SPECTRAL ANALYSIS OF NONSTATIONARY TIME SERIES

DAVID S. STOFFER
Department of Statistics
University of Pittsburgh

Ori Rosen
UTEP

Sally Wood
U.Melbourne
Original problem was to do local spectral analysis for a multivariate point process, $X_t = e_j$, for $j = 1, \ldots, p$, where $e_j = (0, \ldots, 0, 1, 0, \ldots, 0)'$.

It ain’t easy.

Series of papers:

IN THE BEGINNING …

Original problem was to do local spectral analysis for a multivariate point process, $X_t = e_j$, for $j = 1, \ldots, p$, where $e_j = (0, \ldots, 0, 1, 0, \ldots, 0)'$.

It ain’t easy.

Series of papers:

IN THE BEGINNING …

Original problem was to do local spectral analysis for a multivariate point process, $X_t = e_j$, for $j = 1, \ldots, p$, where $e_j = (0, \ldots, 0, 1, 0, \ldots, 0)'$.

It ain’t easy.

Series of papers:

El Niño – Southern Oscillation

The graph shows the Southern Oscillation Index from 1880 to 2000. The x-axis represents the year, and the y-axis represents the Southern Oscillation Index. The graph distinguishes between El Niño and La Niña events, with red indicating El Niño years and blue indicating La Niña years. The data is provided by David Stoffer for his book "Spectral Analysis of the Future."
Annual Global Temperature Anomalies
1950 - 2011

- El Nino
- La Nina
- Other
Spectral Estimation – Stationary Case

For a stationary time series with bounded spectral density \( f(\nu) \), given mean-zero data \( \{X_1, \ldots, X_n\} \), let \( P_n(\nu_k) \) denote the periodogram. For large \( n \), approximately

\[
P_n(\nu_k) = f(\nu_k) U_k
\]

where \( U_k \overset{iid}{\sim} \text{Gamma}(1, 1) \).

Taking logs, we have a GLM

\[
y(\nu_k) = g(\nu_k) + \epsilon_k
\]

where \( y(\nu_k) = \log P_n(\nu_k) \), \( g(\nu_k) = \log f(\nu_k) \) and \( \epsilon_k \) are iid \( \log(\chi^2_2/2) \) s.

Want to fit the model with the constraint that \( g(\nu) \) is smooth. Wahba (1980) suggested smoothing splines. This can be done in a Bayesian framework.
SPECTRAL ESTIMATION – STATIONARY CASE

For a stationary time series with bounded spectral density $f(\nu)$, given mean-zero data $\{X_1, \ldots, X_n\}$, let $P_n(\nu_k)$ denote the periodogram. For large $n$, approximately

$$P_n(\nu_k) = f(\nu_k) U_k$$

where $U_k \sim \text{Gamma}(1, 1)$.

Taking logs, we have a GLM

$$y(\nu_k) = g(\nu_k) + \epsilon_k$$

where $y(\nu_k) = \log P_n(\nu_k)$, $g(\nu_k) = \log f(\nu_k)$ and $\epsilon_k$ are iid $\log(\chi^2_2/2)$s.

Want to fit the model with the constraint that $g(\nu)$ is smooth. Wahba (1980) suggested smoothing splines. This can be done in a Bayesian framework.
SPECTRAL ESTIMATION – STATIONARY CASE

For a stationary time series with bounded spectral density $f(\nu)$, given mean-zero data $\{X_1, \ldots, X_n\}$, let $P_n(\nu_k)$ denote the periodogram. For large $n$, approximately

$$P_n(\nu_k) = f(\nu_k) U_k$$

where $U_k \sim \text{Gamma}(1, 1)$.

Taking logs, we have a GLM

$$y(\nu_k) = g(\nu_k) + \epsilon_k$$

where $y(\nu_k) = \log P_n(\nu_k)$, $g(\nu_k) = \log f(\nu_k)$ and $\epsilon_k$ are iid $\log(\chi^2_2/2)$s.

Want to fit the model with the constraint that $g(\nu)$ is smooth. Wahba (1980) suggested smoothing splines. This can be done in a Bayesian framework.
**How It’s Done:**

(1) Need a Likelihood

**Whittle Likelihood**

Spectral density $f(\nu)$, and mean-zero data $x = \{X_1, \ldots, X_n\}$, for large $n$,

$$\log \mathcal{L}(f \mid x) \approx -\sum_{k=1}^{m} \left\{ \log f(\nu_k) + \frac{P_n(\nu_k)}{f(\nu_k)} \right\},$$

where $P_n(\nu_k)$ is the periodogram, $\nu_k = k/n$, and $k = 1, \ldots, m = \left\lfloor \frac{n-1}{2} \right\rfloor$. 
HOW IT’S DONE:
(2) Need Priors

Place a linear smoothing spline prior on $\log f(\nu)$:

recall $y_n(\nu_k) = \log P_n(\nu_k)$ and $g(\nu_k) = \log f(\nu_k)$

- $g(\nu_k) = \alpha_0 + \alpha_1 \nu_k + h(\nu_k)$ linear $[\alpha] +$ nonlinear $[h(\cdot)]$
- $h(\nu) = \tau \int_0^\nu W(u) \, du \Rightarrow W(u)$ is a GP with kernel depending on $\nu$.
- $\alpha_0 \sim N(0, \sigma_\alpha^2)$, $\alpha_1 \equiv 0$, since $(\partial g(\nu)/\partial \nu)|_{\nu=0} = 0$.
- $h = D\beta$, is a linear combination of basis functions where $h = (h(\nu_1), \ldots, h(\nu_m))'$, and the $j$th column of $D$ is $\sqrt{2} \cos(j \pi \nu)$, $\nu = (\nu_1, \ldots, \nu_m)'$.
- $\beta \sim N(0, \tau^2 I)$
- $\tau^2 \sim U(0, c_{\tau^2})$ is the smoothing parameter.
HOW IT’S DONE:
(2) Need Priors

Place a linear smoothing spline prior on $\log f(\nu)$:

recall $y_n(\nu_k) = \log P_n(\nu_k)$ and $g(\nu_k) = \log f(\nu_k)$

- $g(\nu_k) = \alpha_0 + \alpha_1 \nu_k + h(\nu_k) \quad \text{linear } [\alpha] + \text{nonlinear } [h(\cdot)]$
- $h(\nu) = \tau \int_0^\nu W(u) \, du \Rightarrow W(u)$ is a GP with kernel depending on \( \nu \).
  - $\alpha_0 \sim N(0, \sigma_\alpha^2)$, $\alpha_1 \equiv 0$, since $(\partial g(\nu) / \partial \nu)|_{\nu=0} = 0$.
  - $h = D\beta$, is a linear combination of basis functions where
    $h = (h(\nu_1), \ldots, h(\nu_m))'$, and the $j$th column of $D$ is $\sqrt{2} \cos(j\pi \nu)$,
    $\nu = (\nu_1, \ldots, \nu_m)'$.
  - $\beta \sim N(0, \tau^2 I)$
  - $\tau^2 \sim U(0, c \tau^2)$ is the smoothing parameter.
Place a linear smoothing spline prior on $\log f(\nu)$:

recall $y_n(\nu_k) = \log P_n(\nu_k)$ and $g(\nu_k) = \log f(\nu_k)$

- $g(\nu_k) = \alpha_0 + \alpha_1 \nu_k + h(\nu_k)$ \hspace{1cm} linear $[\alpha]$ + nonlinear $[h(\cdot)]$

- $h(\nu) = \tau \int_0^\nu W(u) \, du \Rightarrow W(u)$ is a GP with kernel depending on $\nu$.

- $\alpha_0 \sim N(0, \sigma^2_\alpha)$, $\alpha_1 \equiv 0$, since $(\partial g(\nu)/\partial \nu)|_{\nu=0} = 0$. 

- $h = D\beta$, is a linear combination of basis functions where $h = (h(\nu_1), \ldots, h(\nu_m))'$, and the $j$th column of $D$ is $\sqrt{2} \cos(j\pi\nu)$, $\nu = (\nu_1, \ldots, \nu_m)'$.

- $\beta \sim N(0, \tau^2 I)$

- $\tau^2 \sim U(0, c_{\tau^2})$ is the smoothing parameter.
How It's Done:

(2) Need Priors

Place a linear smoothing spline prior on $\log f(\nu)$:

recall $y_n(\nu_k) = \log P_n(\nu_k)$ and $g(\nu_k) = \log f(\nu_k)$

- $g(\nu_k) = \alpha_0 + \alpha_1 \nu_k + h(\nu_k)$  \quad \text{linear [$\alpha$] + nonlinear [$h(\cdot)$]}

- $h(\nu) = \tau \int_0^\nu W(u) du \Rightarrow W(u)$ is a GP with kernel depending on $\nu$.

- $\alpha_0 \sim N(0, \sigma_\alpha^2)$, $\alpha_1 \equiv 0$, since $(\partial g(\nu) / \partial \nu)|_{\nu=0} = 0$.

- $h = D\beta$, is a linear combination of basis functions where $h = (h(\nu_1), \ldots, h(\nu_m))'$, and the $j$th column of $D$ is $\sqrt{2} \cos(j \pi \nu)$, $\nu = (\nu_1, \ldots, \nu_m)'$.

- $\beta \sim N(0, \tau^2 I)$

- $\tau^2 \sim U(0, c_{\tau^2})$ is the smoothing parameter.
Place a linear smoothing spline prior on $\log f(\nu)$:

recall $y_n(\nu_k) = \log P_n(\nu_k)$ and $g(\nu_k) = \log f(\nu_k)$

- $g(\nu_k) = \alpha_0 + \alpha_1 \nu_k + h(\nu_k)$  linear $[\alpha]$ + nonlinear $[h(\cdot)]$
- $h(\nu) = \tau \int_0^\nu W(u) du \Rightarrow W(u)$ is a GP with kernel depending on $\nu$.
- $\alpha_0 \sim N(0, \sigma_\alpha^2)$, $\alpha_1 \equiv 0$, since $(\partial g(\nu)/\partial \nu)|_{\nu=0} = 0$.
- $h = D\beta$, is a linear combination of basis functions where
  $h = (h(\nu_1), \ldots, h(\nu_m))'$, and the $j$th column of $D$ is $\sqrt{2} \cos(j \pi \nu)$,
  $\nu = (\nu_1, \ldots, \nu_m)'$.
- $\beta \sim N(0, \tau^2 I)$
- $\tau^2 \sim U(0, c\tau^2)$ is the smoothing parameter.
How It’s Done:

(2) Need Priors

Place a linear smoothing spline prior on $\log f(\nu)$:

recall $y_n(\nu_k) = \log P_n(\nu_k)$ and $g(\nu_k) = \log f(\nu_k)$

- $g(\nu_k) = \alpha_0 + \alpha_1 \nu_k + h(\nu_k)$ \quad linear [\alpha] + nonlinear [h(\cdot)]$
- $h(\nu) = \tau \int_{0}^{\nu} W(u) \, du$ \Rightarrow $W(u)$ is a GP with kernel depending on $\nu$.
- $\alpha_0 \sim N(0, \sigma_{\alpha}^2)$, $\alpha_1 \equiv 0$, since $(\partial g(\nu)/\partial \nu)|_{\nu=0} = 0$.
- $h = D\beta$, is a linear combination of basis functions where $h = (h(\nu_1), \ldots, h(\nu_m))^\prime$, and the $j$th column of $D$ is $\sqrt{2} \cos(j \pi \nu)$, $\nu = (\nu_1, \ldots, \nu_m)^\prime$.
- $\beta \sim N(0, \tau^2 I)$
- $\tau^2 \sim U(0, c\tau^2)$ is the smoothing parameter.
The parameters $\alpha_0$, $\beta$ and $\tau^2$ are drawn from the posterior distribution $p(\alpha_0, \beta, \tau^2 \mid y)$, where $y = (y_n(\nu_1), \ldots, y_n(\nu_m))'$, using MCMC:

- $\alpha_0$ and $\beta$ are sampled jointly via an M-H step from

$$
\log p(\alpha_0, \beta \mid \tau^2, y) \propto - \sum_{k=1}^{m} \left[ \alpha_0 + d_k' / \beta + \exp (y_n(\nu_k) - \alpha_0 - d_k' / \beta) \right] - \frac{\alpha_0^2}{2\sigma_0^2} - \frac{1}{2\tau^2} \beta' \beta,
$$

where $d_k'$ is the $k$th row of $D$.

- $\tau^2$ is sampled from the truncated inverse gamma distribution,

$$
p(\tau^2 \mid \beta) \propto (\tau^2)^{-m/2} \exp \left( - \frac{1}{2\tau^2} \beta' \beta \right), \quad \tau^2 \in (0, c_{\tau^2}].
$$

A consistency result can be established for the sampled posterior under restrictive assumptions: The time series is short-memory Gaussian and $\tau$ is known + other minor details. Choudhuri et al. (2004). Bayesian Estimation of the Spectral Density of a Time Series, JASA.
The parameters $\alpha_0$, $\beta$ and $\tau^2$ are drawn from the posterior distribution $p(\alpha_0, \beta, \tau^2 \mid y)$, where $y = (y_n(\nu_1), \ldots, y_n(\nu_m))'$, using MCMC:

- $\alpha_0$ and $\beta$ are sampled jointly via an M-H step from

$$
\log p(\alpha_0, \beta \mid \tau^2, y) \propto - \sum_{k=1}^{m} \left[ \alpha_0 + d'_k/\beta + \exp(y_n(\nu_k) - \alpha_0 - d'_k/\beta) \right] - \frac{\alpha_0^2}{2\sigma_0^2} - \frac{1}{2\tau^2} \beta'\beta,
$$

where $d'_k$ is the $k$th row of $D$.

- $\tau^2$ is sampled from the truncated inverse gamma distribution,

$$
p(\tau^2 \mid \beta) \propto (\tau^2)^{-m/2} \exp\left(-\frac{1}{2\tau^2} \beta'\beta\right), \quad \tau^2 \in (0, c_{\tau^2}].
$$

A consistency result can be established for the sampled posterior under restrictive assumptions: The time series is short-memory Gaussian and $\tau$ is known + other minor details. Choudhuri et al. (2004). Bayesian Estimation of the Spectral Density of a Time Series, JASA.
**SAMPLING SCHEME**

The parameters $\alpha_0$, $\beta$ and $\tau^2$ are drawn from the posterior distribution $p(\alpha_0, \beta, \tau^2 \mid y)$, where $y = (y_n(\nu_1), \ldots, y_n(\nu_m))'$, using MCMC:

- $\alpha_0$ and $\beta$ are sampled jointly via an M-H step from

  $$\log p(\alpha_0, \beta \mid \tau^2, y) \propto -\sum_{k=1}^{m} \left[ \alpha_0 + d'_k/\beta + \exp(y_n(\nu_k) - \alpha_0 - d'_k/\beta) \right] - \frac{\alpha_0^2}{2\sigma^2_\alpha} - \frac{1}{2\tau^2} \beta'\beta,$$

  where $d'_k$ is the $k$th row of $D$.

- $\tau^2$ is sampled from the truncated inverse gamma distribution,

  $$p(\tau^2 \mid \beta) \propto (\tau^2)^{-m/2} \exp\left(-\frac{1}{2\tau^2} \beta'\beta\right), \quad \tau^2 \in (0, c_{\tau^2}].$$

A consistency result can be established for the sampled posterior under restrictive assumptions: The time series is short-memory Gaussian and $\tau$ is known + other minor details. Choudhuri et al. (2004). Bayesian Estimation of the Spectral Density of a Time Series, JASA.
AR(2) Example

- $\log P_n$
- true $g = \log f$
- --- 95% credible intervals
LOCAL STATIONARITY  Swimming the Thames Approach

- Divide time \( N \) into \( S \) segments of equal size \( n = N/S \) (small).

- For a given segmentation, \( S \), assume stationarity and estimate the log-spectrum, \( g_s(\nu) \), from the data, \( \{X_s(t); t = 1, \ldots, n; s = 1, \ldots, S\} \) in that segment.

- Model \( g_s(\nu) \) as a mixture of an unknown number, \( J \), of spectra:

\[
g_s(\nu) = \sum_{j=1}^{J} \pi_{sj} \log f_j(\nu)
\]

where the \( f_j(\nu) \) are spectra and the \( \pi_{sj} \) are associated weights.

- In a given segment \( s \), we calculate the log periodogram \( y_s(\nu_k) \), where \( \nu_k = k/n \). We model these as:

\[
y_s(\nu_k) = g_s(\nu_k) + \epsilon_k
\]

where the \( \epsilon_k \)'s for \( k = 1, \ldots, \lfloor n/2 \rfloor \) are distributed as \( \log(\chi_{2/2}^2) \).
LOCAL STATIONARITY  Swimming the Thames Approach

- Divide time \((N)\) into \(S\) segments of equal size \(n = N/S\) (small).

- For a given segmentation, \(S\), assume stationarity and estimate the log-spectrum, \(g_s(\nu)\), from the data, \(\{X_s(t); t = 1, \ldots, n; s = 1, \ldots, S\}\) in that segment.

- Model \(g_s(\nu)\) as a mixture of an unknown number, \(J\), of spectra \(\pi_{sj}\) log \(f_j(\nu)\)

\[
g_s(\nu) = \sum_{j=1}^{J} \pi_{sj} \log f_j(\nu)
\]

where the \(f_j(\nu)\) are spectra and the \(\pi_{sj}\) are associated weights.

- In a given segment \(s\), we calculate the log periodogram \(y_s(\nu_k)\), where \(\nu_k = k/n\). We model these as

\[
y_s(\nu_k) = g_s(\nu_k) + \epsilon_k
\]

where the \(\epsilon_k\)'s for \(k = 1, \ldots, [n/2]\) are distributed as \(\log(\chi^2_{2}/2)\).
LOCAL STATIONARITY  Swimming the Thames Approach

- Divide time \((N)\) into \(S\) segments of equal size \(n = N/S\) (small).

- For a given segmentation, \(S\), assume stationarity and estimate the log-spectrum, \(g_s(\nu)\), from the data, \(\{X_s(t); t = 1, \ldots, n; s = 1, \ldots, S\}\) in that segment.

- Model \(g_s(\nu)\) as a mixture of an unknown number, \(J\), of spectra

\[
g_s(\nu) = \sum_{j=1}^{J} \pi_{sj} \log f_j(\nu)
\]

where the \(f_j(\nu)\) are spectra and the \(\pi_{sj}\) are associated weights.

- In a given segment \(s\), we calculate the log periodogram \(y_s(\nu_k)\), where \(\nu_k = k/n\). We model these as

\[
y_s(\nu_k) = g_s(\nu_k) + \epsilon_k
\]

where the \(\epsilon_k\)'s for \(k = 1, \ldots, [n/2]\) are distributed as \(\log(\chi^2_{2}/2)\).
LOCAL STATIONARITY  Swimming the Thames Approach

- Divide time \( N \) into \( S \) segments of equal size \( n = N/S \) (small).
  \[ - - | - - | - - | - - | - - | - - | - - | - - | - - | - - | - - | \]
- For a given segmentation, \( S \), assume stationarity and estimate the log-spectrum, \( g_s(\nu) \), from the data, \( \{X_s(t); t = 1, \ldots, n; s = 1, \ldots, S\} \) in that segment.
- Model \( g_s(\nu) \) as a mixture of an unknown number, \( J \), of spectra
  \[ g_s(\nu) = \sum_{j=1}^{J} \pi_{sj} \log f_j(\nu) \]
  where the \( f_j(\nu) \) are spectra and the \( \pi_{sj} \) are associated weights.
- In a given segment \( s \), we calculate the log periodogram \( y_s(\nu_k) \), where \( \nu_k = k/n \). We model these as
  \[ y_s(\nu_k) = g_s(\nu_k) + \epsilon_k \]
  where the \( \epsilon_k \)'s for \( k = 1, \ldots, \lfloor n/2 \rfloor \) are distributed as \( \log(\chi^2_2/2) \).
Fitting the Model

Estimation is similar to the stationary case but with the additional mixing assumption. The number of spectral components in the model is \( J \), which we take as unknown, with prior \( \Pr(J) \).

For the mixing probabilities we use the multinomial linear logit model,

\[
\pi_{js} = \frac{\exp(\delta_j' u_s)}{\sum_{j=1}^J \exp(\delta_j' u_s)}
\]

with parameters \( \delta_j, j = 1, \ldots, J \). Here, \( u_s = (1, s/S)' \) and \( \delta_j = (\delta_{0j}, \delta_{1j})' \). Such logistic weights are also used in the mixtures-of-experts model.

Estimate the log of the spectral density in segment \( s \), for \( s = 1, \ldots, S \) by its posterior mean using RJ-MCMC to perform the required multidimensional integration. The posterior mean is

\[
E\{g_s(\nu) \mid y\} = \sum_J \int E\{g_s(\nu) \mid y, \theta_J, J\} p(\theta_J \mid y, J) d\theta_J \Pr(J \mid y)
\]

where \( \theta_J \) are the hyper-parameters and \( \Pr(J \mid y) \) is the posterior probability that there are \( J \) components.
Fitting the Model

Estimation is similar to the stationary case but with the additional mixing assumption. The number of spectral components in the model is $J$, which we take as unknown, with prior $\Pr(J)$.

For the mixing probabilities we use the multinomial linear logit model,

$$
\pi_{js} = \exp(\delta_j' u_s) / \sum_{j=1}^{J} \exp(\delta_j' u_s)
$$

with parameters $\delta_j$, $j = 1, \ldots, J$. Here, $u_s = (1, s/S)'$ and $\delta_j = (\delta_{0j}, \delta_{1j})'$. Such logistic weights are also used in the mixtures-of-experts model.

Estimate the log of the spectral density in segment $s$, for $s = 1, \ldots, S$ by its posterior mean using RJ-MCMC to perform the required multidimensional integration. The posterior mean is

$$
E\{g_s(\nu)|y\} = \sum_{J} \int E\{g_s(\nu)|y, \theta_J, J\} p(\theta_J|y, J) d\theta_J \Pr(J|y)
$$

where $\theta_J$ are the hyper-parameters and $\Pr(J|y)$ is the posterior probability that there are $J$ components.
Fitting the Model

Estimation is similar to the stationary case but with the additional mixing assumption. The number of spectral components in the model is \( J \), which we take as unknown, with prior \( \Pr(J) \).

For the mixing probabilities we use the multinomial linear logit model,

\[
\pi_{js} = \frac{\exp(\delta'_j u_s)}{\sum_{j=1}^J \exp(\delta'_j u_s)}
\]

with parameters \( \delta_j, j = 1, \ldots, J \). Here, \( u_s = (1, s/S)' \) and \( \delta_j = (\delta_{0j}, \delta_{1j})' \). Such logistic weights are also used in the mixtures-of-experts model.

Estimate the log of the spectral density in segment \( s \), for \( s = 1, \ldots, S \) by its posterior mean using RJ-MCMC to perform the required multidimensional integration. The posterior mean is

\[
E\{g_s(\nu)|y\} = \sum_J \int E\{g_s(\nu)|y, \theta_J, J\} p(\theta_J|y, J) d\theta_J \Pr(J|y)
\]

where \( \theta_J \) are the hyper-parameters and \( \Pr(J|y) \) is the posterior probability that there are \( J \) components.
We generated a time series of length 1024 from the following piecewise stationary model

\[
x_t = \begin{cases} 
0.9x_{t-1} + \epsilon_t & \text{if } 1 \leq t \leq 450 \\
-0.9x_{t-1} + \epsilon_t & \text{if } 451 \leq t \leq 1024 
\end{cases}
\]

where \( \epsilon_t \sim N(0, 1) \).

We use \( S = 16 \) segments of \( n = 64 \) data points. The change is in the 7th segment.

We assume we know there are \( J = 2 \) components. The log spectra in segment \( s \) for \( s = 1, \ldots, 16 \) is modeled by \( g_s(\nu) = \pi_s \log f_1(\nu) + (1 - \pi_s) \log f_2(\nu) \).
ADAPTSPEC — AN ADAPTIVE METHOD

Piecewise Stationary

Suppose $x = \{X_1, \ldots, X_n\}$ is a time series with an unknown number of stationary segments.

- $m$: unknown number of segments ($m = 1$ means stationary)
- $n_{j,m}$: number of observations in the $j$th segment, $n_{j,m} \geq t_{\text{min}}$.
- $\xi_{j,m}$: location of the end of the $j$th segment, $j = 0, \ldots, m$, $\xi_{0,m} \equiv 0$ and $\xi_{m,m} \equiv n$.
- $f_{j,m}$: spectral densities
- $P_{n_{j,m}}$: periodograms at $\nu_{k_j} = k_j / n_{j,m}$, $0 \leq k_j \leq n_{j,m} - 1$. 
Prior Distributions

- Priors on $g_{j,m}(\nu) = \log f_{j,m}(\nu)$, $j = 1, \ldots, m$, as before.

- $\Pr(\xi_{j,m} = t \mid m) = 1/p_{jm}$, for $j = 1, \ldots, m - 1$, where $p_{jm}$ is the number of available locations for split point $\xi_{j,m}$.

- The prior on the number of segments

\[
\Pr(m = k) = 1/M, \quad \text{for} \quad k = 1, \ldots, M.
\]
Within-model moves: (location of end points)

- Given \( m, \xi_{k^*}, m \) is proposed to be relocated.
- The corresponding and \( \beta \)s are updated (absorb \( \alpha_0 \)s into \( \beta \)s).
- These two steps are jointly accepted or rejected in a M-H step.
- The \( \tau^2 \)s are then updated in a Gibbs step.
Between-model moves: \(^{(\text{number of segments})}\)

- \(m^p = m^c + 1 \dagger\)
  - Select a segment to split
  - Select a new split point in this segment.
  - Two new \(\tau^2\)'s are formed from the current \(\tau^2\)
  - Two new \(\beta\)'s are drawn.

- \(m^p = m^c - 1\)
  - Select a split point to be removed.
  - A single \(\tau^2\) is then formed from the current \(\tau^2\)'s
  - A new \(\beta\) is proposed.

Accept or Reject in a RJ-MCMC to move between models.

\(\dagger\) \(c=\)current, \(p=\)proposed
EXAMPLE — the benefits of model averaging

Consider two tvAR(1) models $X_t = a_t X_{t-1} + W_t$ for $t = 1, \ldots, 500$

(blue) $a_t = t/500 - .5$  There is no optimal segmentation in this case.
(green) $a_t = .5 \text{sign}(t - 250)$
In each case, \( m = 2 \) is the modal value [posteriors in paper] on the number of partitions. Plotted below are \( \Pr(\xi_{1,2} = t \mid data) \) and \( \Pr(\xi_{1,2} < t \mid data) \), where \( \xi_{1,2} \) is the change point when \( m = 2 \).
We're so sorry, Uncle ENSO.

Think globally – act locally.

AdaptSPEC.

Motivate this. It's all the same.

D avid Stoffer

Spectral analysis of the future.

Change-point →

Time-varying →
Plots of (a) SOI from 1876–2011; (b) Niño3.4 index from 1950-2011; (c) DSLPA from 1951–2010.
The Posteriors – \( \Pr(\#\text{segments} = k \mid x) \)

<table>
<thead>
<tr>
<th>( k )</th>
<th>SOI</th>
<th>Niño3.4</th>
<th>DSLPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.95</td>
<td>0.93</td>
<td>0.99</td>
</tr>
<tr>
<td>2</td>
<td>0.05</td>
<td>0.07</td>
<td>0.01</td>
</tr>
<tr>
<td>3</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>
Simulation

\[
x_t = \begin{cases} 
\sum_{k=1}^{6} \phi_{1k} x_{t-k} + \sigma_1 \epsilon_t & \text{for } 1 \leq t \leq 200 \\
\sum_{k=1}^{6} \phi_{2k} x_{t-k} + \sigma_2 \epsilon_t & \text{for } 201 \leq t \leq 1000 \\
\sum_{k=1}^{6} \phi_{3k} x_{t-k} + \sigma_3 \epsilon_t & \text{for } 1001 \leq t \leq 1300 \\
\sum_{k=1}^{6} \phi_{4k} x_{t-k} + \sigma_4 \epsilon_t & \text{for } 1301 \leq t \leq 1600 \\
\sum_{k=1}^{6} \phi_{5k} x_{t-k} + \sigma_5 \epsilon_t & \text{for } 1601 \leq t \leq 2000,
\end{cases}
\]
THANKS TO THE ORGANIZERS

ALL YOU NEED IS LOVE – WITH BREAKS