ADAPTIVE SPECTRAL ANALYSIS OF REPLICATED NONSTATIONARY TIME SERIES

ROBERT KRAFTY
Department of Biostatistics
University of Pittsburgh

Joint work with:
Scott Bruce
Department of Statistical Science, Temple University

Daniel Buysse & Martica Hall
Department of Psychiatry, University of Pittsburgh

Overview

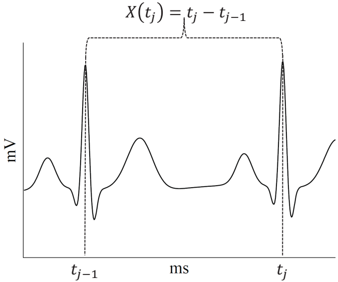
1. Motivating Application: AgeWise
2. Methodological Background
3. CABS: Conditional Adaptive Bayesian Spectral Analysis
4. Simulated Example
5. Analysis of AgeWise
6. Concluding Remarks

AgeWise Study

AgeWise Cargeiver Study: Study of older adults who are the primary cargiver for their spouse with dementia.

- 43% of older adults have problems sleeping.
- Poor sleep is connected to impaired daytime functioning & many negative clinical outcomes.
- $N = 30$ men and women aged 65–89.
- Two types of data:
  1. Self-report predictor:
     PSQI: Measure of self-reported sleep disturbances and effects on life during a one-month period.
  2. Polysomnography:
     Continuous recording of electrophysiological time series during sleep.

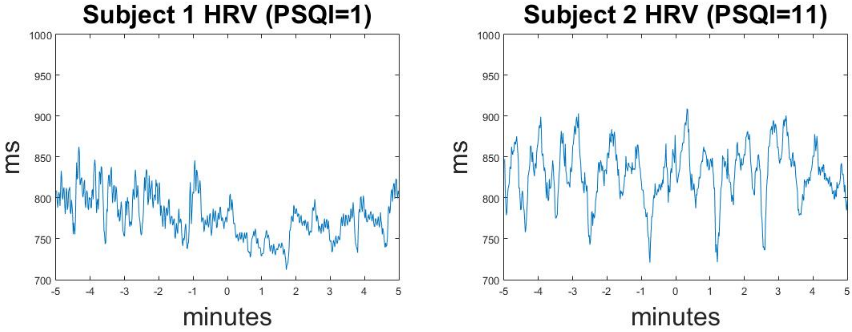
PSG: Polysomnography



HRV During Sleep

- Heart rate variability (HRV): Elapsed time between consecutive heart beats.
  - HFnu: percent power from high-frequencies (.15-.40 Hz)
  - HFnu: measure of parasympathetic nervous system activity
  - HFnu: inverse measure of stress and arousal.
- Body cycles through different physiological states during sleep: REM and NREM.
  - Essential to rejuvenation of nervous, muscular, endocrine & skeletal systems.
  - Healthy people: NREM dominated by parasympathetic activity, REM is not.
- Goal: Understand how self-reported sleep quality (PSQI) predicts dynamics of physiology (HRV), which can inform treatment.

Data From Two Subjects



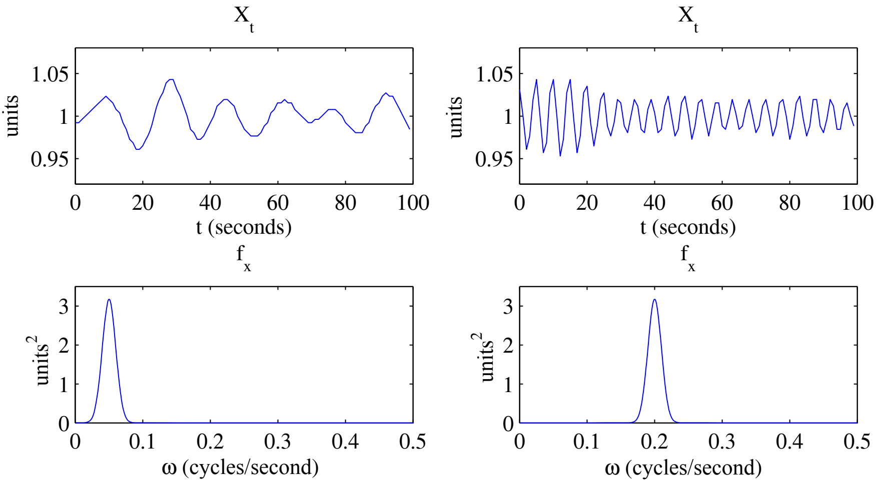
Stationary Time Series

- Consider a stationary time series X_t
- Cramér Representation:

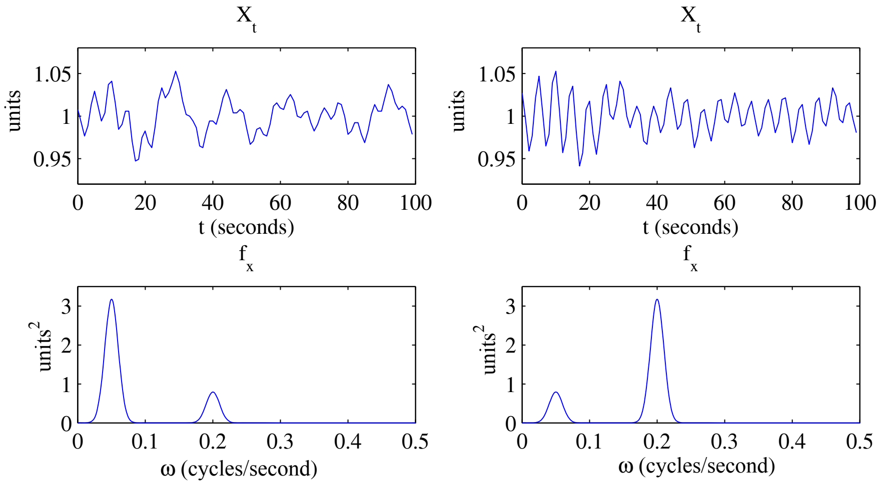
$$X_t = \int_{-1/2}^{1/2} A(\nu) \exp(2\pi\nu t) dZ(\nu).$$

- Power spectrum: $f(\nu) = |A(\nu)|^2$.
- Type of frequency ANOVA
- Variability of X_t attributable to periodic signals at frequency $\omega \in \mathbb{R}$.
- $\text{var}(x_t) = \int_{-1/2}^{1/2} f(\omega) d\omega$

TWO SIMULATED EXAMPLES



Two More Simulated Examples



Periodogram

- Periodogram from $\mathbf{X} = (X_1, \ldots, X_T)'$:

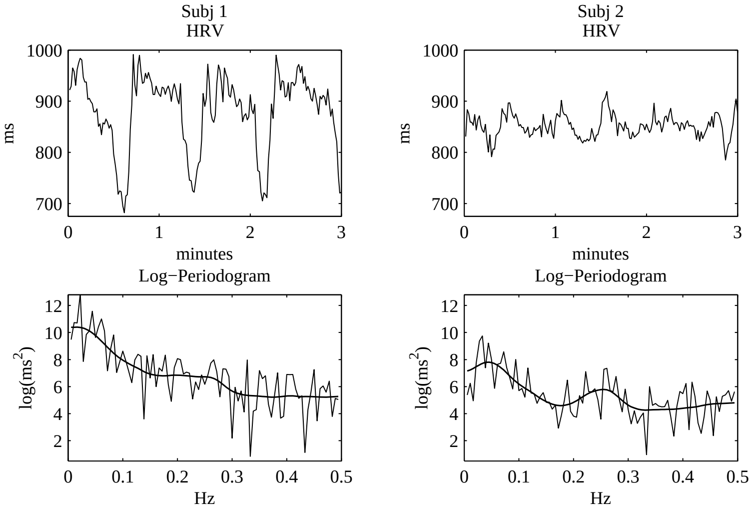
$$Y(\nu) = \frac{1}{T} \left| \sum_{t=1}^{T} X_t \exp(-2\pi i \nu t) \right|^2$$

- Unbiased but noisy estimates of $f(\nu)$.
- Approximately distributed as scaled χ^2 to provide the Whittle likelihood:

$$p(\mathbf{x}|\mathbf{f}) \approx (2\pi)^{-n/2} \prod_{k=0}^{n} \exp\left\{ -\frac{1}{2} [\log f(\nu_k) + Y(\nu_k)/f(\nu_k)] \right\}$$

- $\nu_k = k/T$, $k = 1, \ldots, n = \lfloor T/2 \rfloor$.

Periodogram Examples



Smoothing

- Periodogram can be smoothed to obtain a consistent estimate.
- One approach - Bayesian penalized linear spline:

$$
\begin{aligned}
\log f(\nu) &= \alpha + \sum_{k=1}^{K} \beta_k \cos(2\pi k \nu) \\
\beta = (\beta_1, \ldots, \beta_n)' &\sim N\left(0, \tau^2 D_\beta\right) \\
D_\beta &= \mathtt{diag}\left[(2\pi 1)^{-2}, \ldots, (2\pi K)^{-2}\right] \\
\tau &\sim \text{half-t}
\end{aligned}
$$

- Note: $\int_{-1/2}^{1/2} \left\{[\log f(\nu)]'\right\}^2 d\nu \sim \tau^2 \chi^2_K$

Nonstationary Time Series

Time-varying Cramér Representation (Priestly 1969; Dahlhaus 1997)

$$X_t = \int_{-1/2}^{1/2} A(t/T, \nu) \exp(2\pi \nu t) dZ(\nu).$$

Time-Varying Power spectrum: $f(u, \nu) = |A(u, \nu)|^2$.

- Approaches to estimation:
  - Piecewise Stationary: Adak 1998, Ombao et al. 2003, Davis et al. 2006
  - Continuous: Dahlhaus 1997, Guo et al. 2003
  - Abrupt & Continuous: Yang et al. 2016, Rosen et al. 2012

Replicated Nonstationary Time Series

- In practice: ad hoc.
- In statistics literature: very few methods for replicated nonstationary time series in conjunction with scalar predictors.
- Qin et al. 2009 and Fiecas & Ombao 2016
  - Continuous temporal changes and continuous predictor effects.
- Goal: Method that can capture both smooth and abrupt changes and conduct inference on predictor effect and dynamics.

Conidtional Locally Stationary Time Series

- Observe $\{X_{\ell 1}, \ldots, X_{\ell T}\}$ and w_ℓ from $\ell = 1, \ldots, N$ subjects.

$$X_{\ell t} = \int_{-1/2}^{1/2} A(t/T, w_\ell, \nu) \exp(2\pi\nu t) dZ_\ell(\nu).$$

- Covariate-indexed time-varying spectrum

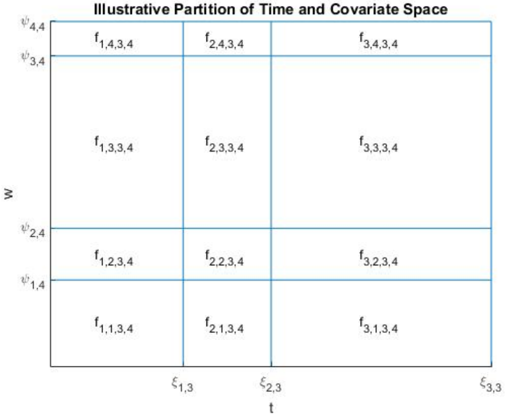
$$f(u, w, \nu) = |A(u, w, \nu)|^2 .$$

Overview of Proposed Approach

- Adaptively partition predictor-time grid into random stationary blocks.
- Bayesian penalized spline for local spectra within blocks.
- Number and location of partitions are random and adaptively estimated from data.
- Reversible jump Markov chain and Hamiltonian Monte Carlo techniques to sample from posterior.
- Inference averaged over distribution of partitions.

Partitioning

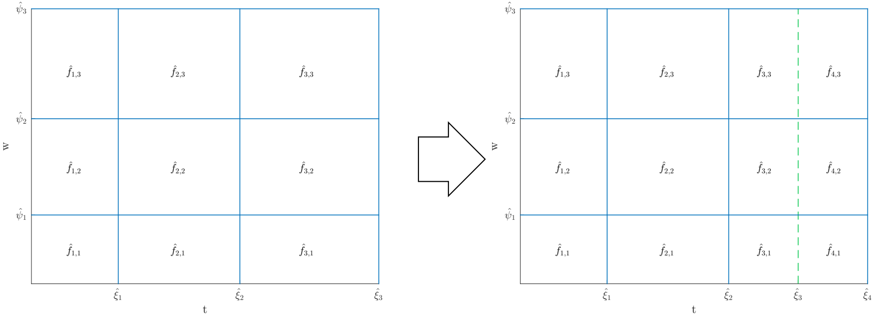
- m time segments $\xi_m = (\xi_{0,m}, \ldots, \xi_{m,m})'$
- p predictor segments $\psi_p = (\psi_{0,p}, \ldots, \psi_{p,p})'$
- m and p have uninform priors
- ξ_m and ψ_p have uniform priors conditional on m and p



Sampling Scheme

- Iterations consist of 2 types of moves:
  1. Between-model moves: the number of partition points changes.
  2. Within-model moves: number of partitions is constant but location changes.

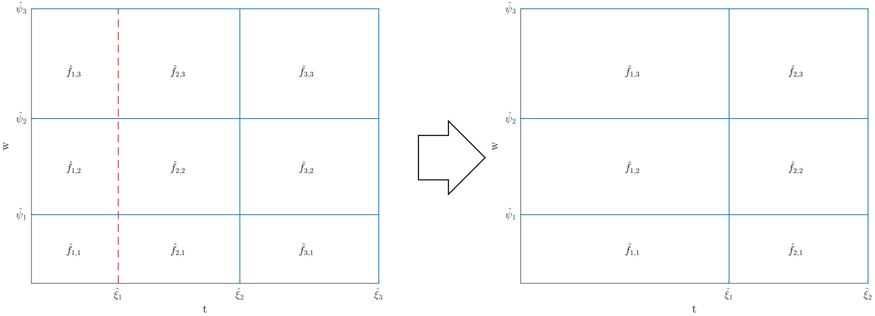
Birth: a new partition point is proposed.



Sampling Scheme

- Iterations consist of 2 types of moves:
  1. Between-model moves: the number of partition points changes.
  2. Within-model moves: number of partitions is constant but location changes.

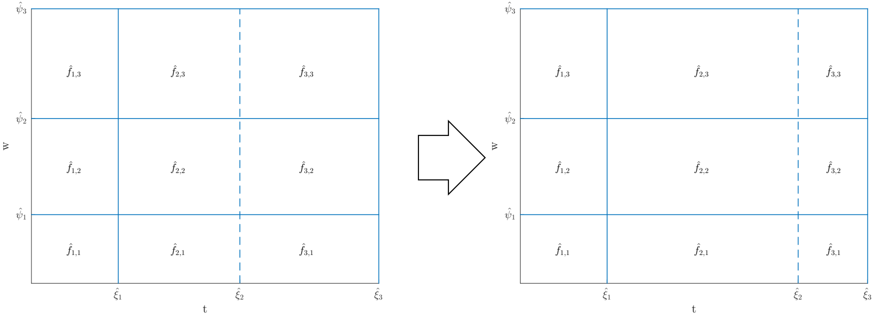
Death: a partition point is proposed for removal.



Sampling Scheme

- Iterations consist of 2 types of moves:
  1. Between-model moves: the number of partition points changes.
  2. Within-model moves: number of partitions is constant but location changes.

Randomly select a partition point and propose a new location.



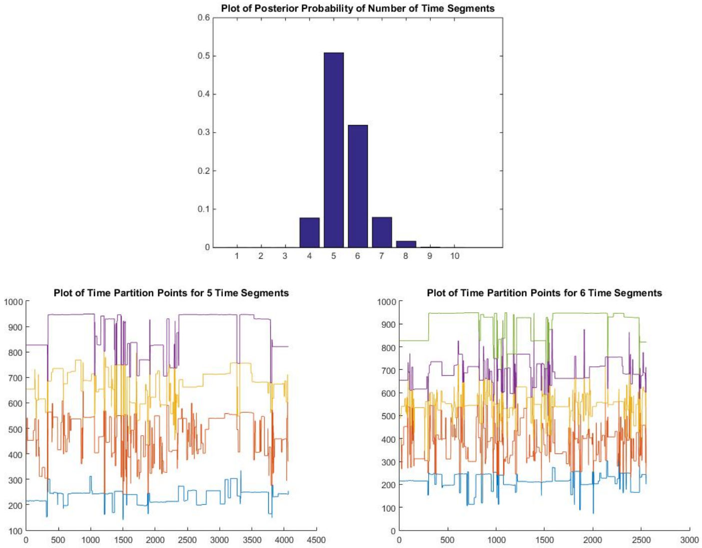
Sampling Scheme

- Iterations consist of 2 types of moves:
  1. Between-model moves: the number of partition points changes.
     - Birth: a new partition point is proposed.
     - Death: a partition point is proposed for removal.
  2. Within-model moves: number of partitions is constant but location changes.
     - Randomly select a partition point and propose a new location.
- Affected spline coefficients and smoothing parameters are drawn after each move using Hamiltonian Monte Carlo.
- Move parameters are jointly accepted or rejected.
- For time segments, then for predictor.

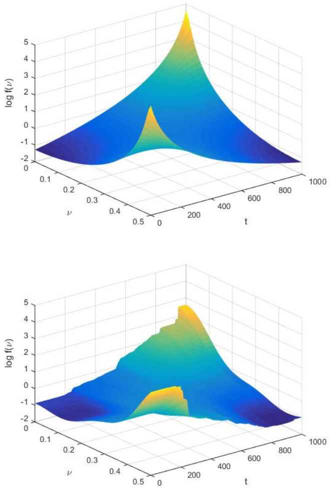
Simulated Smooth Example

- $x_{\ell t} = \phi_{\ell t} x_{\ell t-1} + \epsilon_{\ell t}$
- $\phi_{\ell t} = \begin{cases} -0.5 + t/1000 & \text{for } 0 \leq w_\ell \leq 0.5 \\ -0.9 + 9t/5000 & \text{for } 0.5 < w_\ell \leq 1 \end{cases}$
- $\ell = 1, \cdots, L = 8$ subjects
- $w_\ell = (\ell - 1)/(L - 1)$
- $\epsilon_{\ell t} \overset{\text{iid}}{\sim} N(0, 1)$.

Time Point Posterior Distribution



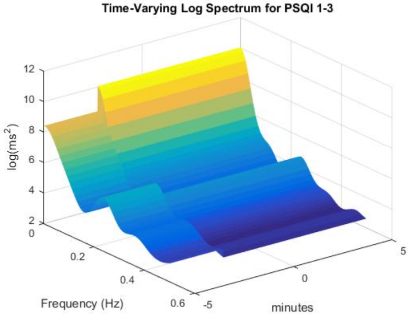
Estimated and True Conditional Time-Varying Spectra

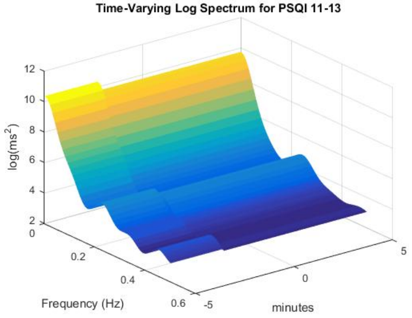


AgeWise Analysis

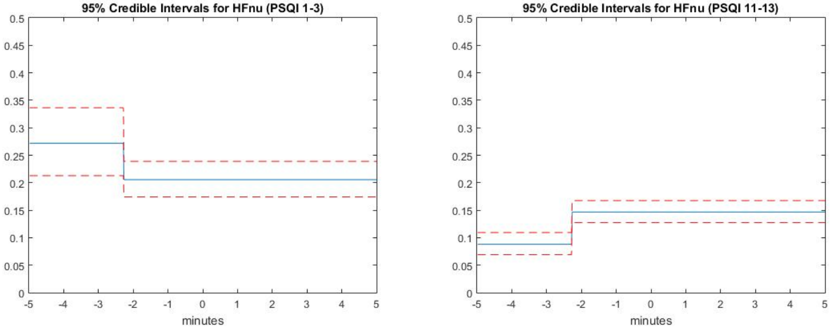
- Two clear temporal segments:
  - Pr$(m = 2) = 99\%$
  - Change point 2.3 minutes before start of REM with Pr$\left(\xi_{1,2} = -2.3 \mid m = 2\right) = 98\%$.
  - The heart leads the brain! (Jurysta et al., 2003)
- Four clear PSQI segments with probability 95%:
  - PSQI 1–3: excellent sleepers
  - PSQI 4–6: good sleepers
  - PSQI 7–10: fair sleeper
  - PSQI 11–13: poor sleepers

Extreme Groups





HFnu


95% Credible Intervals for HFnu (PSQI 1-3)
minutes
95% Credible Intervals for HFnu (PSQI 11-13)
minutes

Conclusion

- Discussed an approach to the automated predictor-dependent time-frequency analysis of replicated time series.
- Can capture abrupt and smooth changes.
- Can conduct inference on associations that accounts for partitioning.
- Paper in press and available on Biometrics' website.
- Matlab program on GitHub: `github.com/sbruce23/CABS`
- Current work:
  - How to deal with multiple covariates?
  - How to analyze multivariate time series?

THANK YOU!

Conclusion

- Discussed an approach to the automated predictor-dependent time-frequency analysis of replicated time series.
- Can capture abrupt and smooth changes.
- Can conduct inference on associations that accounts for partitioning.
- Paper in press and available on Biometrics' website.
- Matlab program on GitHub: `github.com/sbruce23/CABS`
- Current work:
  - How to deal with multiple covariates?
  - How to analyze multivariate time series?

THANK YOU!