Assessing Characteristic Scales Using Wavelets

Don Percival

Applied Physics Laboratory
Department of Statistics
University of Washington
Seattle, Washington, USA

NSF-sponsored collaborative effort with Mike Keim

overheads for talk available at

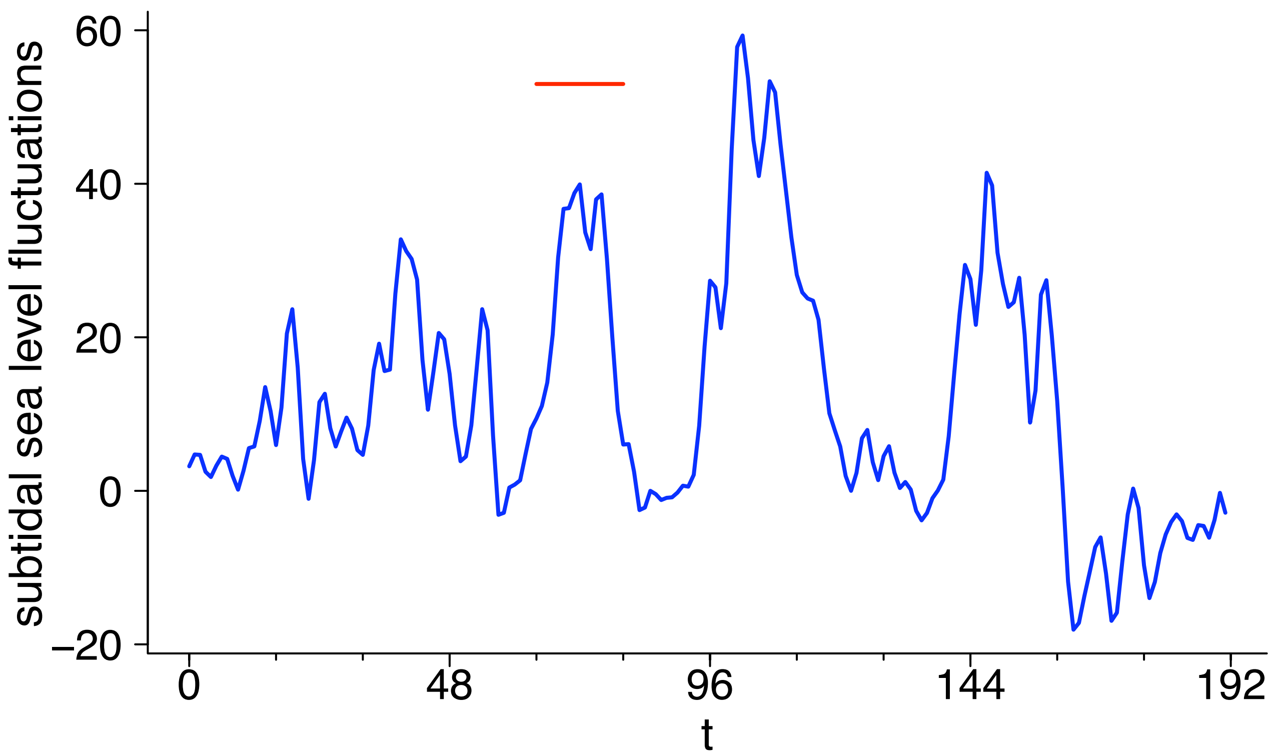
`http://faculty.washington.edu/dbp/talks.html`

Overview

- will discuss notion of characteristic scale, noting its lack of a standard definition
- will consider new definition in terms of discrete wavelet transform (DWT) as formulated by Daubechies, Mallat and others
- will present background on DWT, the wavelet variance (WV) and its sampling theory for intrinsically stationary Gaussian processes
- will then formulate definition of characteristic scale in terms of WV and present large-sample theory for its estimation
- will give four examples of time series of geophysical interest
- will conclude with comments on extensions of potential interest in telemedicine

Characteristic Scale: I

- geophysical time series often seem describable as a series of 'states' or 'events' whose durations tend to cluster around a value known as a characteristic scale (CS)



Characteristic Scale: II

- CS not defined outside of summary statistics used to extract it from a time series X_t, $t \in \mathbb{Z}$ (set of all integers)
- von Storch and Zwiers (1999) discuss several definitions when X_t is taken to be a stochastic process
- one definition quantifies 'memory' of process
- assuming $\mathbf{P}[X_{t+\tau} > 0 \,|\, X_t > 0] > 0.5$ for small lags $\tau > 0$, but $\mathbf{P}[X_{t+\tau} > 0 \,|\, X_t > 0] = 0.5$ at large lags, CS is smallest τ such that latter holds; i.e., CS is length of time needed for process to 'forget' its current positive state
- definition intuitively appealing, but of limited use: e.g., if X_t is a first order autoregressive process (AR(1)), $\tau = \infty$

Characteristic Scale: III

- second attempt at a definition (assumes X_t is stationary process)
- suppose $\text{var}\,\{X_t\} = \sigma^2$
- form sample mean of $X_1, X_2, \ldots, X_N$, i.e.,

$$\overline{X} = \frac{1}{N}\sum_{t=1}^{N} X_t, \text{ and consider its variance var}\,\{\overline{X}\}$$

- suppose $Y_1, Y_2, \ldots, Y_N$ are independent & identically distributed random variables, also with variance σ^2
- implies $\text{var}\,\{\overline{Y}\} = \sigma^2/N$
- equivalent sample size N' defined by setting $\text{var}\,\{\overline{X}\} = \sigma^2/N'$
- limit of N/N' as $N \to \infty$ defines a decorrelation time τ_D, a reasonable definition of CS for some – but not all – time series

Characteristic Scale: IV

- can show that
$$\tau_D = \sum_{k \in \mathbb{Z}} \rho_k,$$
where $\rho_k = \mathrm{cov}\{X_{t+k}, X_t\}/\sigma^2$ is autocorrelation sequence (ACS) for X_t
- for an AR(1) process, above reduces to $\tau_D = (1+\rho_1)/(1-\rho_1)$
- for other stationary processes, can use estimated ACS to estimate τ_D (doing so can be tricky!)
- von Storch and Zwiers (1999) note that variance of other sample statistics can be used in a similar manner to define a CS
- other definitions in literature involve structure functions (semi-variograms), fractal dimension, detrended fluctuation analysis, ...

Wavelet-based Definition of Characteristic Scale

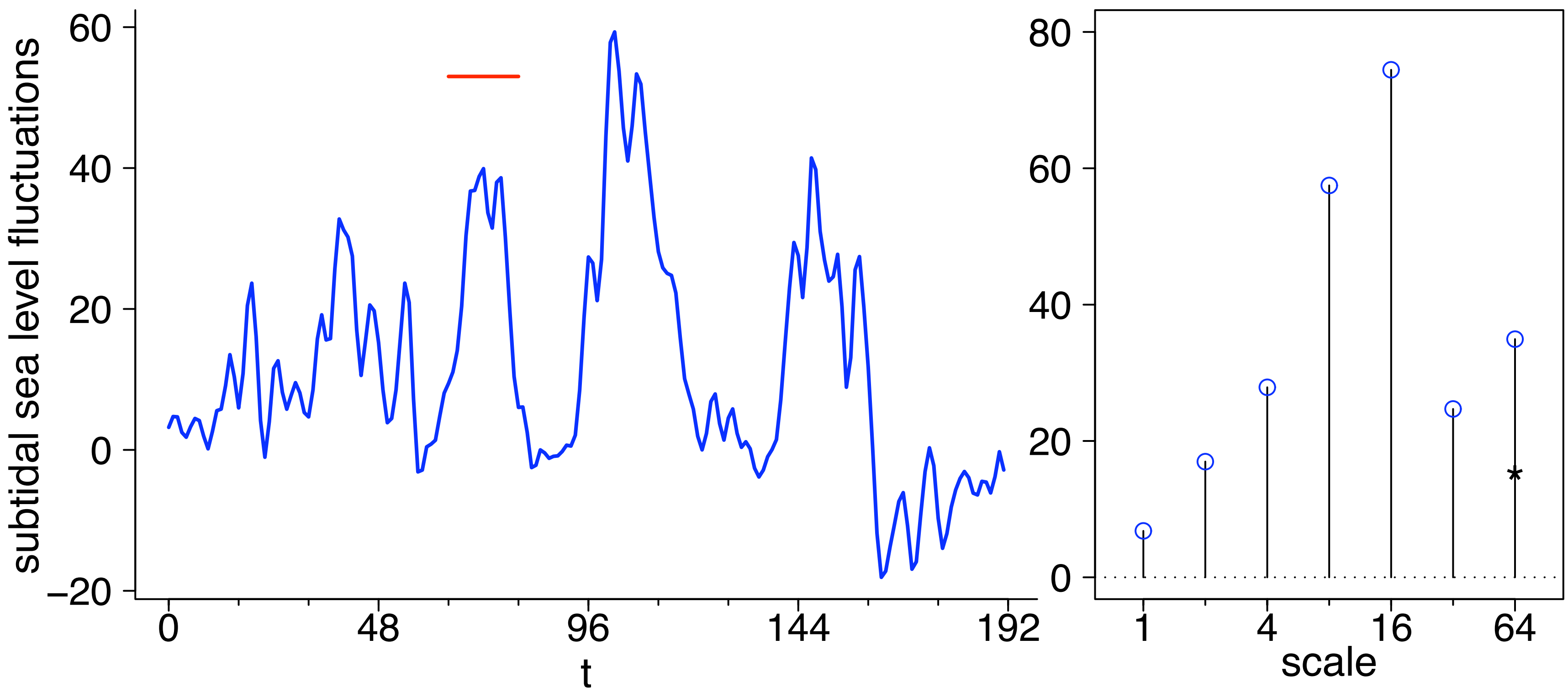
- theme of talk: a wavelet-based definition of CS
- discrete wavelet transform (DWT) describable as scale-based
- to fix ideas, consider Haar maximal overlap DWT, which yields wavelet coefficients $W_{\tau,t}^{\text{Haar}}$ for integer scale τ as follows:

$$2W_{\tau,t}^{\text{Haar}} = \frac{1}{\tau}\sum_{l=0}^{\tau-1} X_{t-l} - \frac{1}{\tau}\sum_{l=0}^{\tau-1} X_{t-\tau-l},$$

  i.e., a difference in adjacent averages spanning τ values
- if 'event' near X_t is of duration τ, $|W_{\tau,t}^{\text{Haar}}|$ will tend to be large
- summary of 'largeness' of $|W_{\tau,t}^{\text{Haar}}|$ is var $\{W_{\tau,t}^{\text{Haar}}\}$, which is known as the wavelet variance (WV)
- large var $\{W_{\tau,t}^{\text{Haar}}\}$ should provide basis for useful CS definition

Wavelet Variance for Subtidal Sea-level Time Series

- right-hand plot shows decomposition of sample variance of X_t via empirical wavelet variances at scales 1, 2, 4, 8, 16, 32 and 64, plus a term (the asterisk) summarizing all higher scales



Background on Wavelet Variance: I

- wavelet variance well-defined for an intrinsically stationary process X_t of integer order $d \geq 0$, so let's review this concept
- if $d = 0$, X_t is a stationary process; i.e., both $E\{X_t\}$ and $\text{cov}\,\{X_{t+\tau}, X_t\} = s_\tau$ – its autocovariance sequence (ACVS) – exist and are finite and independent of t
- if $d > 0$, subjecting X_t to dth order backward difference filter yields a stationary process with ACVS $s_\tau^{(d)}$, namely,

$$X_t^{(d)} = \sum_{k=0}^{d} \binom{d}{k} (-1)^k X_{t-k} = \begin{cases} X_t - X_{t-1} \text{ for } d = 1; \\ X_t - 2X_{t-1} + X_{t-2} \text{ for } d = 2; \\ \quad \vdots, \end{cases}$$

whereas $X_t^{(d-1)}, \ldots, X_t^{(1)}$ & $X_t^{(0)} = X_t$ are all nonstationary (an example being a random walk process)

Background on Wavelet Variance: II

- let $h_{1,l}, l = 0, 1, \ldots, L_1 - 1$, be unit-level Daubechies wavelet filter of even width L_1 with normalization $\sum_l h_{1,l}^2 = 1/2$
- $L_1 = 2$ case is the Haar filter: $h_{1,0} = \frac{1}{2}$, $h_{1,1} = -\frac{1}{2}$
- for $L_1 \geq 4$, use of $h_{1,l}$ equivalent to subjecting output from order $\frac{L_1}{2}$ backward difference filter to width $\frac{L_1}{2}$ low-pass filter
- can use unit level filter to construct jth level filter $h_{j,l}$ of width $L_j = (2^j - 1)(L_1 - 1) + 1$
- jth level wavelet coefficient process for X_t given by

$$W_{j,t} = \sum_{l=0}^{L_j - 1} h_{j,l} X_{t-l}$$

Background on Wavelet Variance: III

- coefficient $W_{j,t}$ is proportional to changes in adjacent weighted averages with effective scale (span) of $\tau_j = 2^{j-1}$
- assuming X_t is a dth order intrinsically stationary process and $L_1 \geq 2d$, wavelet coefficient process $W_{j,t}$ is stationary
- wavelet variance ν_j^2 for X_t at scale τ_j is variance of $W_{j,t}$:

$$\nu_j^2 = \text{var}\{W_{j,t}\}$$

- if X_t is stationary,

$$\text{var}\{X_t\} = \sum_{j=1}^{\infty} \nu_j^2,$$

in which case ν_j^2 is contribution to overall variance due to changes in adjacent weighted averages over scale τ_j

Background on Wavelet Variance: IV

- on the hand, $\sum_{j=1}^{J} \nu_j^2 \to \infty$ as $J \to \infty$ if X_t is intrinsically stationary of order $d \geq 1$, but ν_j^2 can still be interpreted as measuring variability of changes in adjacent weighted averages
- can express ν_j^2 in terms of ACVS $s_\tau^{(d)}$ for underlying stationary component $X_t^{(d)}$ for X_t:

$$\nu_j^2 = s_0^{(d)} \sum_{l=0}^{L_j-d-1} \left(h_{j,l}^{(d)}\right)^2 + 2 \sum_{\tau=1}^{L_j-d-1} s_\tau^{(d)} \sum_{l=0}^{L_j-d-1-\tau} h_{j,l}^{(d)} h_{j,l+\tau}^{(d)},$$

where $h_{j,l}^{(d)}$ is dth-order cumulative summation of $h_{j,l}$ (i.e., $h_{j,l}^{(1)} = \sum_{n=0}^{l} h_{j,n}$, while $h_{j,l}^{(2)} = \sum_{n=0}^{l} h_{j,n}^{(1)}$ and so forth)

Estimation Theory for Wavelet Variance: I

- given a time series $X_0, X_1, \ldots, X_{N-1}$, can compute level j wavelet coefficients for indices $L_j - 1 \leq t \leq N-1$ assuming $M_j = N - L_j + 1 > 0$
- sufficient (but not necessary) condition for $W_{j,t}$ to be a zero mean stationary process is $L_1 > 2d$ ($W_{j,t}$ is necessarily stationary if $L_1 = 2d$, but might not have zero mean)
- assuming L_1 chosen such that $W_{j,t}$ is a zero mean stationary process, then $\nu_j^2 = E\{W_{j,t}^2\}$ and hence

$$\hat{\nu}_j^2 = \frac{1}{M_j} \sum_{t=L_j-1}^{N-1} W_{j,t}^2$$

  is unbiased estimator of wavelet variance ν_j^2

Estimation Theory for Wavelet Variance: II

- can deduce second moment properties of $\hat{\nu}_j^2$ under assumption that $W_{j,t}$'s are multivariate Gaussian: for $j \leq k$, have

$$\operatorname{cov}\{\hat{\nu}_j^2, \hat{\nu}_k^2\} = \frac{2}{M_j} \sum_{\tau=-(M_k-1)}^{M_k-1} \left(1 - \frac{|\tau|}{M_k}\right) s_{j,k,\tau}^2 + \frac{2}{M_j M_k} \sum_{t=L_j-1}^{L_k-2} \sum_{u=L_k-1}^{N-1} s_{j,k,t-u}^2,$$

where $s_{j,k,\tau}$ is cross-covariance sequence for bivariate stationary processes $W_{j,t}$ & $W_{k,t}$:

$$s_{j,k,\tau} = \operatorname{cov}\left\{W_{j,t+\tau}, W_{k,t}\right\} = \sum_{l=0}^{L_j-d-1} h_{j,l}^{(d)} \sum_{m=0}^{L_k-d-1} h_{k,m}^{(d)} s_{\tau-l+m}^{(d)}$$

Estimation Theory for Wavelet Variance: III

- if M_j is not too small (≥ 100 or so), useful approximation is

$$\operatorname{cov}\{\hat{\nu}_j^2, \hat{\nu}_k^2\} \approx \frac{2A_{j,k}}{M_j}, \text{ where } A_{j,k} = \sum_{\tau=-\infty}^{\infty} s_{j,k,\tau}^2$$

- can estimate $A_{j,k}$ (and hence $\operatorname{cov}\{\hat{\nu}_j^2, \hat{\nu}_k^2\}$) via

$$\hat{A}_{j,k} = \frac{1}{2}\left(\hat{\nu}_j^2\hat{\nu}_k^2 + 2\sum_{\tau=1}^{M_k-1} \hat{s}_{j,\tau}\hat{s}_{k,\tau}\right),$$

where $\hat{s}_{j,\tau}$ is biased estimator of ACVS:

$$\hat{s}_{j,\tau} = \frac{1}{M_j}\sum_{t=L_j-1}^{N-1-\tau} W_{j,t+\tau}W_{j,t}, \qquad 0 \leq \tau \leq M_j - 1$$

Estimation Theory for Wavelet Variance: IV

- under mild conditions, $\hat{\nu}_j^2$ is asymptotically normal with mean ν_j^2 and large-sample variance $2A_{j,j}/M_j$
- to summarize, $\hat{\nu}_j^2$ is an unbiased estimator of ν_j with tractable statistical properties
- in particular, can determine var $\{\hat{\nu}_j^2\}$ and cov $\{\hat{\nu}_j^2, \hat{\nu}_k^2\}$ by substituting readily computable estimates $\hat{A}_{j,k}$ for $A_{j,k}$ in expressions above

Wavelet-Based Definition of Characteristic Scale: I

- interpretation of WV as scaled-based analysis of variance motivates the following definition for CS (basic idea is to consider τ_j's where ν_j^2 is large compared to neighboring values)
- suppose X_t is intrinsically stationary with WVs such that $\nu_j^2 \geq \nu_{j\pm1}^2$ for some $j \geq 2$, with strict inequality holding in at least one case
- fit a quadratic $y_k = a + bx_k + cx_k^2$ passing through $(x_k, y_k) = (\log_2(\tau_k), \log_2(\nu_k^2))$, $k = j-1, j, j+1$
- define a *local CS* as location at which quadratic is maximized:

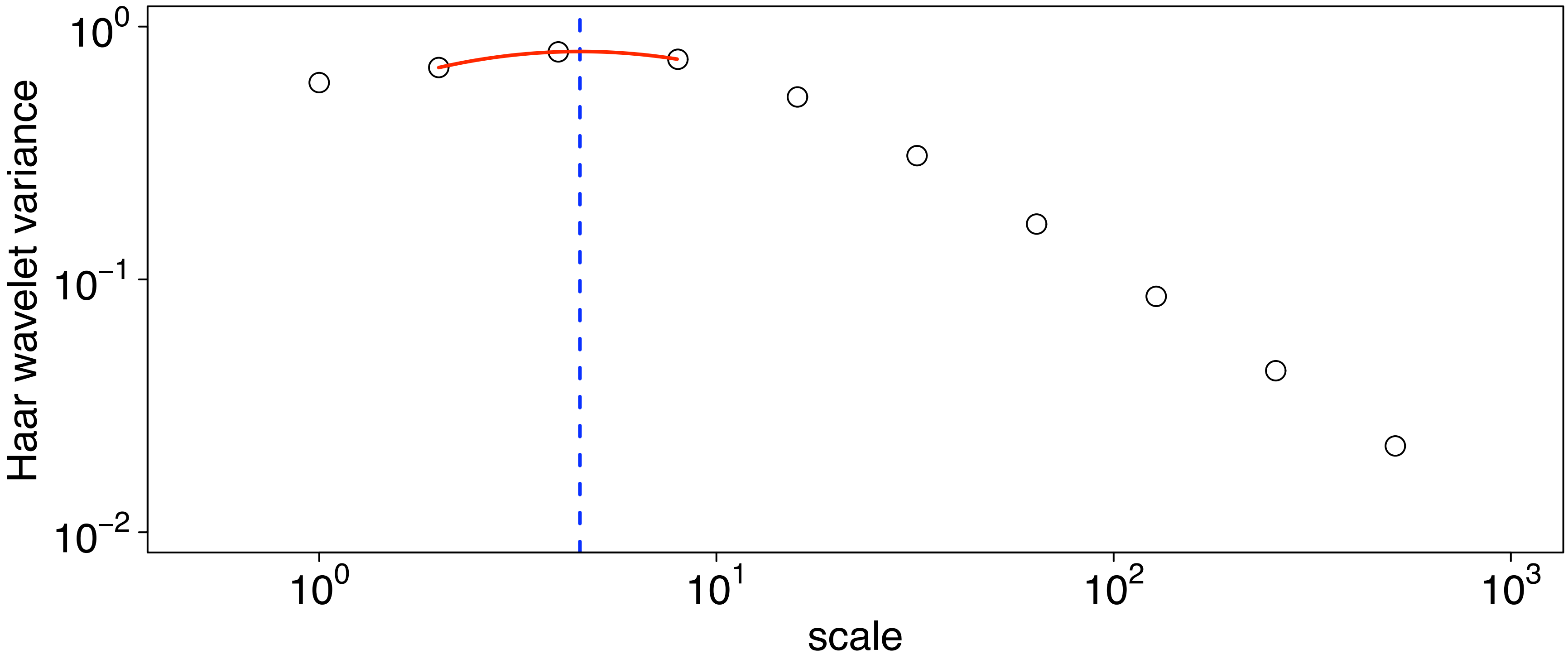
$$\tau_{c,j} = 2^{-\beta_1/\beta_2}\tau_j, \text{ where } \beta_1 = \frac{y_{j+1} - y_{j-1}}{2} \ \& \ \beta_2 = y_{j+1} - 2y_j + y_{j-1}$$

Wavelet-Based Definition of Characteristic Scale: II

- note: $\tau_j/\sqrt{2} \leq \tau_{c,j} \leq \tau_j\sqrt{2}$
- note: $\nu^2_{j-1} < \nu^2_j = \nu^2_{j+1} > \nu^2_{j+2}$ yields $\tau_{c,j} = \tau_{c,j+1} = \tau_j\sqrt{2}$
- in addition, if X_t has local CS $\tau_{c,j}$ such that $\nu^2_j > \nu^2_k$ for all $k \in \mathbb{Z}^+$ (the positive integers) excluding $k = j-1, j, j+1$, then X_t is said to have *global CS* $\tau_c = \tau_{c,j}$
- let's look at some examples of theoretical WV curves possessing local and global CSs

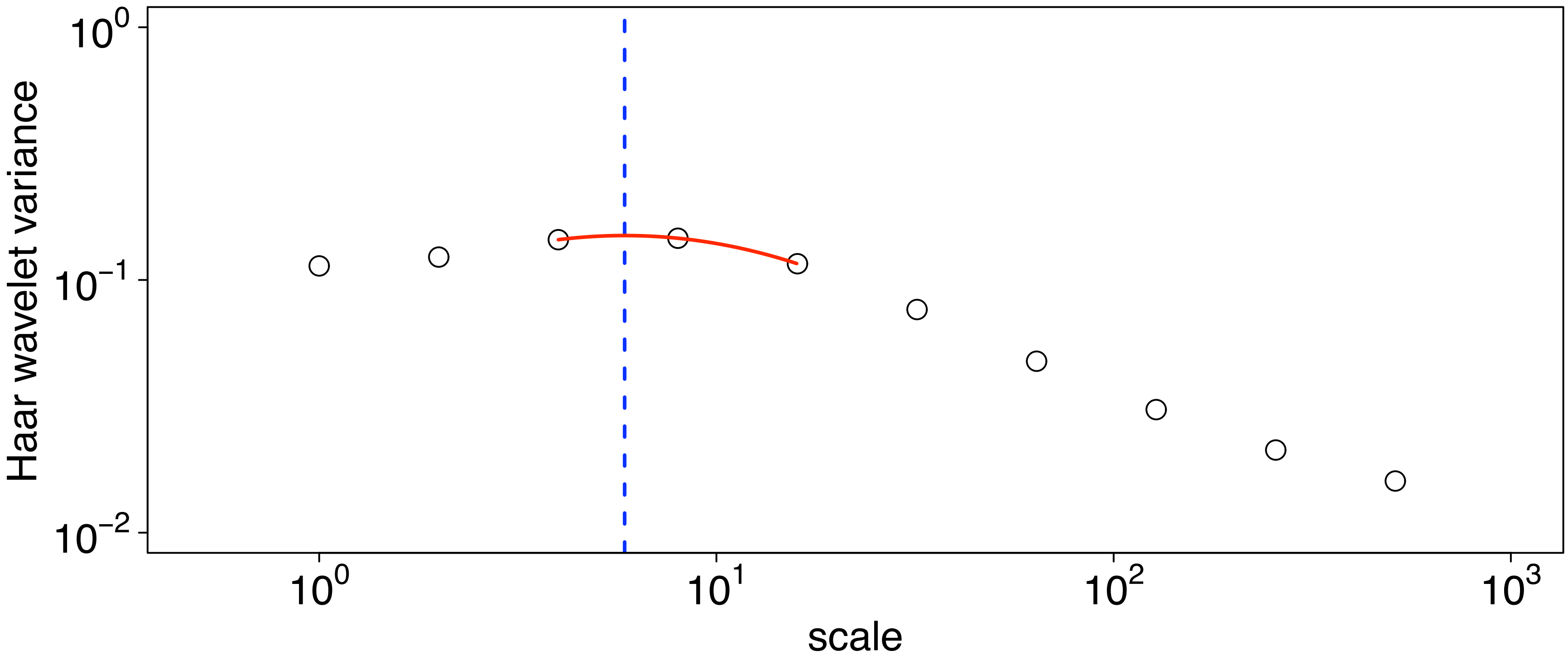
First-Order Autoregressive Process

- AR(1) process defined by $X_t = \phi X_{t-1} + \epsilon_t$, where here $\phi = 0.7$ and ϵ_t is white noise (τ_c and decorrelation time $\tau_D = \frac{1+\phi}{1-\phi}$ similar, with agreement getting better as $\phi \to 1$)



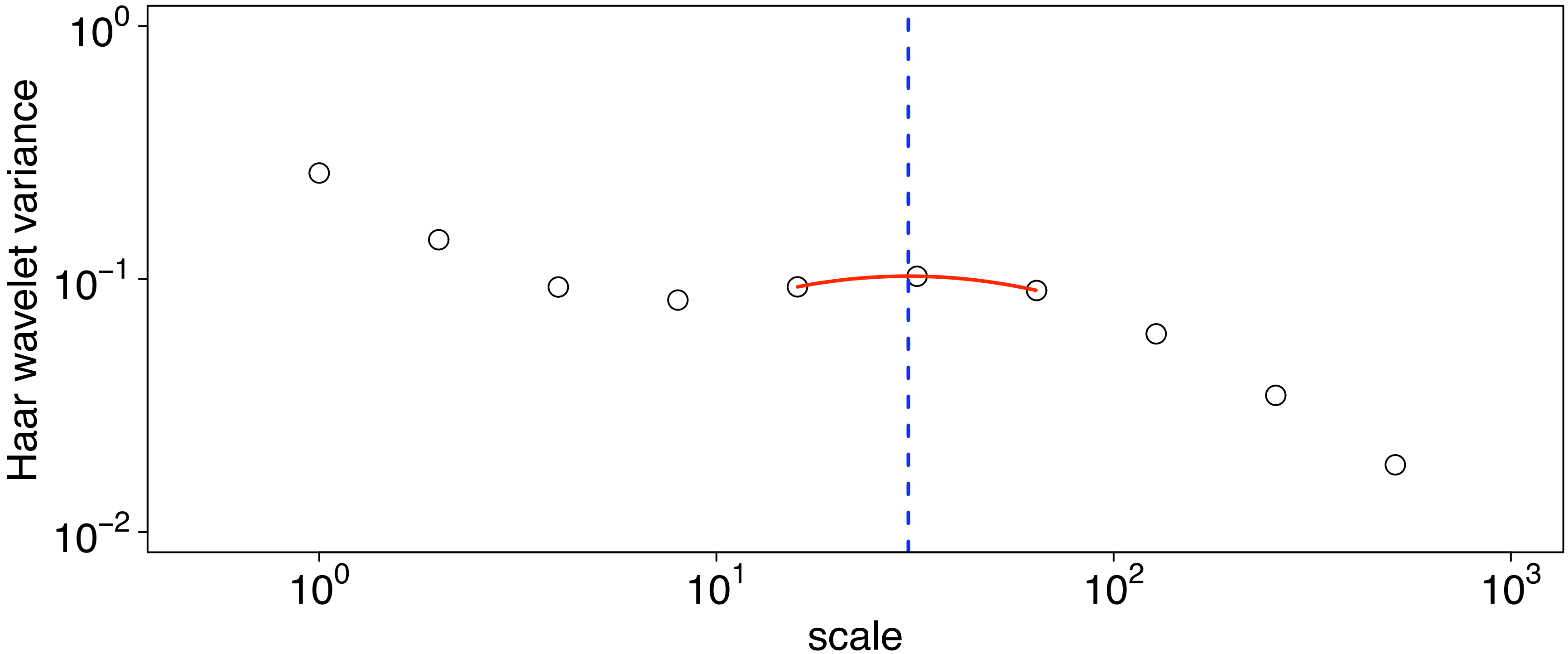
Sum of AR(1) & Fractionally Differenced Processes

- FD process has 'long-memory' – CS measure $\tau_D = \sum_{k \in \mathbb{Z}} \rho_k = \infty$ (reflecting asymptotic decay rate of ρ_k); by contrast, wavelet-based τ_c is finite (concentrates on localized properties of X_t)



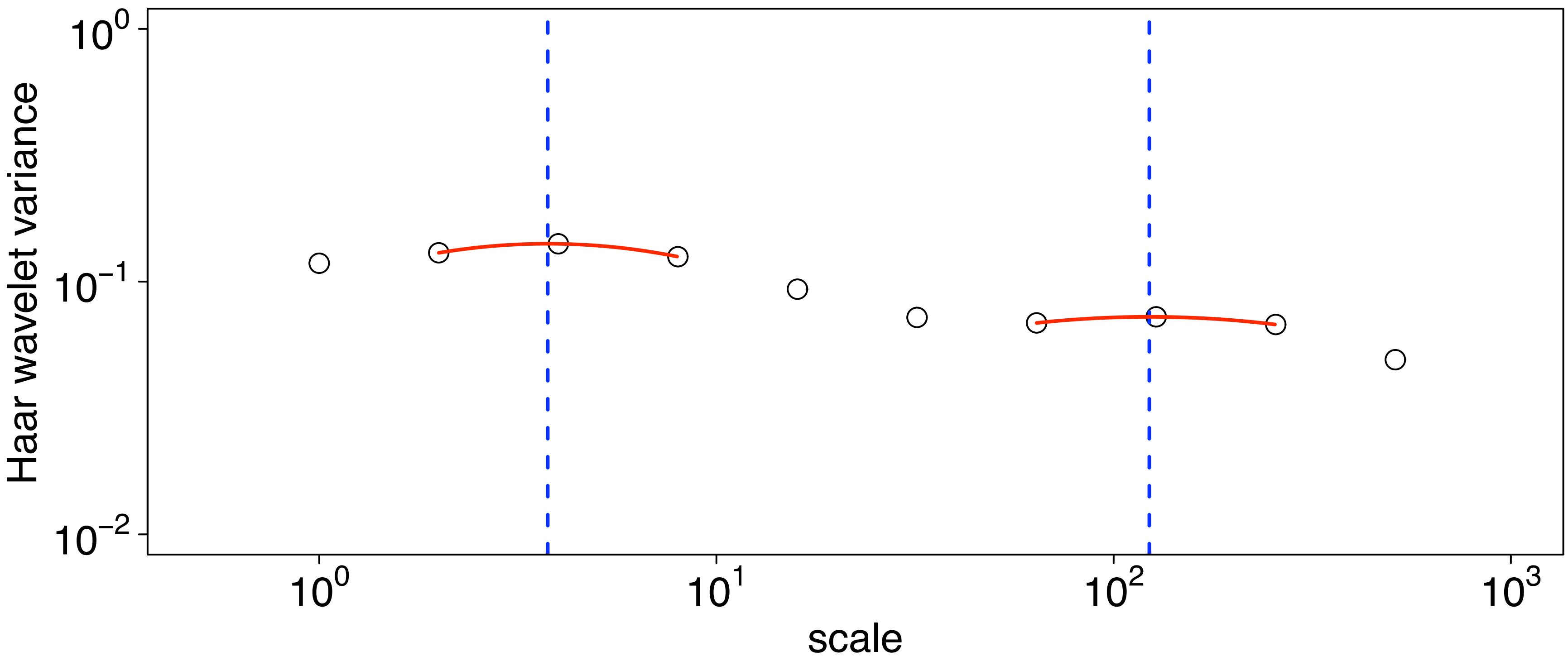
Sum of AR(1) & White Processes

- example of process with a local CS that is not also a global CS



Sum of Two AR(1) Processes

- example of process with two local CSs, one of which is also a global CS; AR(1) parameters are $\phi = 0.65$ and $\phi = 0.99$



Comments on Definition of Wavelet-Based CS

- natural definition of CS would involve WV defined over a continuum of scales from a continuous wavelet transform (CWT)
- for AR(1) processes, CWT-based CS and DWT-based τ_c using Haar wavelet are virtually the same
- DWT-based CS is preferable due to its greater degree of statistical tractability
- τ_c defined in terms of quadratic fit in log/log space, which is preferable to linear/linear, linear/log or log/linear spaces, also because of considerations involving agreement with CWT and statistical tractability

Estimation of Wavelet-Based Characteristic Scale

- suppose we have a time series that is a realization of a portion $X_0, X_1, \ldots, X_{N-1}$ of an intrinsically stationary process of order d
- given an appropriately chosen wavelet filter and an unbiased estimator $\hat{\nu}_j^2$ for levels $j = 1, \ldots, J_0$, suppose there is some $1 < j < J_0$ such that $\hat{\nu}_j^2 \geq \hat{\nu}_{\pm j}^2$ (with strict inequality holding in at least one case)
- conditional on observed pattern and estimating $y_k = \log_2(\nu_k^2)$ using $\hat{y}_k = \log_2(\hat{\nu}_k^2)$, obvious estimator of $\tau_{c,j}$ is

$$\tau_{c,j} = 2^{-\hat{\beta}_1/\hat{\beta}_2}\tau_j, \text{ where } \hat{\beta}_1 = \frac{\hat{y}_{j+1} - \hat{y}_{j-1}}{2} \ \& \ \hat{\beta}_2 = \hat{y}_{j+1} - 2\hat{y}_j + \hat{y}_{j-1}$$

Statistical Properties of CS Estimator: I

- assuming $[\hat{\nu}^2_{j-1}, \hat{\nu}^2_j, \hat{\nu}^2_{j+1}]^T$ is multivariate Gaussian with mean & covariance given by large-sample theory, delta method says

$$\left[\log_2\left(\hat{\nu}^2_{j-1}\right), \log_2\left(\hat{\nu}^2_j\right), \log_2\left(\hat{\nu}^2_{j+1}\right)\right]^T = \left[\hat{y}_{j-1}, \hat{y}_j, \hat{y}_{j+1}\right]^T$$

is approximately Gaussian with mean

$$\left[\log_2\left(\nu^2_{j-1}\right), \log_2\left(\nu^2_j\right), \log_2\left(\nu^2_{j+1}\right)\right]^T = \left[y_{j-1}, y_j, y_{j+1}\right]^T$$

and large-sample covariance matrix Σ whose (m, n)th element is

$$\frac{\operatorname{cov}\{\hat{\nu}^2_{j'+m}, \hat{\nu}^2_{j'+n}\}}{\nu^2_{j'+m}\nu^2_{j'+n}\log^2(2)} + 2\frac{\operatorname{var}\{\hat{\nu}^2_{j'+m}\}\operatorname{var}\{\hat{\nu}^2_{j'+n}\} + (\operatorname{cov}\{\hat{\nu}^2_{j'+m}, \hat{\nu}^2_{j'+n}\})^2}{\nu^4_{j'+m}\nu^4_{j'+n}\log^2(2)};$$

here $j' = j - 1$ and $0 \le (m, n) \le 2$

Statistical Properties of CS Estimator: II

- note that

$$\begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} & 0 & \frac{1}{2} \\ 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} \hat{y}_{j-1} \\ \hat{y}_j \\ \hat{y}_{j+1} \end{bmatrix} \equiv H \begin{bmatrix} \hat{y}_{j-1} \\ \hat{y}_j \\ \hat{y}_{j+1} \end{bmatrix}$$

- hence $[\hat{\beta}_1, \hat{\beta}_2]^T$ is asymptotically Gaussian with mean $[\beta_1, \beta_2]^T$ and covariance $H\Sigma H^T$
- letting $\hat{\kappa} = -\hat{\beta}_1/\hat{\beta}_2$, delta method says $\hat{\kappa}$ is asymptotically Gaussian with mean $-\beta_1/\beta_2$ and large-sample variance

$$\sigma^2_{\hat{\kappa}} = \frac{\operatorname{var}\{\hat{\beta}_1\}}{\beta_2^2} + \frac{\beta_1^2 \operatorname{var}\{\hat{\beta}_2\}}{\beta_2^4} + \frac{\operatorname{var}\{\hat{\beta}_1\}\operatorname{var}\{\hat{\beta}_2\} + 2(\operatorname{cov}\{\hat{\beta}_1, \hat{\beta}_2\})^2}{\beta_2^4} + \frac{3\beta_1^2(\operatorname{var}\{\hat{\beta}_2\})^2}{\beta_2^6} - \frac{2\beta_1 \operatorname{cov}\{\hat{\beta}_1, \hat{\beta}_2\}}{\beta_2^3}$$

Statistical Properties of CS Estimator: III

- final use of delta method says $\hat{\tau}_{c,j}$ is asymptotically normal with mean $\tau_{c,j}$ and large-sample variance $\tau^2_{c,j}\,\sigma^2_{\hat{\kappa}}\,\log^2_e(2)$
- approximate 95% confidence interval (CI) for $\tau_{c,j}$ given by

$$\left[2^{-1.96\sigma_{\hat{\kappa}}}\hat{\tau}_{c,j},\, 2^{1.96\sigma_{\hat{\kappa}}}\hat{\tau}_{c,j}\right]$$

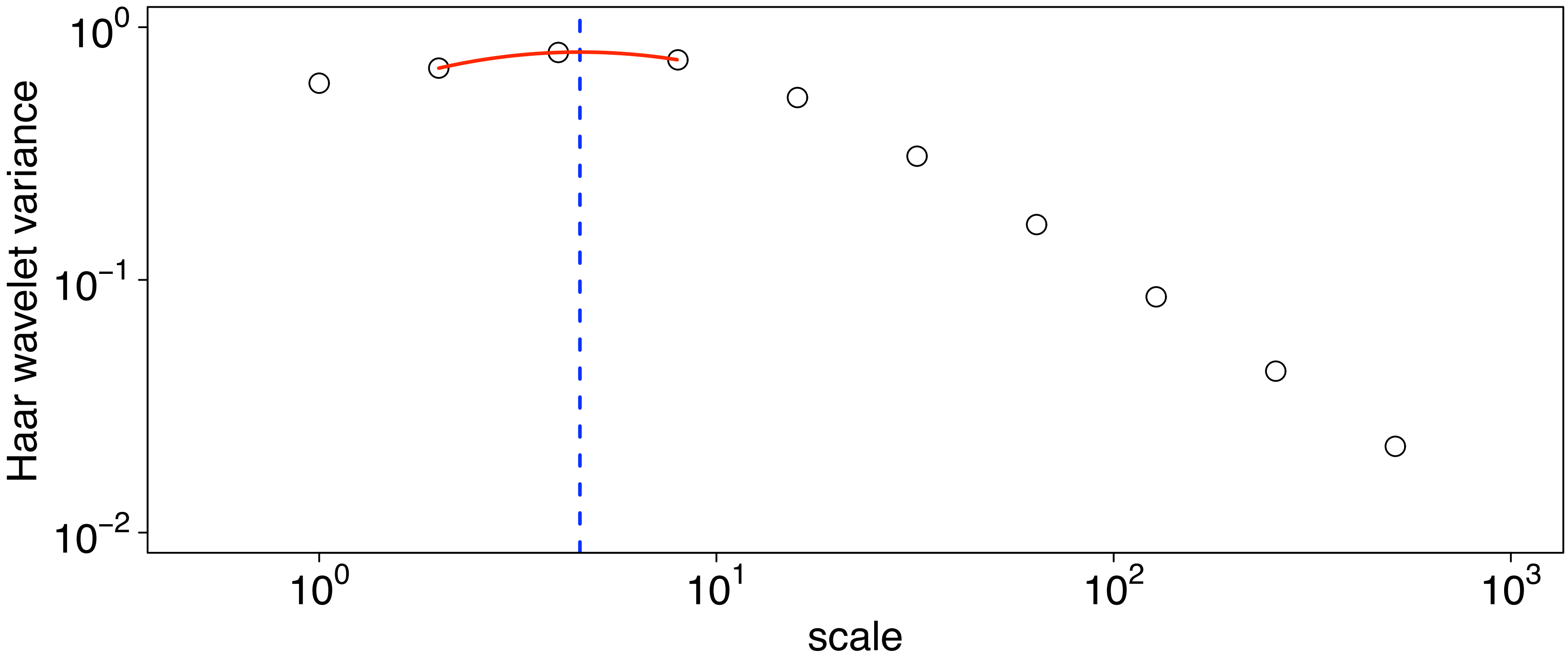
- in practical applications, can use 'plug-in' estimate for $\sigma^2_{\hat{\kappa}}$
- caveat: approach conditioned upon $\hat{\nu}^2_j \geq \hat{\nu}^2_{j\pm1}$ correctly indicating the presence of a local CS (asymptotically OK)
- sanity check: generate realizations of trivariate Gaussian with mean $[\hat{\nu}^2_{j-1}, \hat{\nu}^2_j, \hat{\nu}^2_{j+1}]^T$ & covariance matrix given by large sample theory, and look at proportion of realizations with maximum at j (if large, have some faith observed pattern actually exists)

Monte Carlo Experiments: I

- for each of four processes and for sample sizes N ranging from 256 to 8192, generated 1000 realizations using exact simulation method for AR and FD processes
- recorded number of realizations M for which peak there was a peak in Haar WV curve at either proper level j or $j \pm 1$
- for each of these M realizations, estimated CS and computed 95% CI using plug-in procedure
- following tables show average of estimated CSs and percentage of times CIs trapped true CS

Monte Carlo Experiments: II

- AR(1) process ($\phi = \rho_1 = 0.7$)



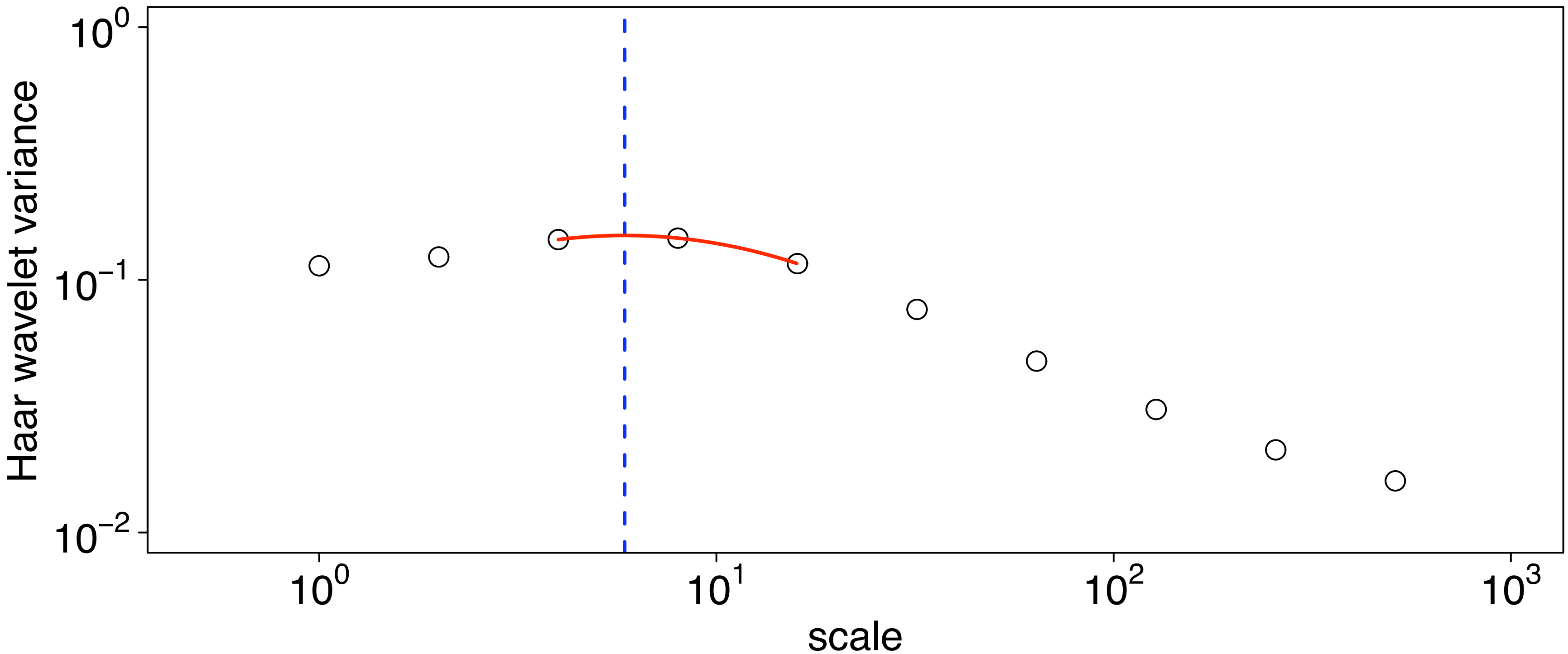
Monte Carlo Experiments: III

- results for AR(1) process ($\phi = \rho_1 = 0.7$)

N	τ_c	$\hat{\tau}_c$	M	% coverage
256	4.53	4.75	953	88.6
512		4.69	992	88.2
1024		4.68	999	87.9
2048		4.66	1000	87.1
4096		4.60	1000	90.0
8192		4.57	1000	94.4

Monte Carlo Experiments: IV

- sum of AR(1) & FD processes ($\phi = 0.75$ & $\delta = 0.45$)



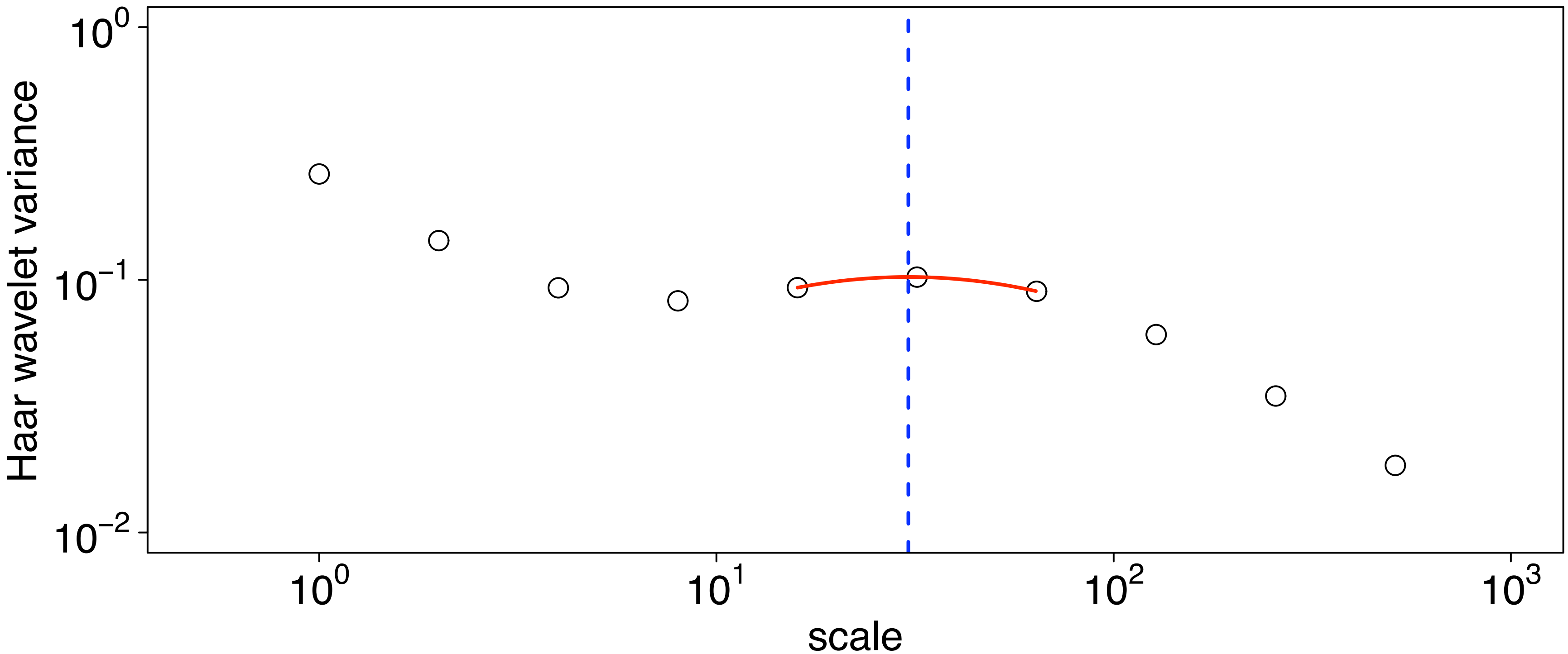
Monte Carlo Experiments: V

- results for sum of AR(1) & FD processes ($\phi = 0.75$ & $\delta = 0.45$)

N	τ_c	$\hat{\tau}_c$	M	% coverage
256	5.87	6.58	942	89.2
512		6.16	981	89.2
1024		5.88	1000	90.8
2048		5.84	1000	90.5
4096		5.81	1000	92.6
8192		5.83	1000	93.8

Monte Carlo Experiments: VI

- sum of AR(1) & white noise processes ($\phi = 0.95$)



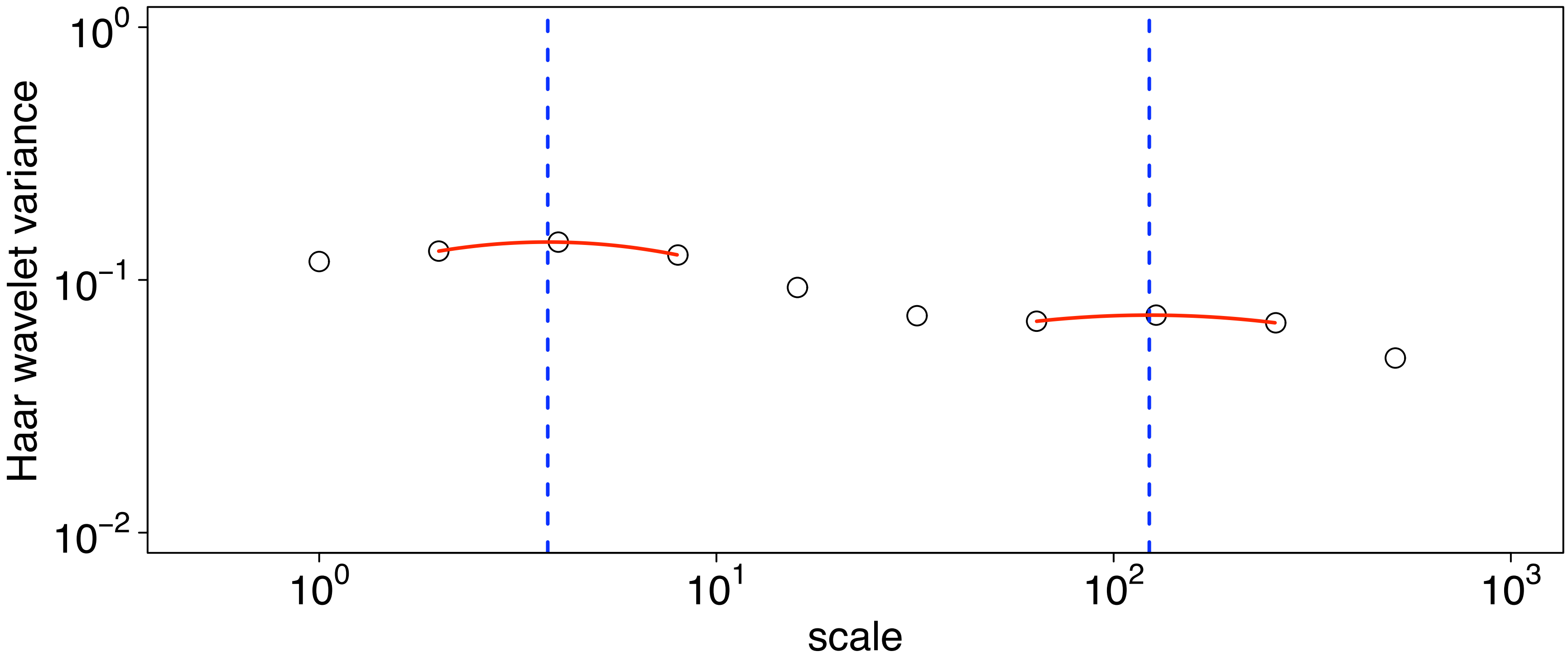
Monte Carlo Experiments: VII

- results for sum of AR(1) & white noise processes ($\phi = 0.95$)

N	τ_c	$\hat{\tau}_c$	M	% coverage
256	30.42	33.51	798	87.0
512		33.51	838	86.8
1024		33.51	900	89.1
2048		32.49	964	88.5
4096		31.73	999	91.4
8192		31.41	1000	93.1

Monte Carlo Experiments: VIII

- sum of two AR(1) processes ($\phi = 0.65$ & 0.99)



Monte Carlo Experiments: IX

- first CS of sum of two AR(1) processes ($\phi = 0.65$ & 0.99)

N	τ_c	$\hat{\tau}_c$	M	% coverage
256	3.76	4.08	935	90.4
512		4.00	971	91.0
1024		3.87	997	92.6
2048		3.84	999	96.7
4096		3.80	1000	97.2
8192		3.76	1000	96.1

Monte Carlo Experiments: X

- second CS of sum of two AR(1) processes ($\phi = 0.65$ & 0.99)

N	τ_c	$\hat{\tau}_c$	M	% coverage
256	122.96	—	—	—
512		89.31	454	86.8
1024		139.31	690	88.0
2048		149.60	703	88.9
4096		152.97	757	89.8
8192		153.43	775	86.7

- note: $8192/\tau_c \doteq 67$; i.e., small number of replications of CS

Monte Carlo Experiments: XI

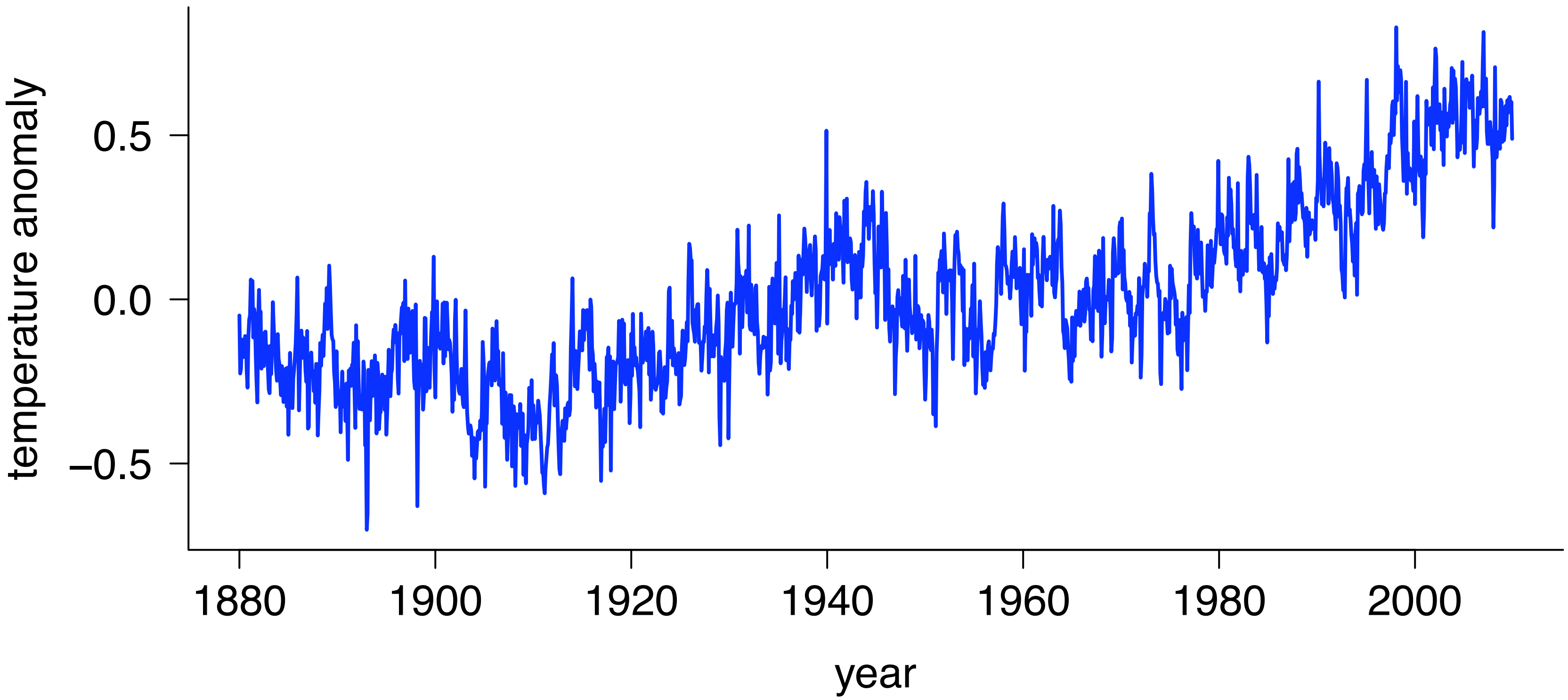
- tendency for $\hat{\tau}_c$ to be positively biased
- closeness of coverage percentage to nominal 95% tends to depend on true τ_c: the smaller τ_c is, the better the coverage rate
- for small sample sizes, cover rate tends to be below 95%
- coverage rates tend to improve with increasing sample size, as asymptotic theory would suggest
- bottom line: experiments indicate large-sample theory gives useful – but admittedly not perfect – approximations to variability in $\hat{\tau}_c$ for moderate sample sizes
- note: similar results hold in experiments using Daubechies wavelet filters of lengths $L_1 = 4$ and 8

Four Real-World Examples

- global temperature record
- coherent structures in river flow
- Madden–Julian atmospheric oscillation
- medium multiyear Arctic sea ice

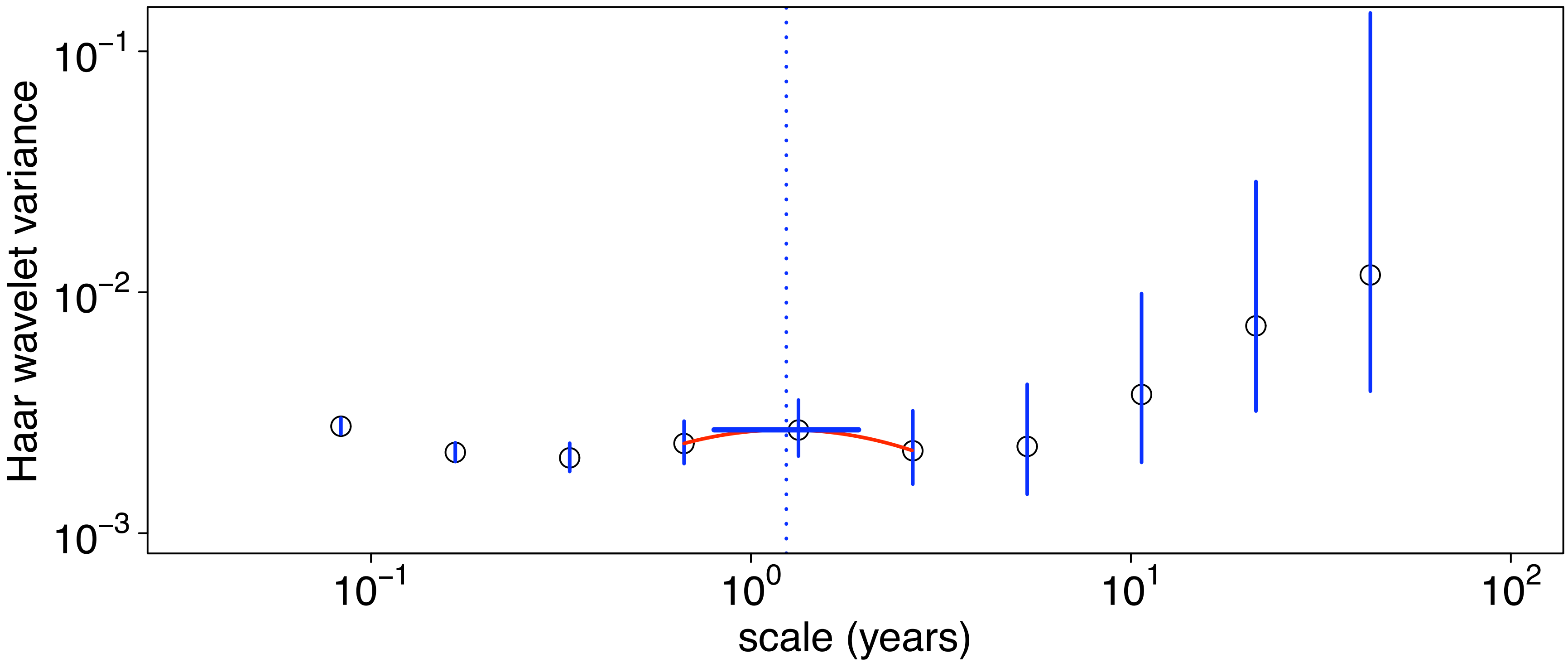
Global Temperature Record: I

- monthly global temperature anomalies (land and water combined) from Jan. 1880 to Dec. 2009 ($N = 1560$; data obtained from NOAA Web site); note prominent trend upwards



Global Temperature Record: II

- $\hat{\tau}_{c,5} \doteq 14.9$ months with associated 95% CI of $[9.6, 23.0]$

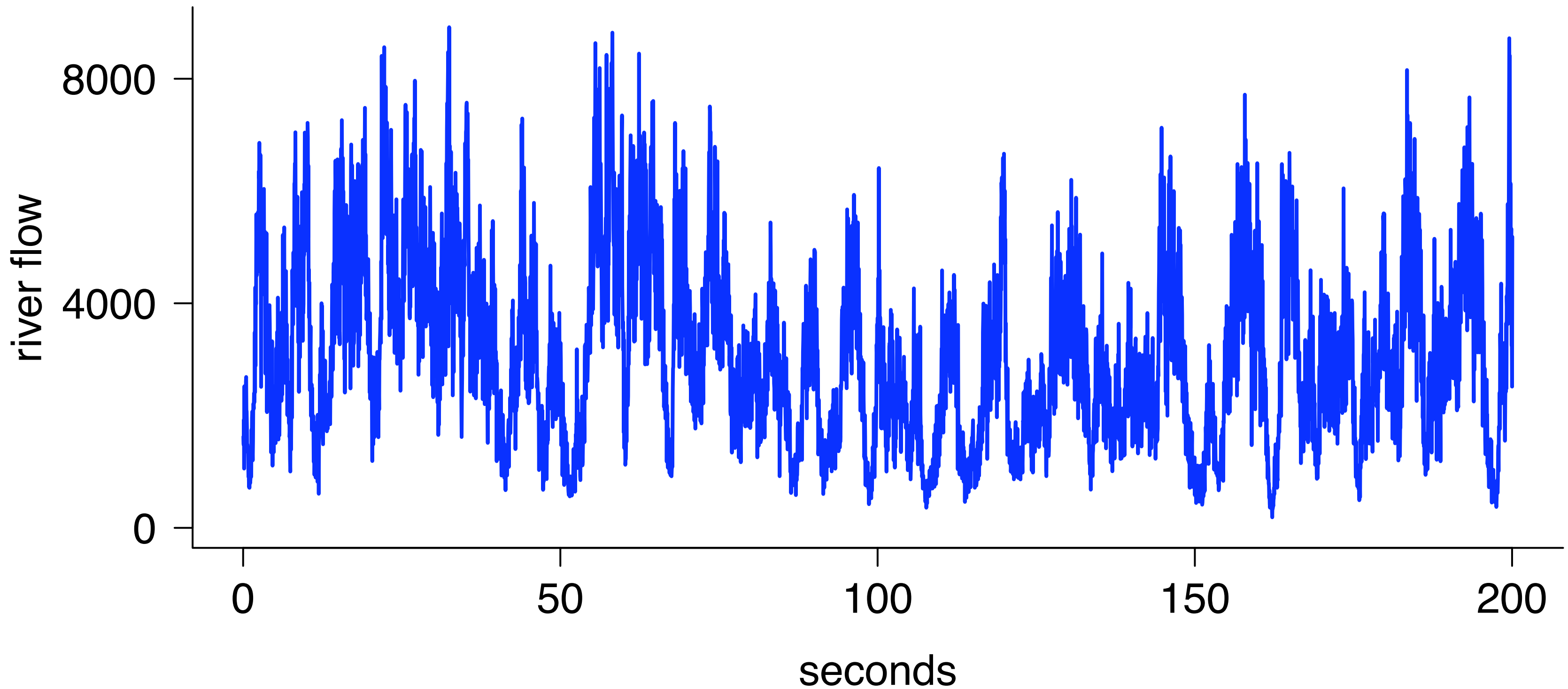


Global Temperature Record: III

- Tsonis et al. (1998) detrended series using singular spectrum analysis and used residuals to estimate a CS of 20 months via involved procedure based on cumulative sums
- interpreted CS as being due to influence of El Niño/La Niña cycles on global temperatures
- 95% CI for wavelet-based CS traps 20 months, but no need for explicit detrending due to differencing operations embedded within wavelet filters (eliminates trends that can be modelled by low-order polynomials, as verified here by use of length $L_1 = 8$ wavelet filter)

Coherent Structures in River Flows: I

- first 5000 values from time series of length $N = 29972$ sampled every $\frac{1}{25}$ sec. from Univ. of WA velocity profiler set on bottom of river estuary downstream from sill pointing upwards

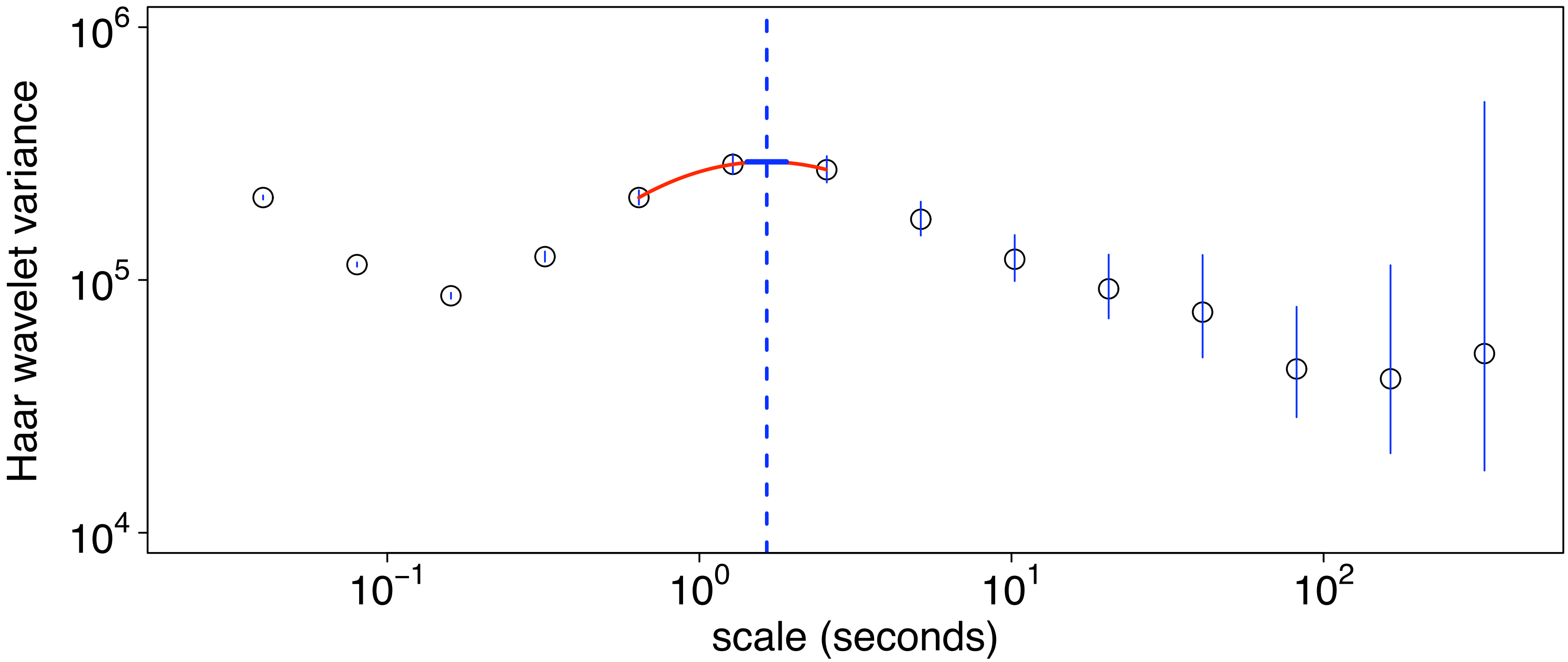


Coherent Structures in River Flows: II

- videos of river surface clearly show quasi-periodic upwellings from river appearing as temporary ‘blobs’ (coherent structures), with each blob dissipating within a second or so, with another one forming immediately afterwards
- quantifying this little-understood phenomenon using standard Fourier-based spectral analysis problematic because it appears as a small perturbation in a low-frequency rolloff
- by contrast, as shown on next overhead, WV clearly displays a peak, rendering phenomenon as interpretable in terms of a CS
- time-evolving properties of blobs can be studied by estimating CSs for portions of time series spanning successive 20-minute time intervals

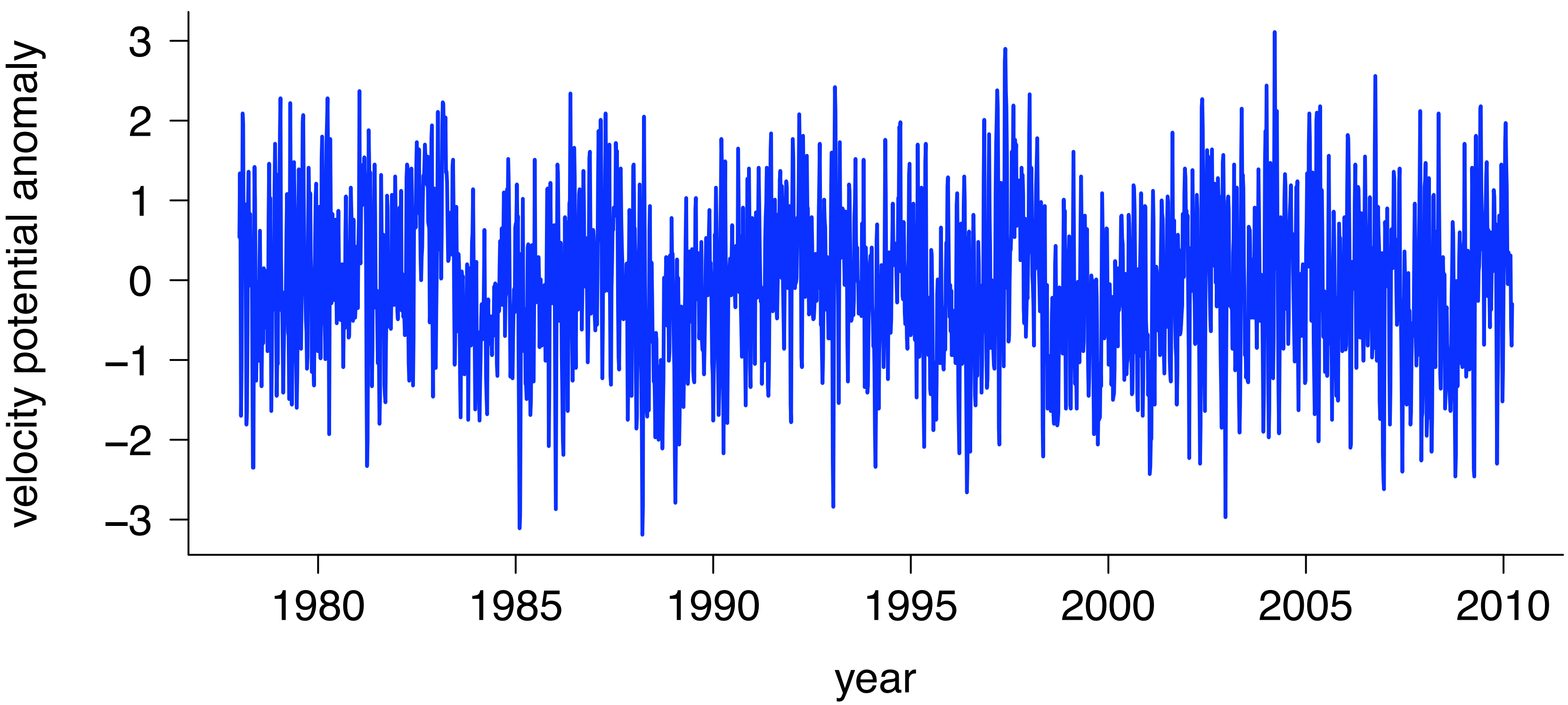
Coherent Structures in River Flows: III

- $\hat{\tau}_{c,6} = 1.6$ seconds with associated 95% CI of $[1.4, 1.9]$



Madden–Julian Atmospheric Oscillation: I

- 200 hectopascal velocity potential anomalies, a Madden–Julian oscillation (MJO) index from NOAA Web site ($N = 2354$ values covering 32+ years with sampling interval of 5 days)

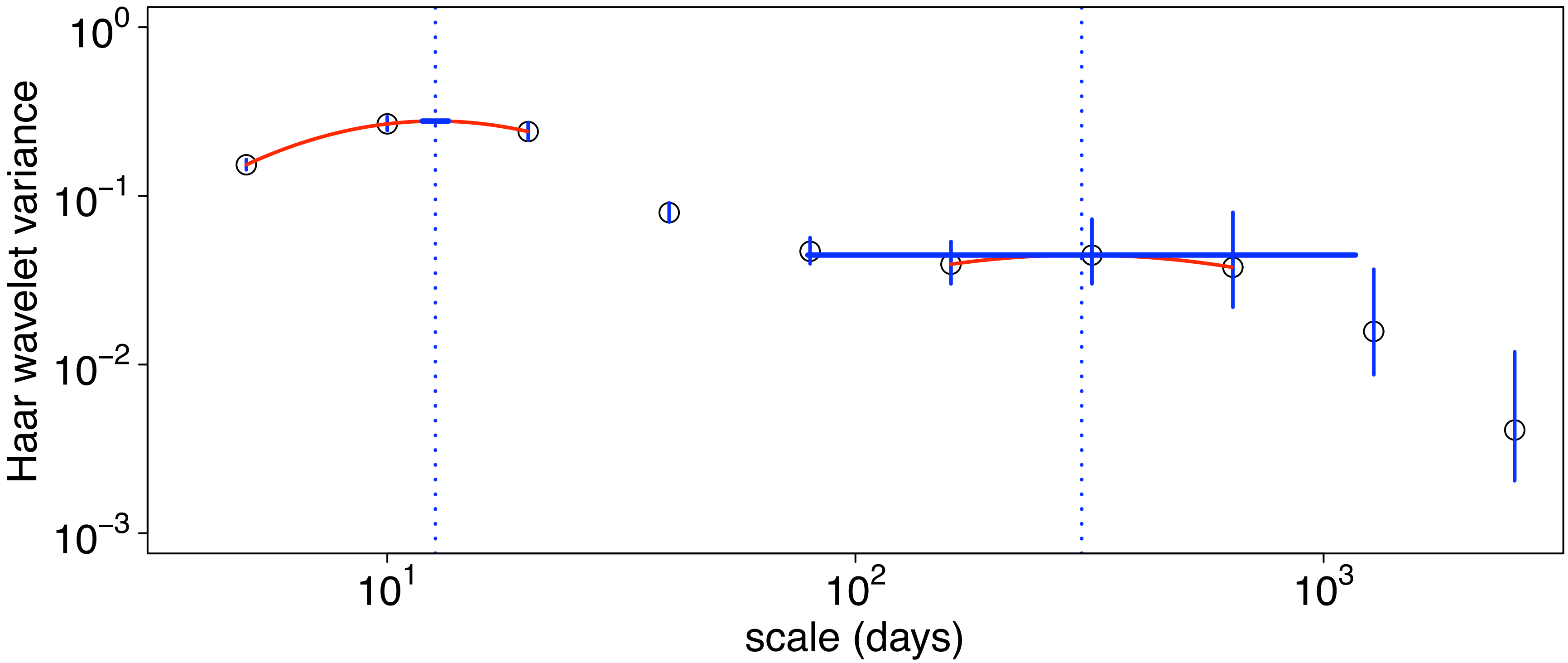


Madden–Julian Atmospheric Oscillation: II

- Madden and Julian (1994) originally discussed 40–50 day oscillation appearing in various atmospheric time series collected in tropics
- MJO now called 30–60 day or intraseasonal oscillation

Madden–Julian Atmospheric Oscillation: III

- $\hat{\tau}_{c,2} = 12.7$ days with associated 95% CI of $[11.9, 13.5]$

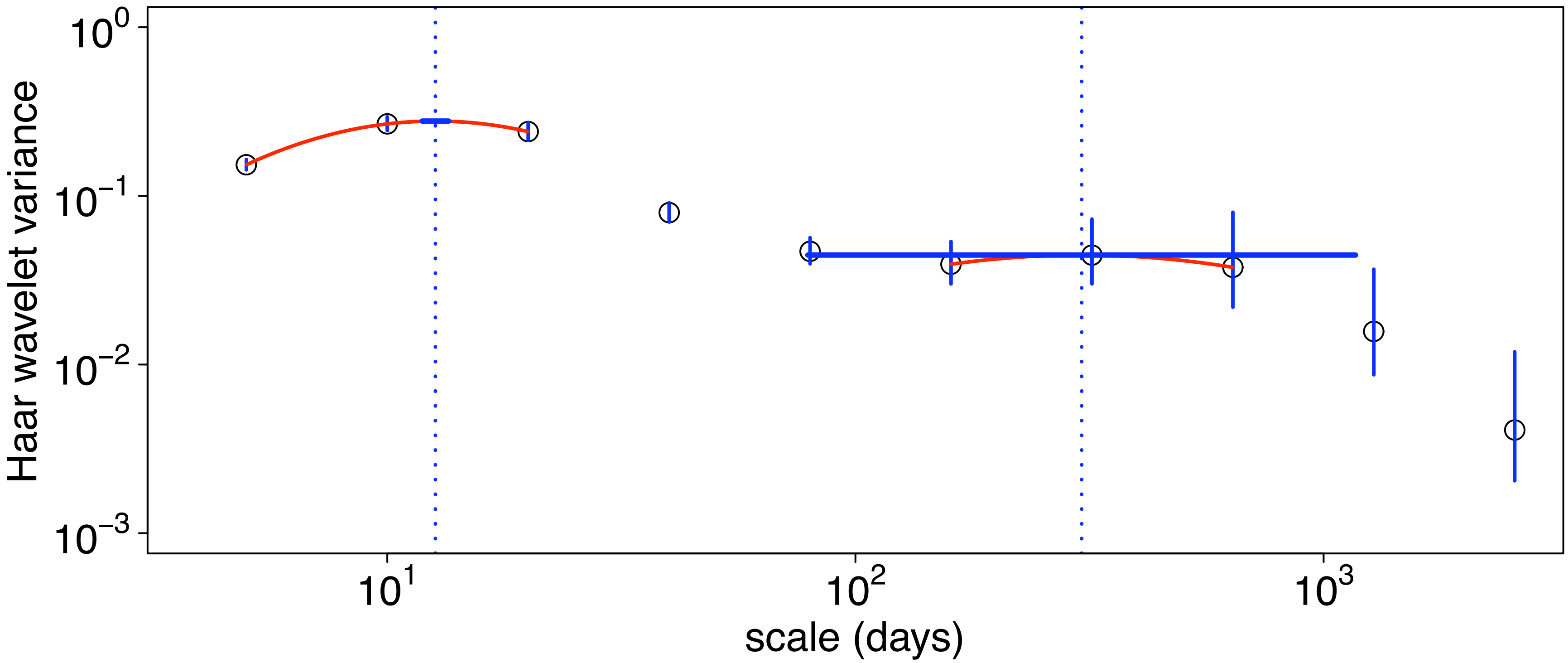


Madden–Julian Atmospheric Oscillation: IV

- since scale of τ is associated with periods in interval $[2\tau, 4\tau]$, CS point estimate of 12.7 days matches up with 25–51 day oscillations and hence with description of MJO as a 30–60 day oscillation
- deduction of MJO from Fourier-based spectral analysis problematic due to lack of standard way to determine beginning/end of frequency interval associated with broadband oscillations
- notion of CS bypasses this difficulty, opening up means of objectively tracking how MJO varies across time and over different time series

Madden–Julian Atmospheric Oscillation: V

- $\hat{\tau}_{c,7} \doteq 304$ days with associated 95% CI of $[79, 1170]$

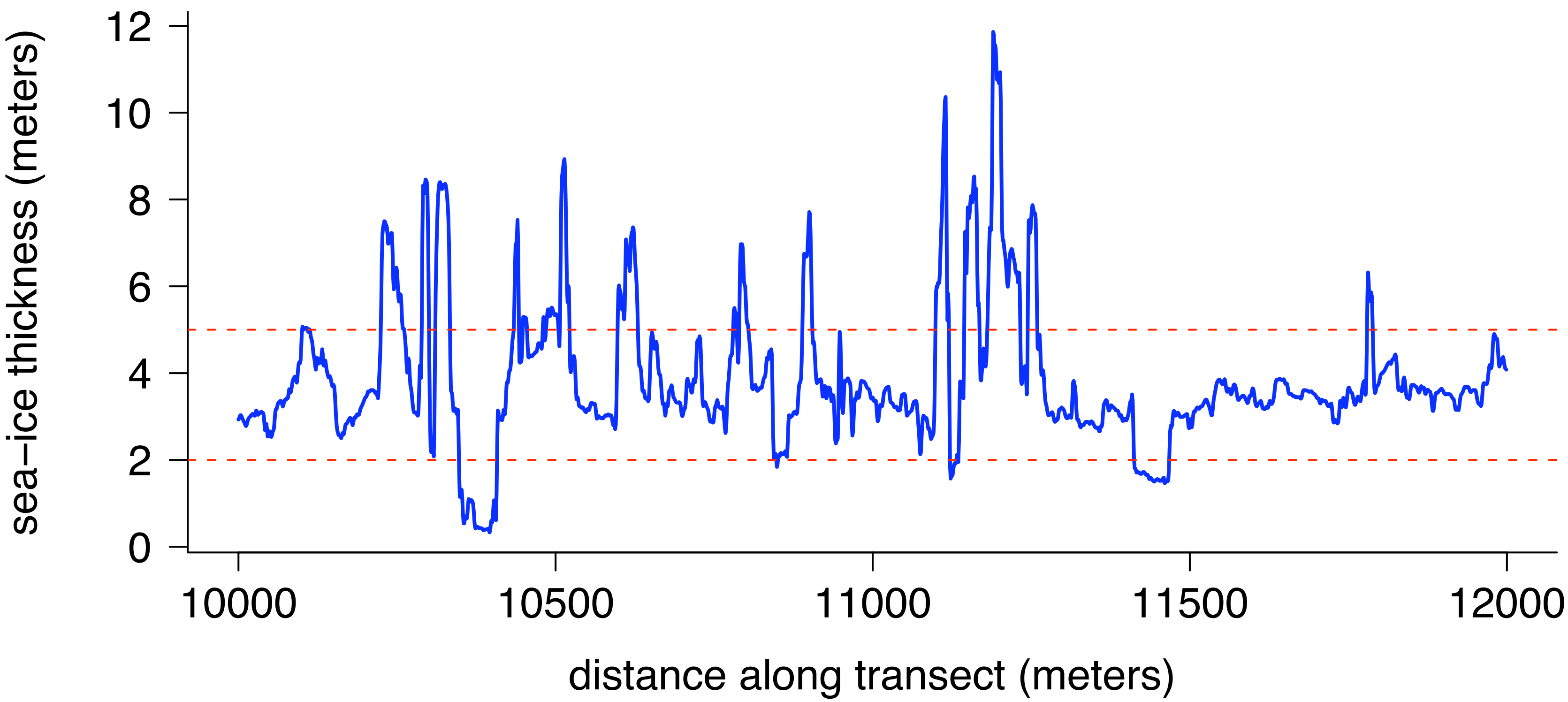


Madden–Julian Atmospheric Oscillation: VI

- interval of periods associated with second CS suggests an oscillation spanning two to three years that is about 5 times weaker than MJO
- presence of this weak CS is conditional upon peak pattern in WV being correct
- as example of use of sanity check, generated $100,000$ realizations of trivariate Gaussian with mean $[\hat{\nu}_6^2, \hat{\nu}_7^2, \hat{\nu}_8^2]^T$ & covariance matrix given by large sample theory
- of these realizations, 60% obeyed observed $\hat{\nu}_6^2 \leq \hat{\nu}_7^2 \geq \hat{\nu}_8^2$ pattern, but remaining 40% did not, casting considerable doubt on validity of observed peak pattern
- similar test on MJO-related CS yielded $99,916$ realizations with observed peak pattern and only 84 without

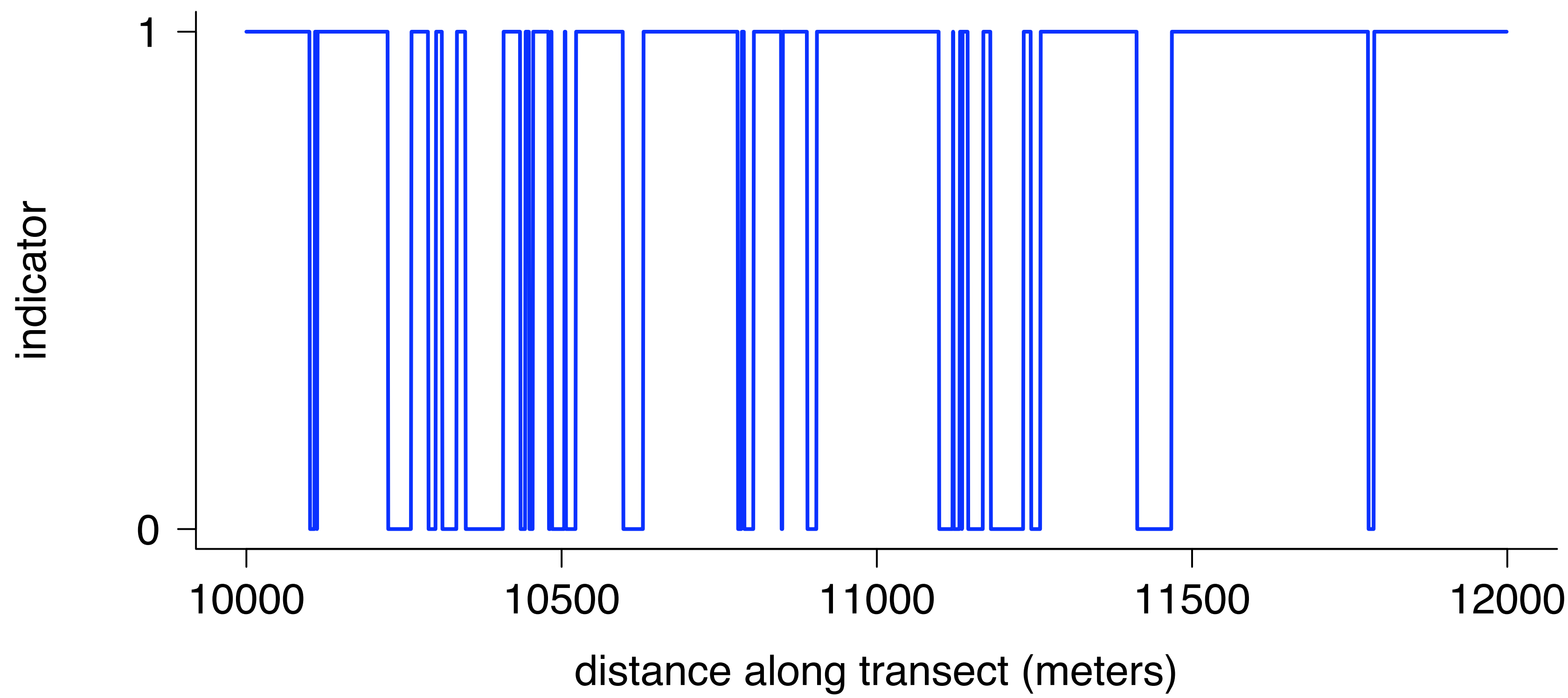
Medium Multiyear Arctic Sea Ice: I

- part of ice thickness measurements spaced a meter apart from a 50 km track collected using an upward-looking sonar on a submarine (data from U.S. National Snow and Ice Data Center)



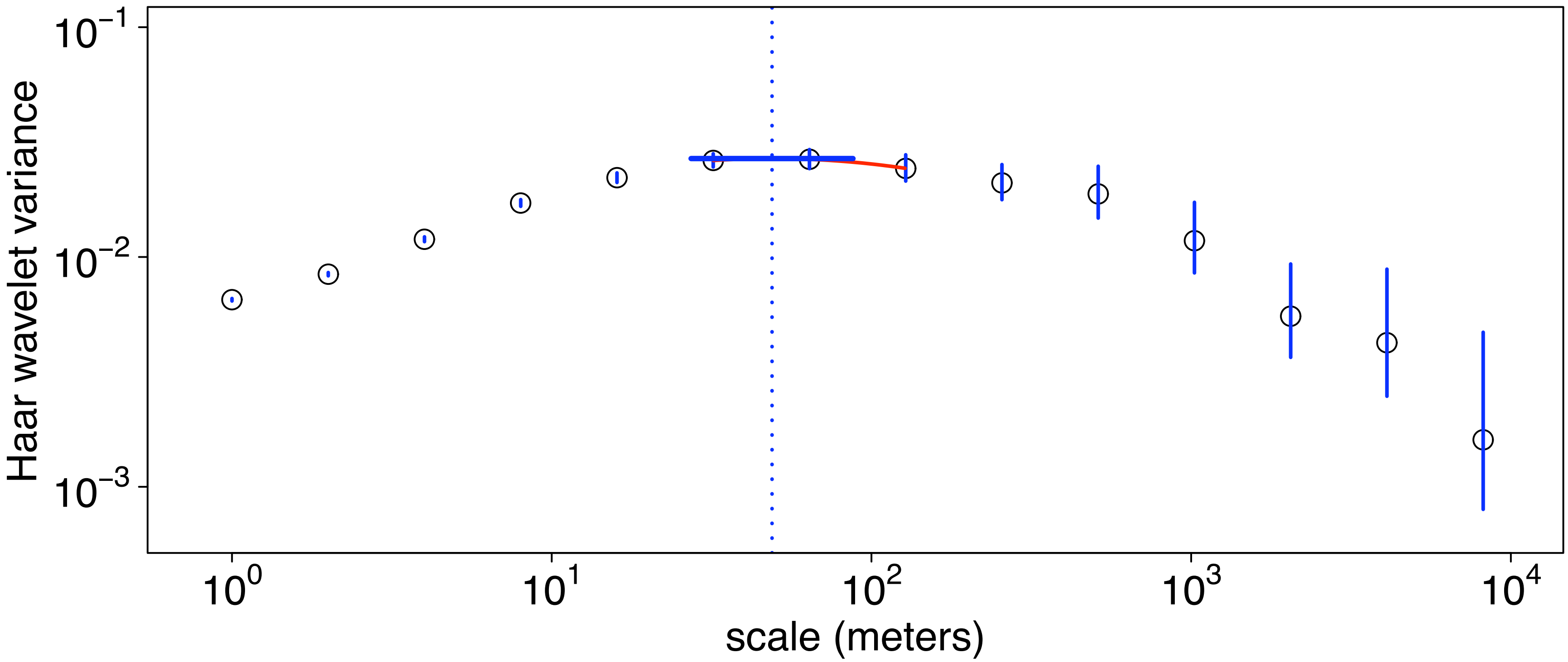
Medium Multiyear Arctic Sea Ice: II

- binary series indicating presence (1) or absence (0) of medium multiyear Arctic sea ice, which by definition has a thickness between 2 and 5 m



Medium Multiyear Arctic Sea Ice: III

- $\hat{\tau}_{c,7} \doteq 48.9$ m with associated 95% CI of $[27.3, 87.5]$



Medium Multiyear Arctic Sea Ice: IV

- CS is indicator of 'typical' extent of medium multiyear ice
- in face of other evidence that Arctic climate is dramatically changing, question of considerable interest is how stable CSs for different ice types are both spatially and temporally
- note: evidence of long-range dependence in ice thickness measurements implies long-range dependence in indicator series also, meaning that decorrelation time τ_D would be infinite for the medium multiyear ice indicators, rendering it useless as a measure of CS
- by contract, wavelet-based CS is finite and provides a useful summary of one aspect of indicator series

Concluding Remarks: I

- CS based upon peaks in WV curves has certain advantages over other definitions for CS, including
  - use of well-defined scale-based decomposition of variance afforded by DWT
  - ability to focus on localized properties of process rather than asymptotic decay rates of autocorrelation sequences
  - ability to handle certain nonstationary processes
  - ability to handle series with trends well approximated by low-order polynomials
  - availability of tractable large-sample theory that is applicable to time series of moderate size

Concluding Remarks: II

- avenues for future research
  - extension to non-Gaussian process (including better handling of indicator series)
  - handling irregularly sampled time series
  - extension to two-dimensional data, with potential applications in medical image processing (telemedicine?)

Thanks to . . .

- Fred Godtliebsen, Kevin Thon and Marc Geilhufe for arranging and facilitating my visit
- National Science Foundation (USA) for support under grant # ARC 0529955 (any opinions, findings and conclusions or recommendations expressed in this talk are those of the authors and do not necessarily reflect the views of the National Science Foundation)

References

- Keim, M. J. and Percival, D. B. (2011), 'Assessing Characteristic Scales Using Wavelets,' submitted
- Madden, R. A. and Julian, P. R. (1994) Observations of the 40–50-day tropical oscillation – a review. *Monthly Weather Review*, **122**, pp. 814–837.
- Tsonis, A. A., Roebber, P. J. and Elsner, J. B. (1998) A characteristic time scale in the global temperature record. *Geophysical Research Letters*, **25**, pp. 2821–2823
- von Storch, H. and Zwiers, F. W. (1999) *Statistical Analysis in Climate Research*. Cambridge: Cambridge University Press