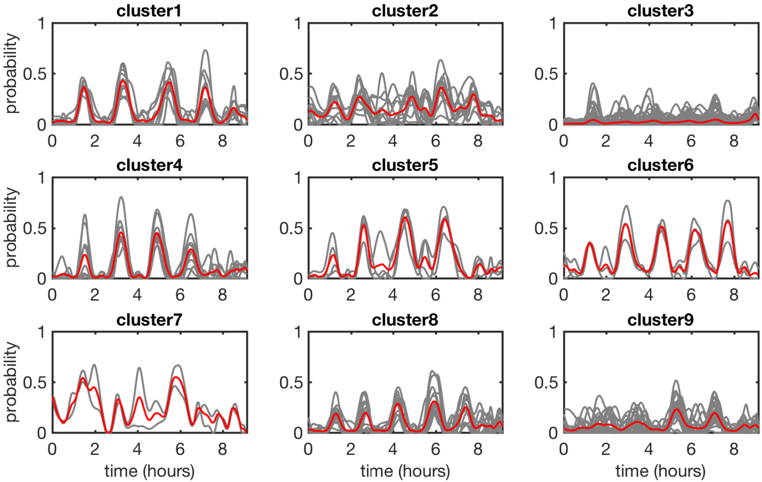
MFPCA for unbalanced data with unordered visits



MFPCA for unbalanced data with unordered visits



Application to the results of cluster analysis



Thank you for your attention.

References I

De Gennaro, L., Ferrara, M., Vecchio, F., Curcio, G., and Bertini, M. (2005).
An electroencephalographic fingerprint of human sleep.
NeuroImage, 26(1):114 – 122.

De Gennaro, L., Marzano, C., Fratello, F., Moroni, F., Pellicciari, M. C., Ferlazzo, F., Costa, S., Couyoumdjian, A., Curcio, G., Sforza, E., Malafosse, A., Finelli, L. A., Pasqualetti, P., Ferrara, M., Bertini, M., and Rossini, P. M. (2008).
The electroencephalographic fingerprint of sleep is genetically determined: A twin study.
Annals of Neurology, 64(4):455–460.

Di, C.-Z., Crainiceanu, C. M., Caffo, B. S., and Punjabi, N. M. (2009).
Multilevel functional principal component analysis.
The Annals of Applied Statistics, 3(1):458–488.

Finelli, L. A., Achermann, P., and Borbély, A. A. (2001).
Individual "fingerprints" in human sleep eeg topography.
Neuropsychopharmacology, 25:57–62.

Gaffney, S. J. and Smyth, P. (2005).
Joint probabilistic curve clustering and alignment.
In Saul, L. K., Weiss, Y., and Bottou, L., editors, *Advances in Neural Information Processing Systems 17*, pages 473–480. MIT Press.

Gervini, D. and Gasser, T. (2004).
Self-modeling warping functions.
Journal of the Royal Statistical Society. Series B, 66(4):959–971.

References II

Klösch, G., Kemp, B., Penzel, T., Schlögl, A., Rappelsberger, P., Trenker, E., Gruber, G., Zeitlhofer, J., Saletu, B., Herrmann, W., Himanen, S., Kunz, D., Barbanoj, M., Röschke, J., Varri, A., and Dorffner, G. (2001).
The SIESTA project polygraphic and clinical database.
Medicine and Biology Magazine, 20(3):51–57.

Lewandowski, A., Rosipal, R., and Dorffner, G. (2012).
Extracting more information from EEG recordings for a better description of sleep.
Computer methods and programs in biomedicine, 108(3):961–972.

Lewandowski, A., Rosipal, R., and Dorffner, G. (2013).
On the individuality of sleep eeg spectra.
Journal of Psychophysiology, 27(3):105–112.

Montero, P. and Vilar, J. A. (2014).
TSclust: An R package for time series clustering.
Journal of Statistical Software, 62(1):1–43.

Müller, H. G. and Tang, R. (2008).
Pairwise curve synchronisation for functional data.
Biometrika, 95(4):875–889.

Rosipal, R., Lewandowski, A., and Dorffner, G. (2013).
In search of objective components for sleep quality indexing in normal sleep.
Biological Psychology, 94(1):210–220.

References III

Rošťáková, Z., Dorffner, G., Aydemir, Ö., and Rosipal, R. (2017).
Estimation of sleep quality by using microstructure profiles.
In ten Teije, A., Popow, C., Holmes, J. H., and Sacchi, L., editors, *Artificial Intelligence in Medicine*, pages 105–115. Springer International Publishing.

Rošťáková, Z. and Rosipal, R. (2018).
Time alignment as a necessary step in the analysis of sleep probabilistic curves.
Measurement Science Review, 18(1):1–6.

Sangalli, L. M., Secchi, P., Vantini, S., and Vitelli, V. (2010).
k–mean alignment for curve clustering.
Computational Statistics and Data Analysis, 54:1219 – 1233.

Tucker, J. D., Wu, W., and Srivastava, A. (2013).
Generative models for functional data using phase and amplitude separation.
Computational Statistics and Data Analysis, 61:50–66.

Wang, K. and Gasser, T. (1997).
Alignment of curves by dynamic time warping.
The Annals of Statistics, 25(3):1251–1276.