Bayesian Nonparametric Spectral Analysis of Multivariate Time Series

Yixuan Liu [1], Claudia Kirch[2], Kate Lee[1], Renate Meyer[1]

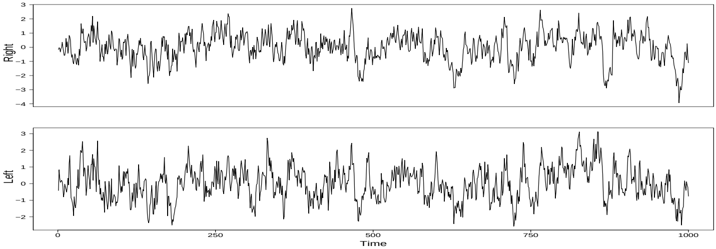
[1]Dept of Statistics, University of Auckland, New Zealand

[2]Dept of Mathematics, Otto-von-Guericke University, Magdeburg, Germany

Biometrics in the Bay of Islands, 27/11–01/12/2023

Bivariate EEG Time Series

Rat A



Source: Quiroga et al. (2002), Phys. Rev. E, 65, 041903

- intra-temporal variability
- inter-temporal interactions
- Not independent nor identically distributed.
- Cyclic behaviour, but not deterministic.
  - Different heights of peaks
  - Different lengths of periods

Temporal Dependance Structure

Assumptions:

- observations are not iid but stationary
  (same stochastic behaviour everywhere)
- dependence decreases over time

Usually, this is no problem for any point estimates (e.g. sample average for expectation)

BUT: The quantification of uncertainty depends critically on this dependence structure!

Temporal Dependance Structure

Assumptions:

- observations are not iid but stationary
  (same stochastic behaviour everywhere)
- dependence decreases over time

Usually, this is no problem for any point estimates (e.g. sample average for expectation)

BUT: The quantification of uncertainty depends critically on this dependence structure!

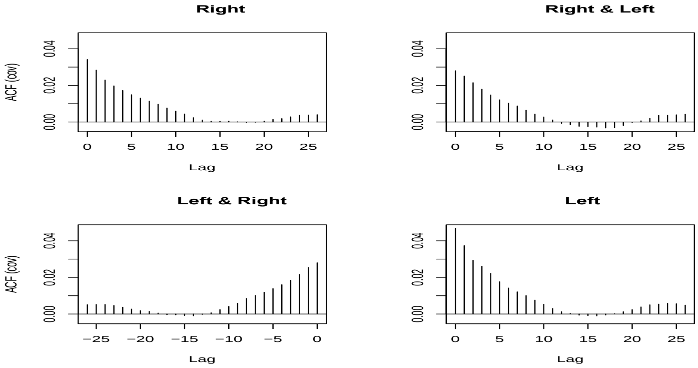
How to quantify uncertainty without parametric assumptions?

Multivariate Time Series

Let $\mathbf{Z}_t = (Z_t^{(1)}, \ldots, Z_t^{(d)})^T$ d-dim, mean-centered, stationary

$\mathbf{\Gamma}(h) = \text{Cov}(\mathbf{Z}_{t+h}, \mathbf{Z}_t)$ independent of t for all lags h,

$\mathbf{\Gamma}(h)$ matrix-valued autocovariance function, positive definite



ACF of EEG time series for Rat A.

Characterisation by Spectral Density Matrix

Herglotz Lemma

Autocovariance Function $\longleftrightarrow$ Spectral Density Matrix

$$\mathbf{\Gamma}(h) = \int_0^{2\pi} e^{ih\lambda}\mathbf{f}(\lambda)d\lambda \qquad \longleftrightarrow \qquad \mathbf{f}(\omega) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\mathbf{\Gamma}(k)e^{-ik\omega}$$

Characterisation by Spectral Density Matrix

Herglotz Lemma

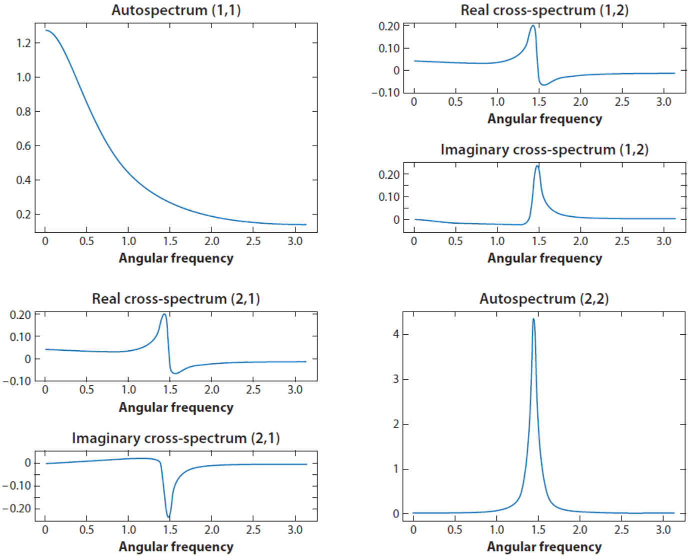
Autocovariance Function $\longleftrightarrow$ Spectral Density Matrix

$$\mathbf{\Gamma}(h) = \int_0^{2\pi} e^{ih\lambda}\mathbf{f}(\lambda)d\lambda \qquad \longleftrightarrow \qquad \mathbf{f}(\omega) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\mathbf{\Gamma}(k)e^{-ik\omega}$$

$$\mathbf{f}(\omega) = \begin{pmatrix} f_{11}(\omega) & \cdots & f_{1d}(\omega) \\ \vdots & \ddots & \vdots \\ f_{d1}(\omega) & \cdots & f_{dd}(\omega) \end{pmatrix}$$ is a function on $[0, \pi]$:

- matrix-valued
- 2π-periodic
- Hermitian
- positive definite

Spectral Density of VARMA(2,2)



Nonparametric Spectral Analysis

Given observations: $\mathbf{Z}_1, \ldots, \mathbf{Z}_n$

Nonparametric estimation of spectral density

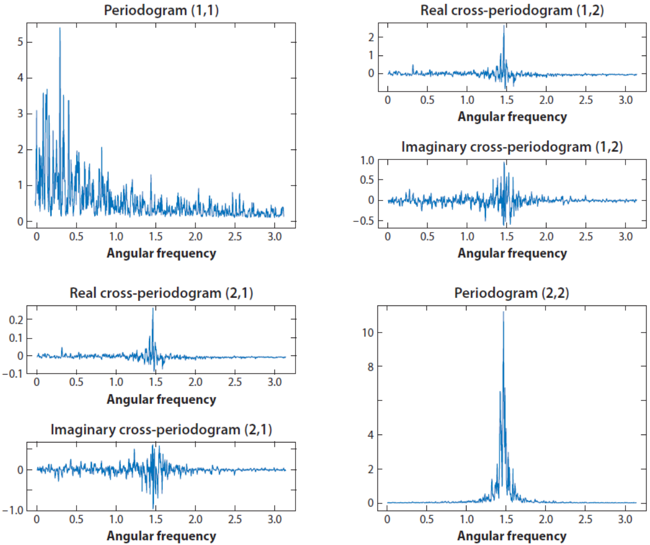
$$\mathbf{f}(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \mathbf{\Gamma}(k) e^{-ik\omega}$$

is based on the **periodogram matrix**

$$\mathbf{I}_n(\omega) = \frac{1}{2\pi n} \tilde{\mathbf{Z}}(\omega) \tilde{\mathbf{Z}}(\omega)^*$$

where $\tilde{\mathbf{Z}}(\omega) = \sum_{t=1}^{n} \mathbf{Z}_t e^{-it\omega}$ is the Fourier transform of the time series

Simulated VARMA(2,2), $n = 4096$



Smoothing the Periodogram

Periodogram is asympt. unbiased but not consistent, i.e.,

$$\begin{aligned} \mathrm{E}(\mathbf{I}_n(\omega)) &\rightarrow \mathbf{f}(\omega) \\ \mathrm{Var}(\mathbf{I}_n(\omega)) &\rightarrow \mathbf{f}^2(\omega) \end{aligned}$$

$\longrightarrow$ smoothing techniques in frequentist literature:

- kernel-based
- nearest neighbour
- multi-taper
- spline-based
- wavelet-based

For a recent review, see e.g. von Sachs (2020), Annu. Rev. Stat. Appl 7: 361–86

Bayesian Nonparametric Approach

Use asymptotic properties of periodograms:

- $\mathbf{I}_n(\omega_j)$ asymptotically independent, $\omega_j = \frac{2\pi j}{n}$
- $\mathbf{I}_n(\omega_j) \overset{asym}{\sim} \text{Wishart}_d(\mathbf{f}(\omega_j), N)$, $j = 0, \ldots, N = \lfloor (n-1)/2 \rfloor$

Multivariate Whittle likelihood (Whittle, 1957)

$$p^W(\mathbf{Z}|\mathbf{f}) \propto \exp\left\{ -\sum_{j=1}^{N} \left(\log(\det(2\pi \mathbf{f}(\omega_j)) + \frac{1}{2\pi} \tilde{\mathbf{Z}}_j^* \mathbf{f}(\omega_j)^{-1} \tilde{\mathbf{Z}}_j \right) \right\}$$

Bayesian Approaches to Multivariate Time Series

Previous Bayesian nonparametric approaches:

- smoothing splines for Cholesky components (Dai & Guo 2004, Rosen & Stoffer 2007, Zhang 2016, Li & Krafty 2018, Hu& Prado 2023)
- RJMCMC, Variational Bayes

Bayesian Approaches to Multivariate Time Series

Previous Bayesian nonparametric approaches:

- smoothing splines for Cholesky components
  (Dai & Guo 2004, Rosen & Stoffer 2007, Zhang 2016, Li & Krafty 2018, Hu& Prado 2023)
- RJMCMC, Variational Bayes

BUT

- posterior consistency?

Bayesian Approaches to Multivariate Time Series

Previous Bayesian nonparametric approaches:

- smoothing splines for Cholesky components (Dai & Guo 2004, Rosen & Stoffer 2007, Zhang 2016, Li & Krafty 2018, Hu& Prado 2023)
- RJMCMC, Variational Bayes

BUT

- posterior consistency?
- choice of smoothing parameter?

Generalized Whittle Likelihood

Generalized Whittle likelihood

$$p^{GW}(\mathbf{Z}|\mathbf{f}) \propto \det(\mathbf{C}\mathbf{C}^*)^{-1/2} p_{VAR}(\mathbf{F}^*\mathbf{C}^{-1}\mathbf{F}\mathbf{Z})$$

Generalized Whittle Likelihood

Generalized Whittle likelihood

$$p^{GW}(\mathbf{Z}|\mathbf{f}) \propto \det(\mathbf{C}\mathbf{C}^*)^{-1/2} p_{VAR}(\mathbf{F}^*\mathbf{C}^{-1}\mathbf{F}\mathbf{Z})$$

Start with VAR(p) working model with spectral density $\mathbf{f}_{VAR}$.

Generalized Whittle Likelihood

Generalized Whittle likelihood

$$p^{GW}(\mathbf{Z}|\mathbf{f}) \propto \det(\mathbf{C}\mathbf{C}^*)^{-1/2} p_{VAR}(\mathbf{F}^*\mathbf{C}^{-1}\mathbf{F}\mathbf{Z})$$

Start with VAR(p) working model with spectral density $\mathbf{f}_{VAR}$.

time domain

$\mathbf{Z} \sim p_{VAR}$

frequency domain

Generalized Whittle Likelihood

Generalized Whittle likelihood

$$p^{GW}(\mathbf{Z}|\mathbf{f}) \propto \det(\mathbf{C}\mathbf{C}^*)^{-1/2} p_{VAR}(\mathbf{F}^*\mathbf{C}^{-1}\mathbf{F}\mathbf{Z})$$

Start with VAR(p) working model with spectral density $\mathbf{f}_{VAR}$.

<u>time domain</u>		<u>frequency domain</u>
$\mathbf{Z} \sim p_{VAR}$	$\xrightarrow{\text{FT}}$	$\mathbf{F}\,\mathbf{Z}$

Generalized Whittle Likelihood

Generalized Whittle likelihood

$$p^{GW}(\mathbf{Z}|\mathbf{f}) \propto \det(\mathbf{C}\mathbf{C}^*)^{-1/2} p_{VAR}(\mathbf{F}^*\mathbf{C}^{-1}\mathbf{F}\mathbf{Z})$$

Start with VAR(p) working model with spectral density $\mathbf{f}_{VAR}$.

time domain

$\mathbf{Z} \sim p_{VAR}$ $\xrightarrow{\text{FT}}$

frequency domain

$\mathbf{F}\mathbf{Z}$

$\downarrow \; \mathbf{C}(\mathbf{f}, \mathbf{f}_{VAR}) = \text{diag}(\ldots, \mathbf{f}(\omega_i)^{1/2}\mathbf{f}_{VAR}^{-1/2}(\omega_i), \ldots$

$\mathbf{C}\mathbf{F}\mathbf{Z}$

Generalized Whittle Likelihood

Generalized Whittle likelihood

$$p^{GW}(\mathbf{Z}|\mathbf{f}) \propto \det(\mathbf{C}\mathbf{C}^*)^{-1/2} p_{VAR}(\mathbf{F}^*\mathbf{C}^{-1}\mathbf{F}\mathbf{Z})$$

Start with VAR(p) working model with spectral density $\mathbf{f}_{VAR}$.

time domain		frequency domain
$\mathbf{Z} \sim p_{VAR}$	$\xrightarrow{\text{FT}}$	$\mathbf{F}\mathbf{Z}$
		$\downarrow$ $\mathbf{C}(\mathbf{f}, \mathbf{f}_{VAR}) = \text{diag}(\ldots, \mathbf{f}(\omega_i)^{1/2}\mathbf{f}_{VAR}^{-1/2}(\omega_i), \ldots$
$\mathbf{F}^*\mathbf{C}\mathbf{F}\mathbf{Z} \sim p^{GW}$	$\xleftarrow{\text{FT}^{-1}}$	$\mathbf{C}\mathbf{F}\mathbf{Z}$

Properties

Proposition

1. The Whittle likelihood is a special case:
   Generalized Whittle likelihood of a Gaussian VAR(0)
2. If $\mathbf{f} = \mathbf{f}_{VAR}$, then $p^{GW} = p_{VAR}$.
3. The periodogram is asymptotically unbiased for the true spectral density under the Generalized Whittle likelihood.

Bernstein – Hpd Matrix Gamma Process Prior

Prior for $d = 1$
(Choudhuri et al. 2003)

$$f(\pi x) = \sum_{j=1}^{k} \Phi\left(\left(\frac{j-1}{k}, \frac{j}{k}\right]\right) b_{j,k}(x)$$

$b_{j,k}(x)$	:	polynomial basis
$k \sim p(k)$	:	polynomial degree
Φ	:	Gamma process

Increments: $\Phi(dx) \overset{\text{ind}}{\sim} \text{Ga}(\alpha, \beta)$

Bernstein – Hpd Matrix Gamma Process Prior

Prior for $d = 1$
(Choudhuri et al. 2003)

$$f(\pi x) = \sum_{j=1}^{k} \Phi\left(\left(\frac{j-1}{k}, \frac{j}{k}\right]\right) b_{j,k}(x)$$

$b_{j,k}(x)$: polynomial basis
$k \sim p(k)$: polynomial degree
Φ : Gamma process

Increments: $\Phi(dx) \overset{\text{ind}}{\sim} \text{Ga}(\alpha, \beta)$

Prior for $d > 1$
(Meier et al. 2020)

Bernstein – Hpd Matrix Gamma Process Prior

Prior for $d = 1$
(Choudhuri et al. 2003)

$$f(\pi x) = \sum_{j=1}^{k} \Phi\left(\left(\frac{j-1}{k}, \frac{j}{k}\right]\right) b_{j,k}(x)$$

$b_{j,k}(x)$: polynomial basis
$k \sim p(k)$: polynomial degree
Φ : Gamma process

Increments: $\Phi(dx) \overset{\text{ind}}{\sim} \text{Ga}(\alpha, \beta)$

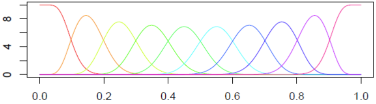
Prior for $d > 1$
(Meier et al. 2020)

$$\mathbf{f}(\pi x) = \sum_{j=1}^{k} \mathbf{\Phi}\left(\left(\frac{j-1}{k}, \frac{j}{k}\right]\right) b_{j,k}(x)$$

$b_{j,k}(x)$: polynomial basis
$k \sim p(k)$: polynomial degree
$\mathbf{\Phi}$: Hpd Gamma process

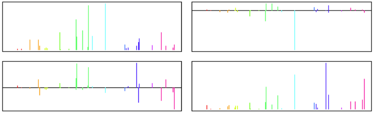
Increments: $\mathbf{\Phi}(dx) \overset{\text{ind}}{\sim} \text{Ga}_{d \times d}(\alpha, \beta)$

Example Polynomial Mixture $d = 2$



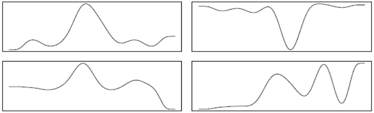
Polynomial basis $b_{j,k}$ for $k = 10$.

Coarsened Bernstein polynomials.



Realization of $\mathbf{\Phi}$.

Hpd Gamma Process.



Mixture $\sum_{j=1}^{k} \mathbf{\Phi}\left(\left(\frac{j-1}{k}, \frac{j}{k}\right]\right) b_{j,k}(x)$.

Hpd Gamma Process

Infinite Series Representation

$$\mathbf{\Phi} = \sum_{j=1}^{\infty} \delta_{x_j} r_j \mathbf{U}_j$$

with independent

$x_j \overset{iid}{\sim} U[0,1]$, $\mathbf{U}_j \overset{iid}{\sim} \alpha^\star$, $r_j = \rho^{-}_{\alpha,\beta}(w_j)$, $w_j = \sum_{i=1}^{j} v_i$, $v_i \overset{iid}{\sim} \text{Exp}(1)$

Simulation: inverse Lévy measure algorithm (Wolpert & Ickstadt, 1998)

Posterior Computation

- Generalized Whittle's Likelihood:

$$p^{GW}(\mathbf{Z}|\mathbf{f}) \propto \det(\mathbf{C}\mathbf{C}^*)^{-1/2} p_{VAR}(\mathbf{F}^*\mathbf{C}^{-1}\mathbf{F}\mathbf{Z})$$

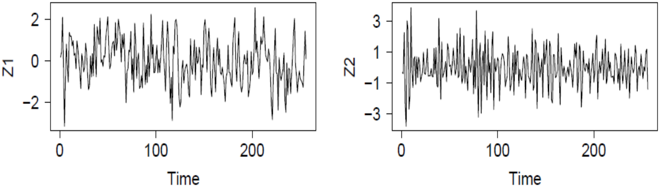
- Prior:
  - noninformative on VAR coefficients
  - Hpd Gamma process prior on $\boldsymbol{Q}(\omega) := \mathbf{f}_{VAR}^{-1/2}(\omega)\mathbf{f}(\omega)\mathbf{f}_{VAR}^{-1/2}(\omega)$

$$\boldsymbol{Q}(\pi x) = \sum_{j=1}^{k} \boldsymbol{\Phi}\left(\left(\frac{j-1}{k}, \frac{j}{k}\right]\right) b_{j,k}(x), \quad \boldsymbol{\Phi} = \sum_{j=1}^{L} \delta_{x_j} r_j \mathbf{U}_j$$

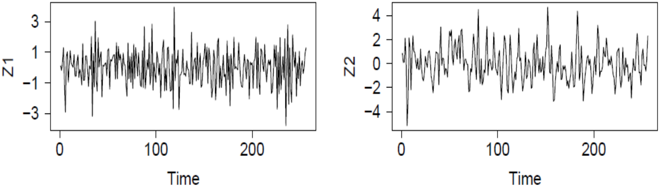
- Adaptive MH-within-Gibbs:
  R-package `beyondWhittle`
  GitHub: `https://github.com/easycure1/vnpctest`

Simulation Study

Data generated from a) VAR(2) and b) VMA(1)



(a)



(b)

Simulation Results

	VAR(2) model								
	$n = 256$			$n = 512$			$n = 1024$		
	GW	W	VAR	GW	W	VAR	GW	W	VAR
L_2-error	0.130	0.136	0.099	0.101	0.106	0.067	0.080	0.084	0.047
Coverage	0.826	0.548	0.908	0.718	0.374	0.898	0.616	0.348	0.886
Width $\boldsymbol{f}_{11}$	0.341	0.314	0.210	0.177	0.168	0.121	0.109	0.104	0.078

	VMA(1) model								
	$n = 256$			$n = 512$			$n = 1024$		
	GW	W	VAR	GW	W	VAR	GW	W	VAR
L_2-error	0.102	0.117	0.189	0.076	0.087	0.144	0.062	0.065	0.110
Coverage	0.888	0.594	0.980	0.876	0.518	0.972	0.690	0.294	0.966
Width $\boldsymbol{f}_{11}$	0.346	0.299	1.313	0.200	0.194	0.661	0.129	0.135	0.406

Simulation Results

	VAR(2) model								
	$n = 256$			$n = 512$			$n = 1024$		
	GW	W	VAR	GW	W	VAR	GW	W	VAR
L_2-error	0.130	0.136	0.099	0.101	0.106	0.067	0.080	0.084	0.047
Coverage	0.826	0.548	0.908	0.718	0.374	0.898	0.616	0.348	0.886
Width $\boldsymbol{f}_{11}$	0.341	0.314	0.210	0.177	0.168	0.121	0.109	0.104	0.078

	VMA(1) model								
	$n = 256$			$n = 512$			$n = 1024$		
	GW	W	VAR	GW	W	VAR	GW	W	VAR
L_2-error	0.102	0.117	0.189	0.076	0.087	0.144	0.062	0.065	0.110
Coverage	0.888	0.594	0.980	0.876	0.518	0.972	0.690	0.294	0.966
Width $\boldsymbol{f}_{11}$	0.346	0.299	1.313	0.200	0.194	0.661	0.129	0.135	0.406

Simulation Results

	VAR(2) model								
	$n = 256$			$n = 512$			$n = 1024$		
	GW	W	VAR	GW	W	VAR	GW	W	VAR
L_2-error	0.130	0.136	0.099	0.101	0.106	0.067	0.080	0.084	0.047
Coverage	0.826	0.548	0.908	0.718	0.374	0.898	0.616	0.348	0.886
Width $\boldsymbol{f}_{11}$	0.341	0.314	0.210	0.177	0.168	0.121	0.109	0.104	0.078

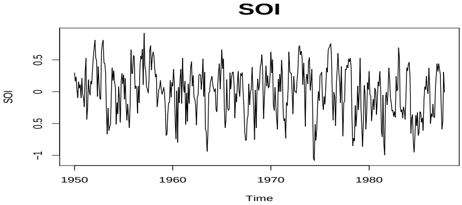
	VMA(1) model								
	$n = 256$			$n = 512$			$n = 1024$		
	GW	W	VAR	GW	W	VAR	GW	W	VAR
L_2-error	0.102	0.117	0.189	0.076	0.087	0.144	0.062	0.065	0.110
Coverage	0.888	0.594	0.980	0.876	0.518	0.972	0.690	0.294	0.966
Width $\boldsymbol{f}_{11}$	0.346	0.299	1.313	0.200	0.194	0.661	0.129	0.135	0.406

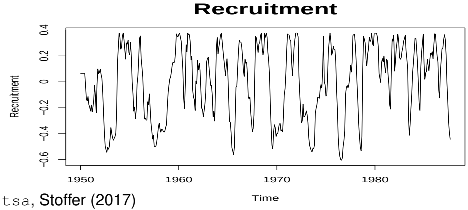
Applications

- Ecology/Oceanography: SOI-Recruitment
- Meteorology: Windspeeds at different locations
- Physiology: EEG

Southern Oscillation Index(SOI)

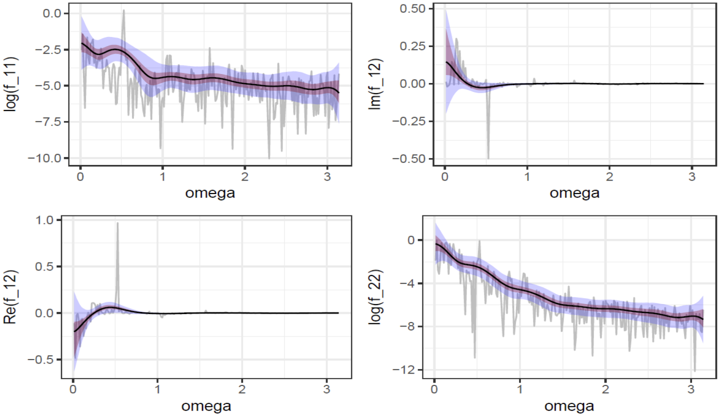
SOI $=$ monthly standardized anomaly of mean sea-level pressure difference between Tahiti and Darwin





Source: R package `astsa`, Stoffer (2017)

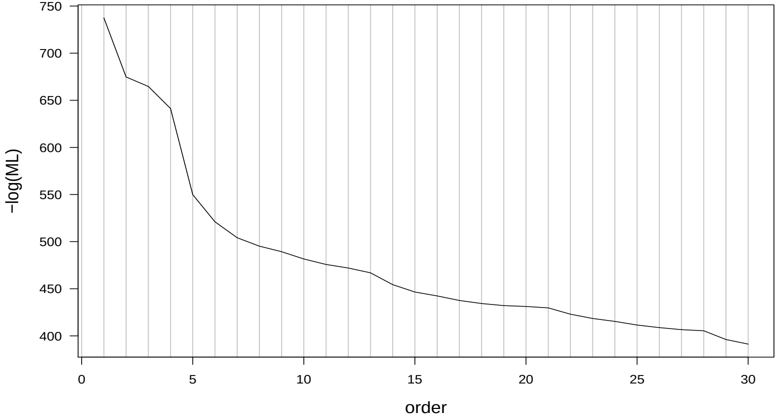
Results with Whittle Likelihood



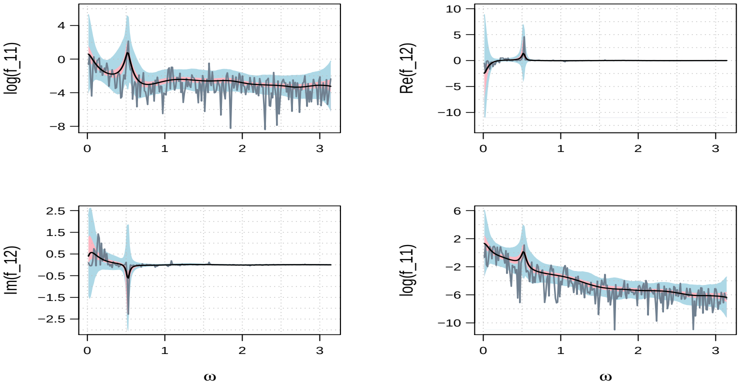
SOI spectrum peaks at $\omega = 0.52 \longrightarrow$ period $2\pi/\omega = 12$ months

strong annual autocorrelation

Choice of VAR order: Elbow Criterion



Results with Generalized Whittle Likelihood with VAR(5)

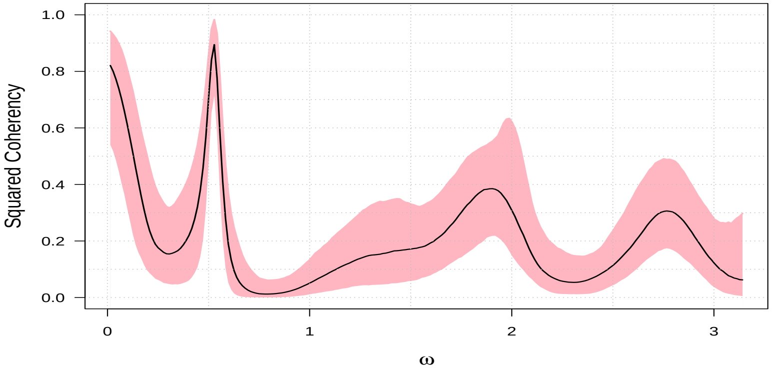


SOI spectrum peaks at $\omega = 0.52 \longrightarrow$ period $2\pi/\omega = 12$ months

strong annual autocorrelation

Squared Coherence

$$\frac{|f_{12}(\omega)|^2}{f_{11}(\omega) f_{22}(\omega)}$$

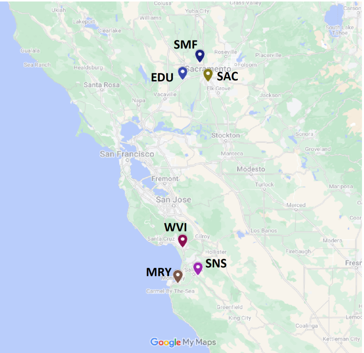


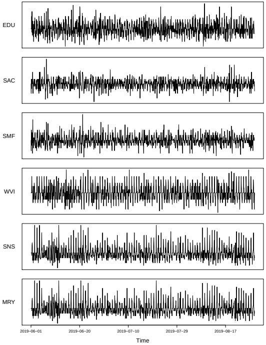
Applications

- Ecology/Oceanography: SOI-Recruitment
- Meteorology: Windspeeds at different locations
- Physiology: EEG

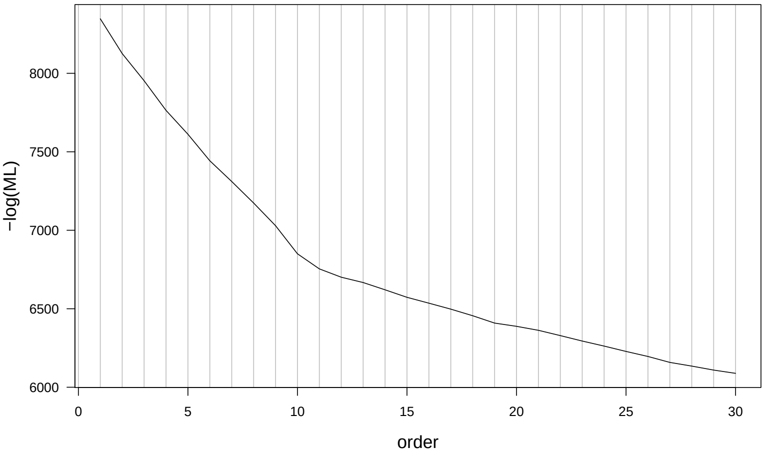
California Windspeed Data

Source: Iowa State University Environmental Mesonet Database, Hu and Prado (2023)

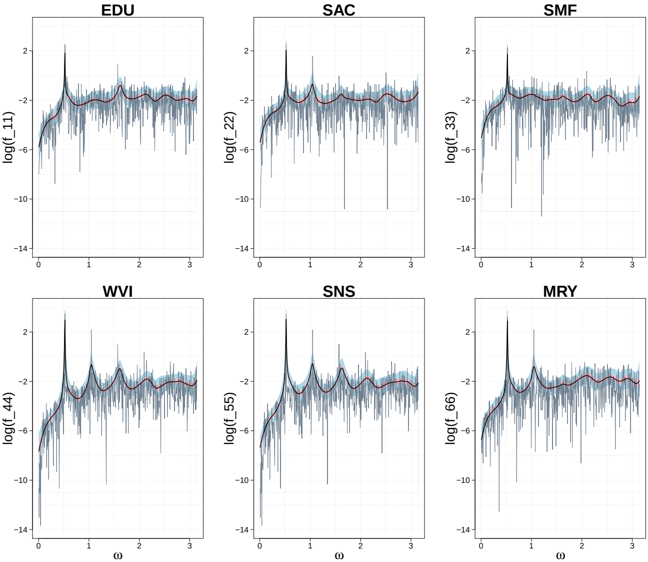




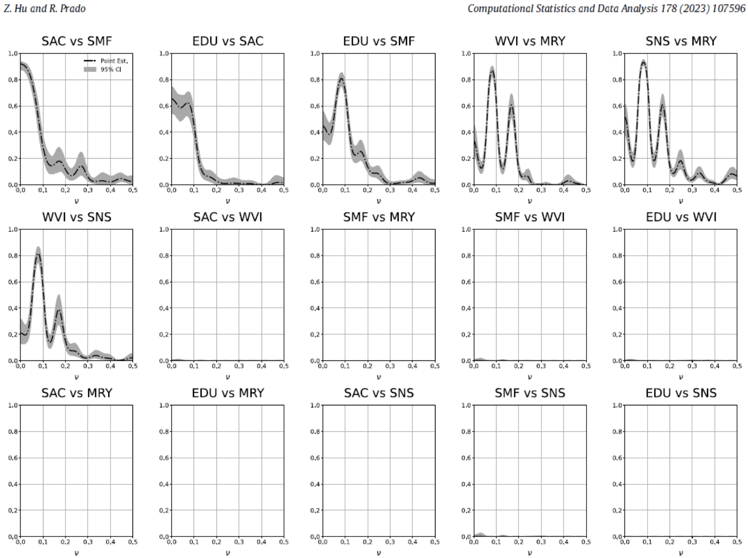
Elbow Criterion



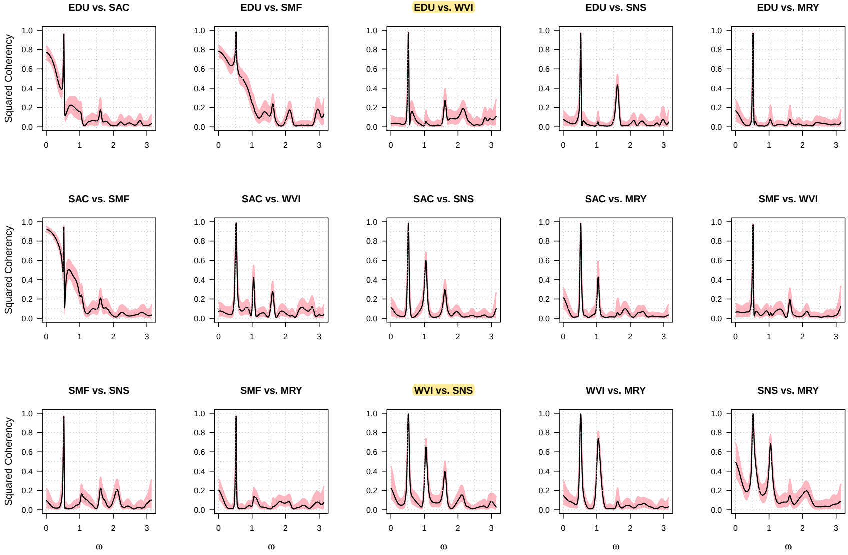
Spectral Density Estimates



Squared Coherences by Hu and Prado (2023)



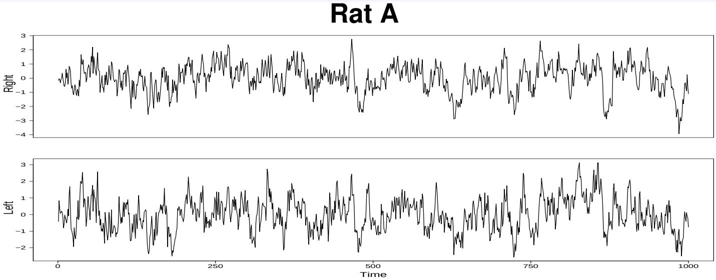
Squared Coherences using Generalized Whittle Likelihood

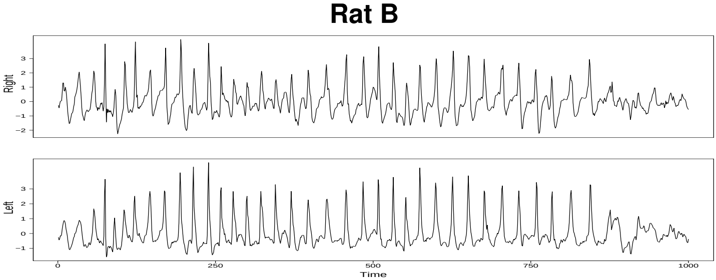


Applications

- Ecology/Oceanography: SOI-Recruitment
- Meteorology: Windspeeds at different locations
- Physiology: EEG

Two-Channel RAT EEG



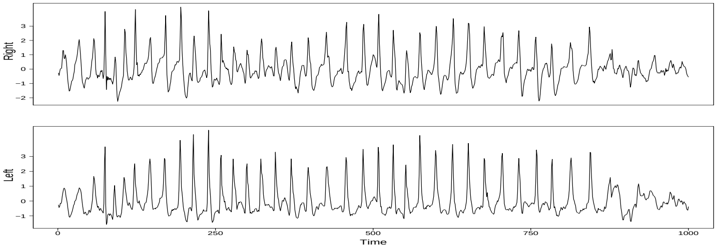


Source: Quiroga et al. (2002), Phys. Rev. E, 65, 041903

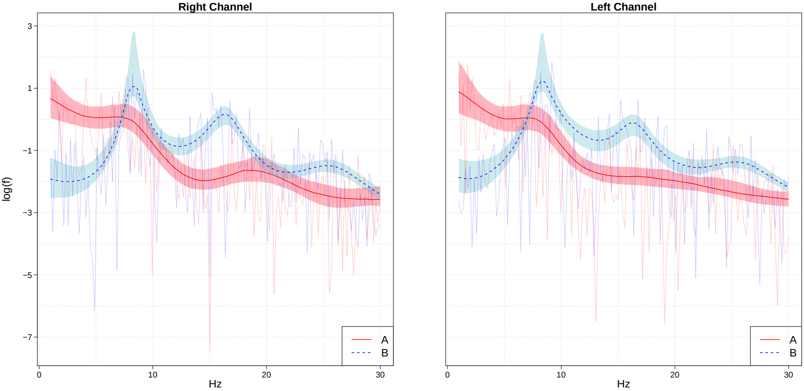
Epilepsy – Synchronization

- Spike discharges
- Synchronization between right and left channels
- Pathological synchronization → epileptic seizure

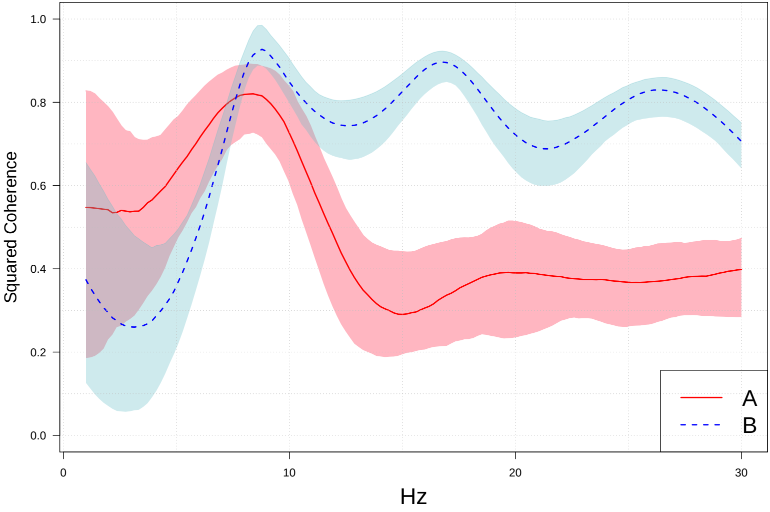
Rat B



GW Estimates of Spectral Densities



GW Estimates of Squared Coherence



References

Liu, Kirch, Lee, Meyer (2023).
A nonparametrically corrected likelihood for Bayesian spectral analysis of multivariate time series
http://arxiv.org/abs/2306.04966

Meier, Kirch, Meyer (2020).
Bayesian Nonparametric Analysis of Multivariate Time Series: A Matrix Gamma Process Approach.
Journal of Multivariate Analysis 175, 104560

Kirch, Edwards, Meier, Meyer (2019).
Beyond Whittle: Nonparametric Correction of a Parametric Likelihood with a Focus on Bayesian Time Series Analysis.
Bayesian Analysis 14, 1037-1073.

Hpd Gamma Distribution (Pérez-Abreu, Stelzer 2014)

- Radial decomposition: For $\mathbf{Z} > \mathbf{0}$ write $\mathbf{Z} = r\mathbf{U}$ with
  - radial part $r = \text{tr}(\mathbf{Z}) > \mathbf{0}$
  - spherical part $\mathbf{U} \in \mathbb{S} = \{\mathbf{U} > \mathbf{0} : \text{tr}\mathbf{U} = \mathbf{1}\}$
- α finite measure on $\mathbb{S}$ and $\beta : \mathbb{S} \longrightarrow (0, \infty)$

Hpd Gamma Distribution (Lévy-Khinchine representation)

$\mathbf{Z} \sim \text{Ga}_{d\times d}(\alpha, \beta)$ if for $\theta > \mathbf{0}$

$$\mathrm{E}e^{-tr(\theta\mathbf{Z})} = \exp\left(-\int_{\mathbb{S}}\int_0^\infty [1 - e^{-tr(r\theta\mathbf{U})}]\nu_{\alpha,\beta}(dr, d\mathbf{U})\right)$$

with Hpd Gamma Lévy measure

$$\nu_{\alpha,\beta}(dr, d\mathbf{U}) = \frac{1}{r}\exp(-r\beta(\mathbf{U}))dr\alpha(d\mathbf{U})$$

Hpd Gamma Process

Consider Poisson process Π on $[0,1] \times \{\mathbf{Z} > \mathbf{0}\}$ with mean measure $\nu_{\alpha,\beta}(dr, d\mathbf{U})\, dx$

Hpd Gamma Process (Kingman's Construction)

$$\boldsymbol{\Phi}(A) = \sum_{(x,\mathbf{Z})\in\Pi} 1_A(x)\,\mathbf{Z}, \quad A \subset [0,1]$$

Then $\boldsymbol{\Phi}(dx) \overset{ind}{\sim} \mathrm{Ga}_{d\times d}(\alpha, \beta)$

Infinite Series Representation

$$\boldsymbol{\Phi} = \sum_{j=1}^{\infty} \delta_{x_j} r_j \mathbf{U}_j$$

with independent

$x_j \overset{iid}{\sim} U[0,1]$, $\mathbf{U}_j \overset{iid}{\sim} \alpha^{\star}$, $r_j = \rho^{-}_{\alpha,\beta}(w_j)$, $w_j = \sum_{i=1}^{j} v_i$, $v_i \overset{iid}{\sim} \mathrm{Exp}(1)$