AdaptSPEC-X: a spectral method for handling spatially-dependent nonstationarity

Michael Bertolacci
michael_bertolacci@uow.edu.au

Centre for Environmental Informatics, University of Wollongong

2021-07-06

Collaborators

Ed Cripps, UWA

John Lau, UWA

Sally Cripps, USyd

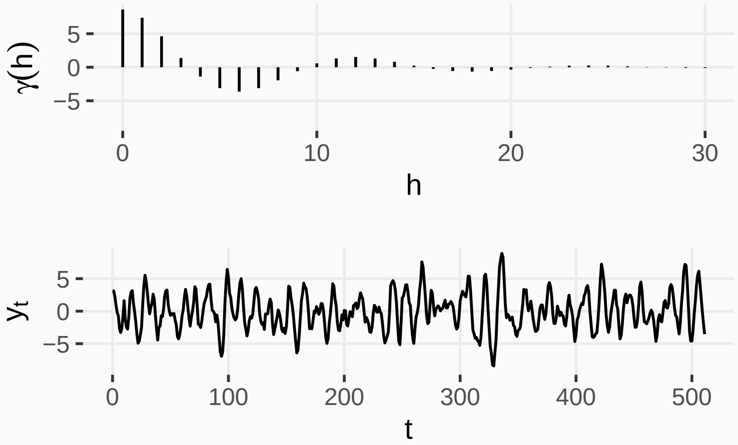
Ori Rosen, UTEP

Background

For y_t a stationary, mean zero time series:

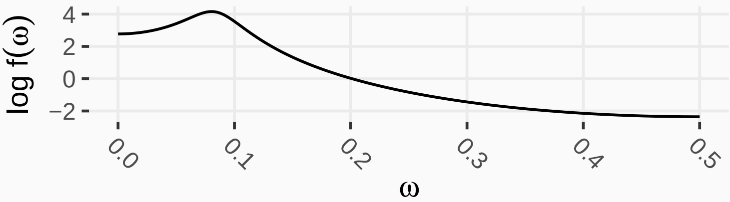
$$\mathrm{Cov}(y_{t+h}, y_t) = \gamma(h)$$

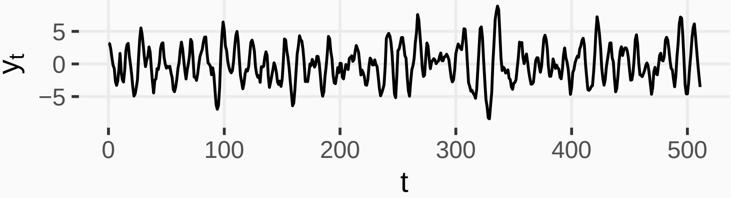
$\gamma(h)$ is the **autocovariance function**.



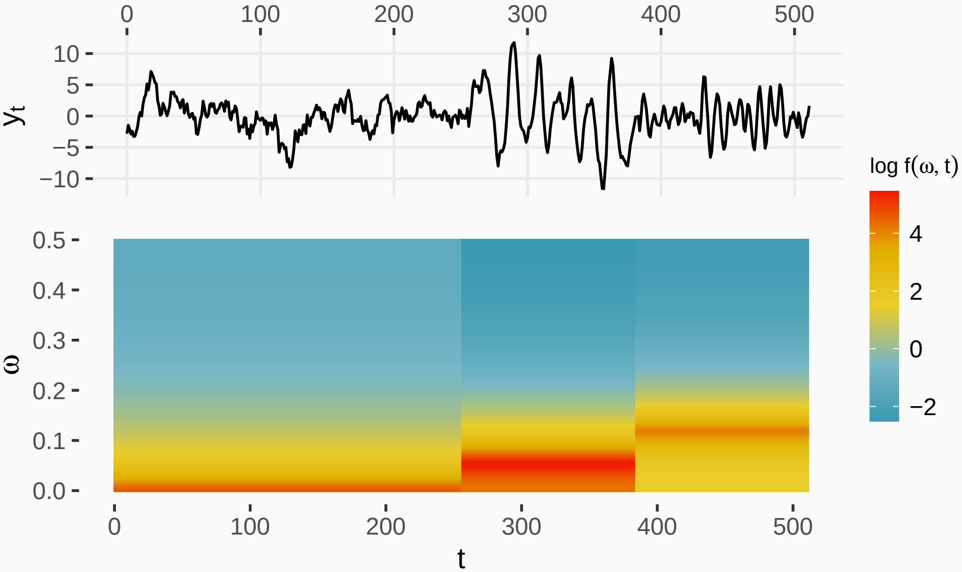
$$\gamma(h) = \int_{-1/2}^{1/2} \exp(2\pi i\omega h) f(\omega) d\omega,$$

$f(\omega)$ is the **spectral density**.





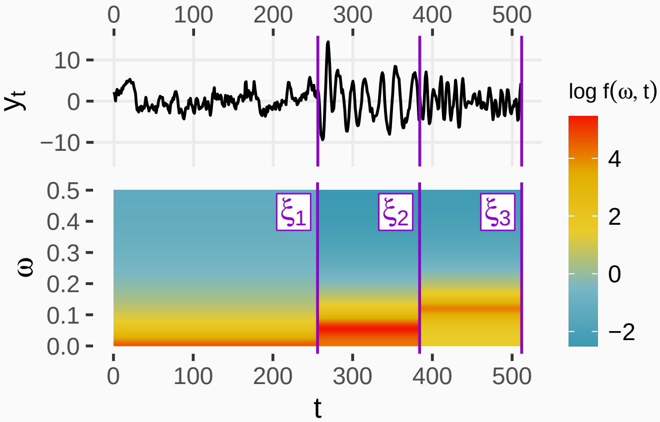
Can define the time-varying spectral density (or evolutionary spectrum), $f(\omega, t)$:



AdaptSPEC: Adaptive Spectral Estimation for Nonstationary Time Series

AdaptSPEC

A Bayesian method to estimate $f(\omega, t)$ for univariate time series (Rosen et al., 2012).



Let $\boldsymbol{y}^s = (y_1^s, \ldots, y_{n^s}^s)'$, $s = 1, \ldots, m$, we the data within a single segment. We assume stationarity. We model the log spectrum as a Gaussian Process (GP):

$$\log f^s(\omega) \sim \mathcal{GP}(\cdot, \cdot).$$

Use the Whittle likelihood (Whittle, 1957):

$$p(\boldsymbol{y}^s \mid f) \approx \prod_{k=1}^{n^s} \frac{1}{\sqrt{f^s(\omega_k)}} \exp\left\{-\frac{1}{2} \frac{|D_k^s|^2}{f^s(\omega_k)}\right\},$$

where $D_k^s = (n^s)^{-1/2} \sum_{t=0}^{n^s-1} (y_t^s - \mu^s) \exp(-2\pi i \omega_k t)$ is the periodogram.

Prior on mean: $\mu^s \sim N(0, \sigma_\mu^2)$.

This structure was pioneered by Wahba (1980).

AdaptSPEC: switches to new spectrum/mean at segment boundaries ξ_m^s.

$$\boldsymbol{y} = \sum_{s=1}^{m} \boldsymbol{y}^s \delta_m^s(t),$$

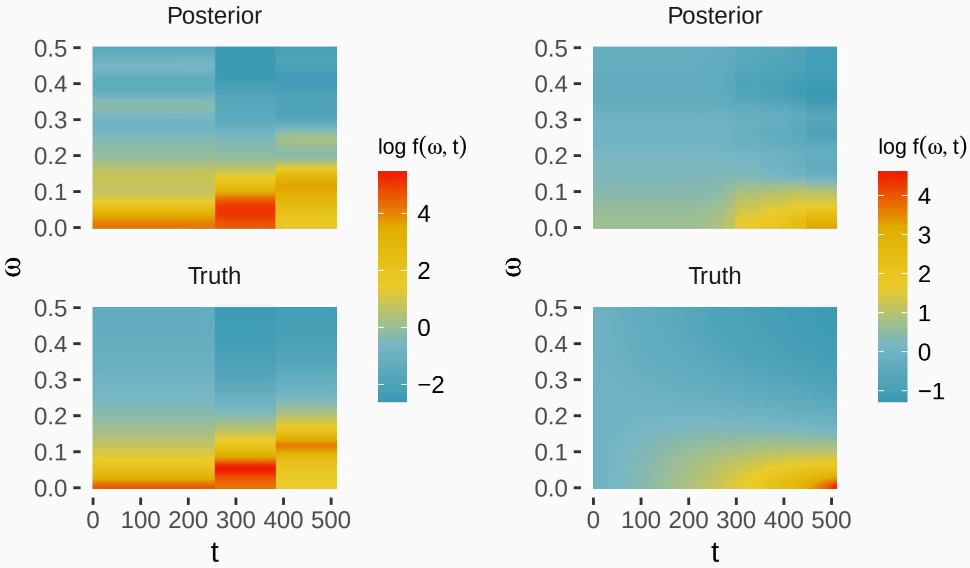
where $\delta_m^s(t) = 1$ iff $t \in (\xi_m^{s-1}, \xi_m^s]$.

Cutpoints ξ_m^s, $s = 1, \ldots, m$ and m assumed unknown.

Group all parameters into $\boldsymbol{\Theta}$, assign a prior $p(\boldsymbol{\Theta})$. Estimation uses Markov chain Monte Carlo (MCMC) for all unknowns, so get posterior distribution on:

- number and location of segment boundaries; and
- mean and spectrum within segments;
- $f(\omega, t)$ (time-varying spectrum) and $\mu(t)$ (time-varying mean);
- missing values.

Estimated time-varying spectra:



AdaptSPEC-X: Nonparametric spectral analysis for multiple time series

Suppose we have a collection of time-series $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_N$ with corresponding covariates $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_N$.

Key objects of interest:

- covariate-dependent time-varying spectrum: $f(\omega, t, \boldsymbol{s})$,
- covariate-dependent time-varying mean: $\mu(t, \boldsymbol{s})$.

Want to be able to predict $f(\omega, t, \boldsymbol{s})$ and $\mu(t, \boldsymbol{s})$ at unobserved $\boldsymbol{s}$.

We use a covariate-dependent mixture of L AdaptSPEC-based components.

Let

$$p(\boldsymbol{y}_j \mid \boldsymbol{\Theta}_1, \ldots, \boldsymbol{\Theta}_L) = \sum_{l=1}^{L} \pi(\boldsymbol{s}_j) g(\boldsymbol{y}_j \mid \boldsymbol{\Theta}_l)$$

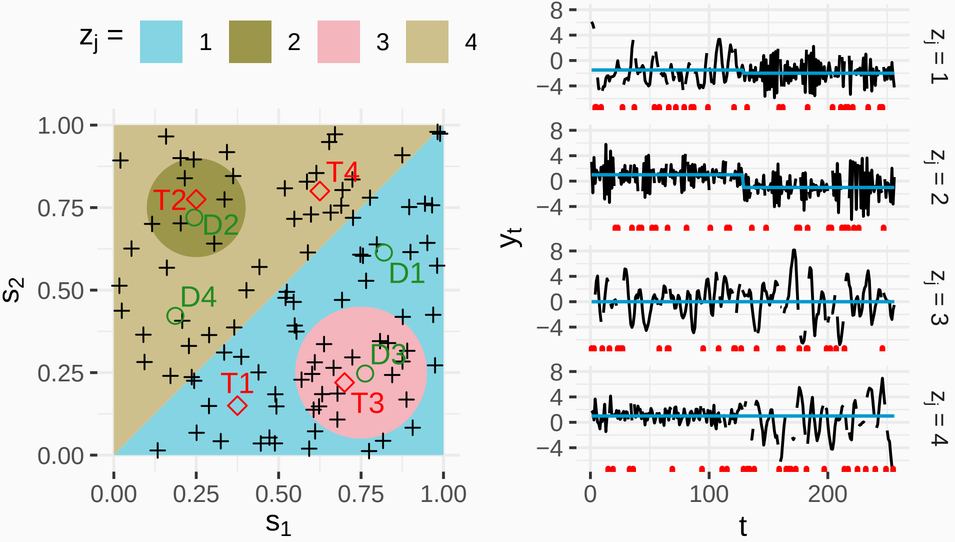
where $g(\cdot \mid \boldsymbol{\Theta}_l)$ is the the density of AdaptSPEC with parameters $\boldsymbol{\Theta}_l$, subject to $\sum_{l=1}^{L} \pi(\boldsymbol{s}_j) = 1$.

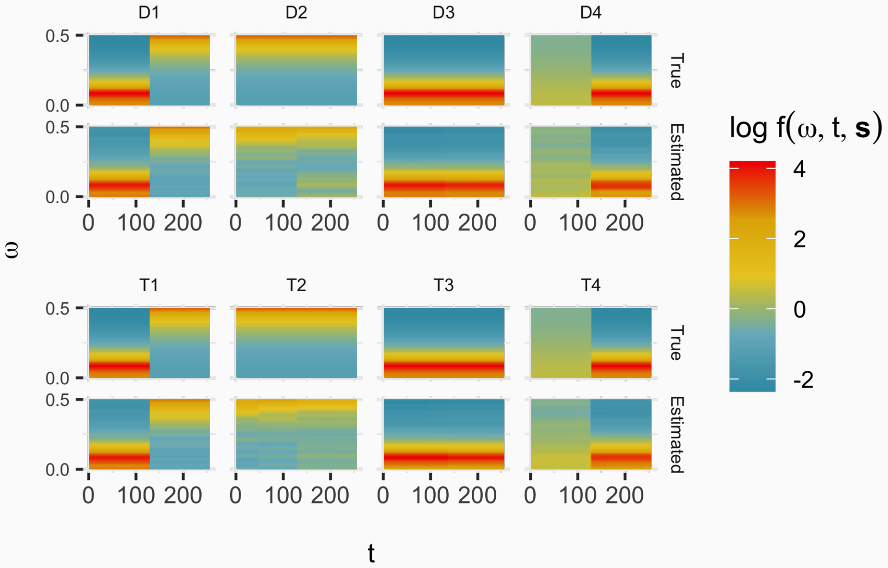
Use the logistic stick-breaking prior (LSBP) (Rigon and Durante, 2021):

$$\pi_l(\boldsymbol{s}) = \nu_l(\boldsymbol{s}) \sum_{k'=1}^{l-1} (1 - \nu_{l'}(\boldsymbol{s})),$$

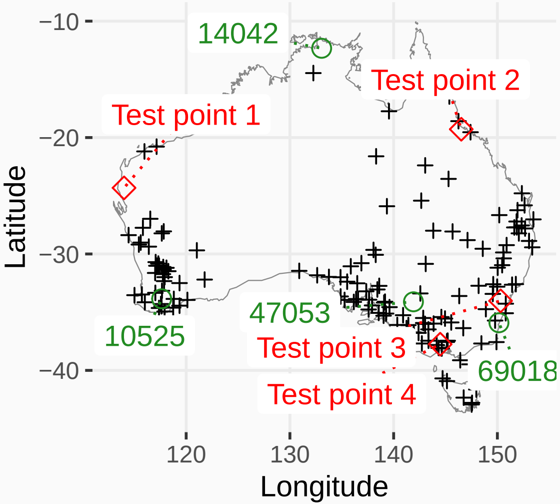
where $\operatorname{logit} \nu_l(\boldsymbol{s}) \sim \mathcal{GP}(\cdot, \cdot)$ for $l < L$ and $\nu_L(\boldsymbol{s}) = 1$.

Simulation study:

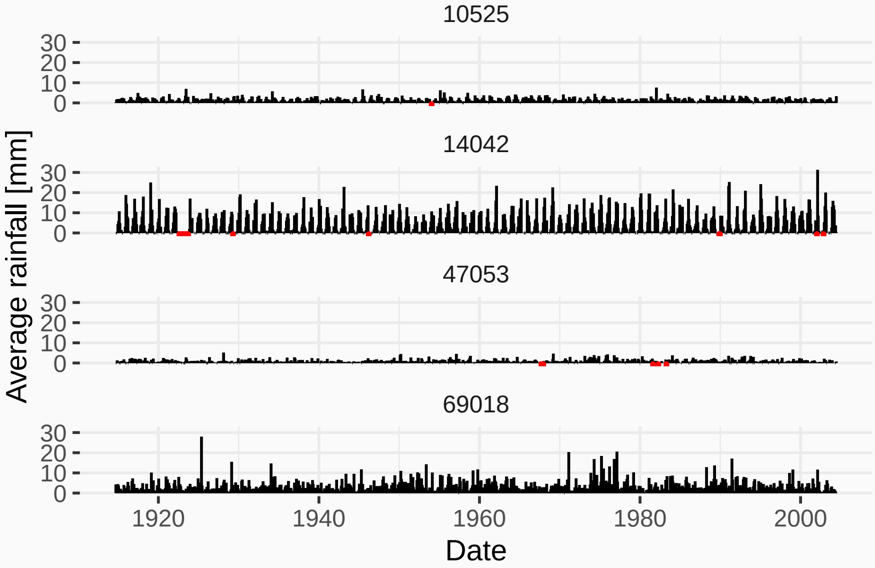


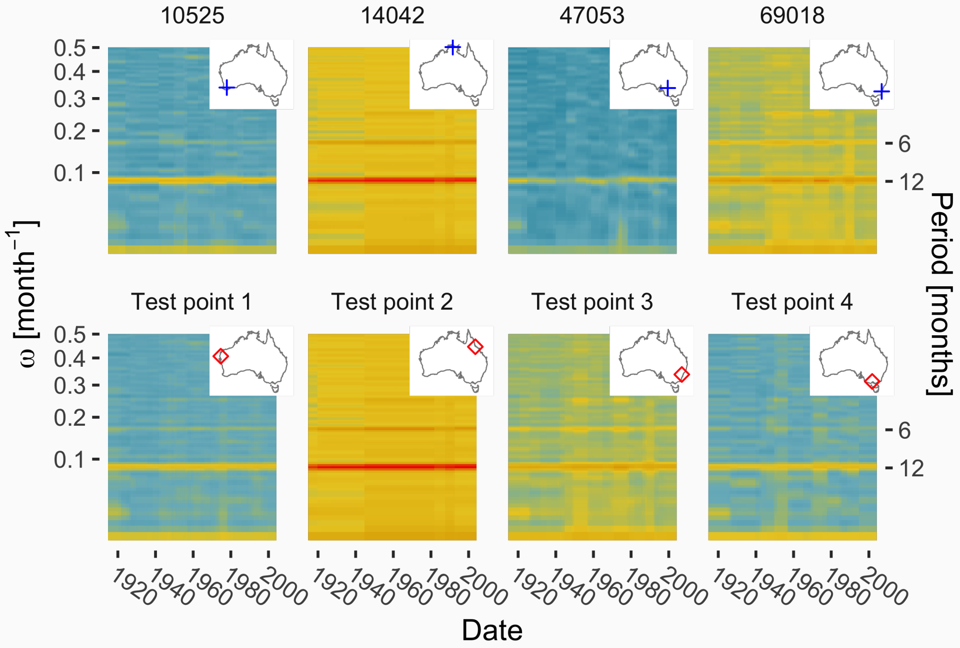


D1
D2
D3
D4
T1
T2
T3
T4
True
Estimated
ω
t
0.5
0.0
0
100
200
log f(ω, t, **s**)
4
2
0
-2



(Data from Bureau of Meteorology, a subset of that analysed by Bertolacci et al., 2019)


10525
14042
47053
69018
30
20
10
0
Average rainfall [mm]
1920
1940
1960
1980
2000
Date


10525
14042
47053
69018
Test point 1
Test point 2
Test point 3
Test point 4
0.5
0.4
0.3
0.2
0.1
ω [month⁻¹]
6
12
Period [months]
1920
1940
1960
1980
2000
Date

Main limitations:

- Whittle likelihood known to be inefficient for non-Gaussian time series and small sample sizes
- No explicit accounting for measurement error
- Covariates (i.e. spatial index) must be time-invariant

Thank you!

Preprint at `https://arxiv.org/abs/1908.06622`

Accepted with minor revisions by JCGS

Code at `https://github.com/mbertolacci/adaptspecx`

Methods will be available in the R package, BayesSpec (`https://cran.r-project.org/package=BayesSpec`)

Thank you ☺

References

Bertolacci, M., Cripps, E., Rosen, O., Lau, J. W., and Cripps, S. (2019). Climate inference on daily rainfall across the Australian continent, 1876–2015. *Annals of Applied Statistics*, 13(2):683–712.

Rigon, T. and Durante, D. (2021). Tractable Bayesian density regression via logit stick-breaking priors. *Journal of Statistical Planning and Inference*, 211:131–142.

Rosen, O., Wood, S., and Stoffer, D. S. (2012). AdaptSPEC: Adaptive spectral estimation for nonstationary time series. *Journal of the American Statistical Association*, 107(500):1575–1589.

Wahba, G. (1980). Automatic smoothing of the log periodogram. *Journal of the American Statistical Association*, 75(369):122–132.

Whittle, P. (1957). Curve and periodogram smoothing. *Journal of the Royal Statistical Society, Series B*, 19:38–47.