An Introduction to the Wavelet Variance and Its Statistical Properties

Don Percival

Applied Physics Laboratory
University of Washington, Seattle

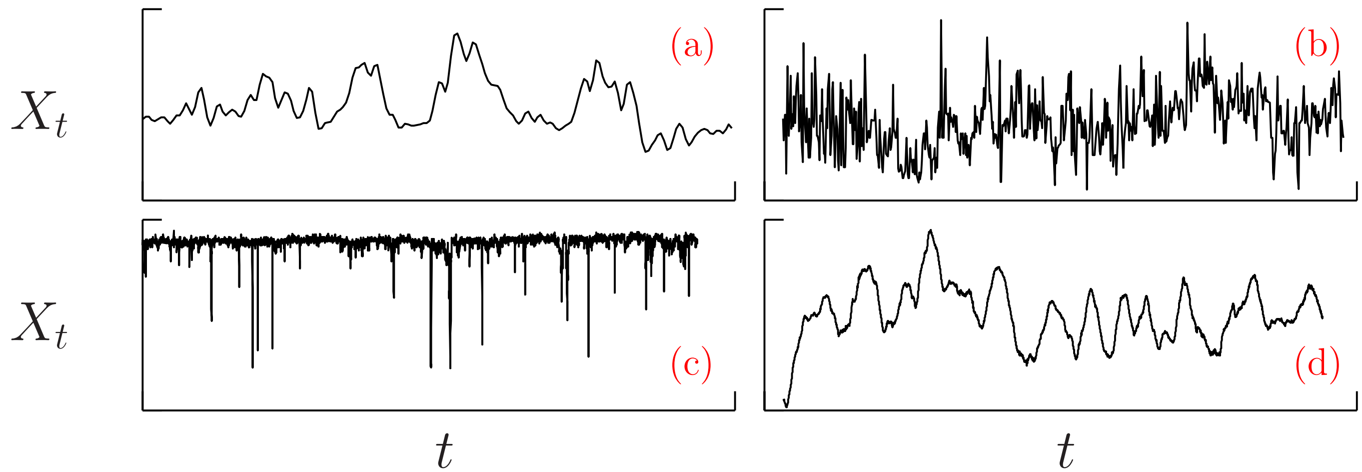
overheads for talk available at

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

Overview

- examples of time series to motivate discussion
- wavelet filters, wavelet coefficients & their interpretation
- decomposition of sample variance using wavelets
- theoretical wavelet variance for stochastic processes
  - stationary processes
  - nonstationary processes with stationary differences
- sampling theory for Gaussian processes with an example
- sampling theory for non-Gaussian processes with an example
- use on time series with time-varying statistical properties
- extensions: covariances, biased estimators, gappy series, fields
- summary

Examples: Time Series X_t Versus Time Index t



(a) subtidal sea levels (2 observations each day, $N = 192$)

(b) Nile River minima (annual, $N = 663$)

(c) surface albedo of arctic ice (25 meters, $N = 8428$)

(d) vertical shear in the ocean (0.1 meters, $N = 4096$)

- four series are visually different
- goal of time series analysis is to quantify these differences

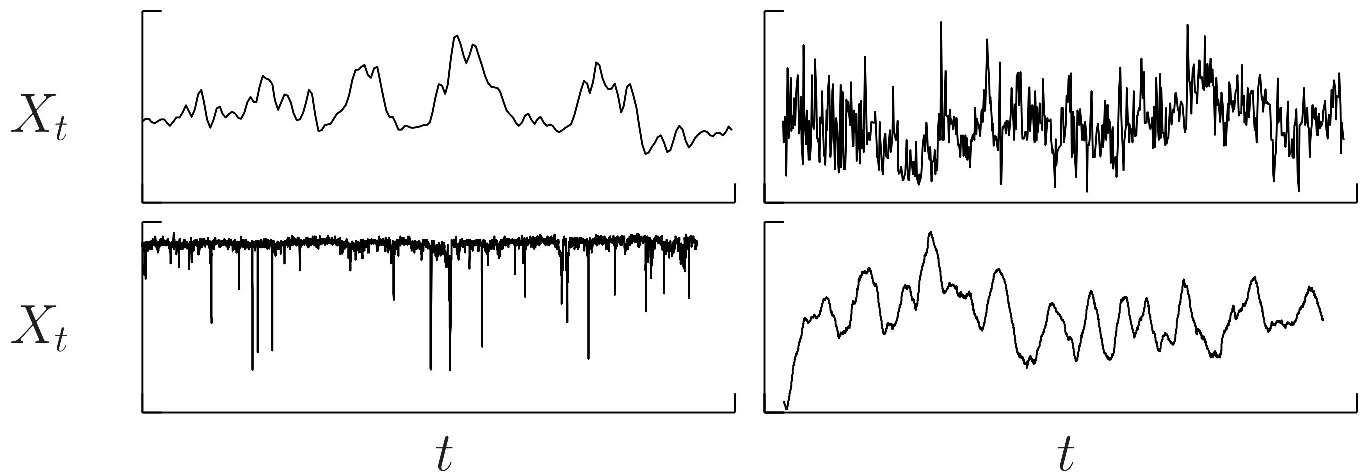
Decomposing Sample Variance of Time Series

- one approach: quantify differences by analysis of variance
- let $X_0, X_1, \ldots, X_{N-1}$ represent time series with N values
- let $\overline{X}$ denote sample mean of X_t's: $\overline{X} \equiv \frac{1}{N} \sum_{t=0}^{N-1} X_t$
- let $\hat{\sigma}_X^2$ denote sample variance of X_t's:

$$\hat{\sigma}_X^2 \equiv \frac{1}{N} \sum_{t=0}^{N-1} \left(X_t - \overline{X}\right)^2$$

- idea is to decompose (analyze, break up) $\hat{\sigma}_X^2$ into pieces that quantify how time series are different
- wavelet variance does analysis based upon differences between (possibly weighted) adjacent averages over 'scales'

Examples Revisited: Notion of Scale



- scale τ refers to the width of a time interval
- scale-based analysis looks at averages over intervals of width τ:

$$\overline{X}_t(\tau) \equiv \frac{1}{\tau} \sum_{l=0}^{\tau-1} X_{t-l}$$

(variation: replace simple average above with weighted average)

- $\overline{X}_t(1) = X_t$ is scale 1 'average', while $\overline{X}_{N-1}(N) = \overline{X}$

Wavelet Coefficients and Filters

- wavelet coefficients tell us about variations in adjacent averages
- use wavelet filter to create wavelet coefficients
- given $X_0, X_1, \ldots, X_{N-1}$, define wavelet coefficients via

$$\widetilde{W}_{j,t} \equiv \sum_{l=0}^{L_j-1} \tilde{h}_{j,l} X_{t-l \bmod N}, \quad t = 0, 1, \ldots, N-1,$$

  where $\tilde{h}_{j,l}$ is a wavelet filter with L_j coefficients, and

$$\begin{aligned} X_{t-l \bmod N} &= X_t, \quad 0 \leq t-l \leq N-1 \\ X_{-1 \bmod N} &= X_{N-1} \\ X_{-2 \bmod N} &= X_{N-2} \text{ etc (`circularity')} \end{aligned}$$

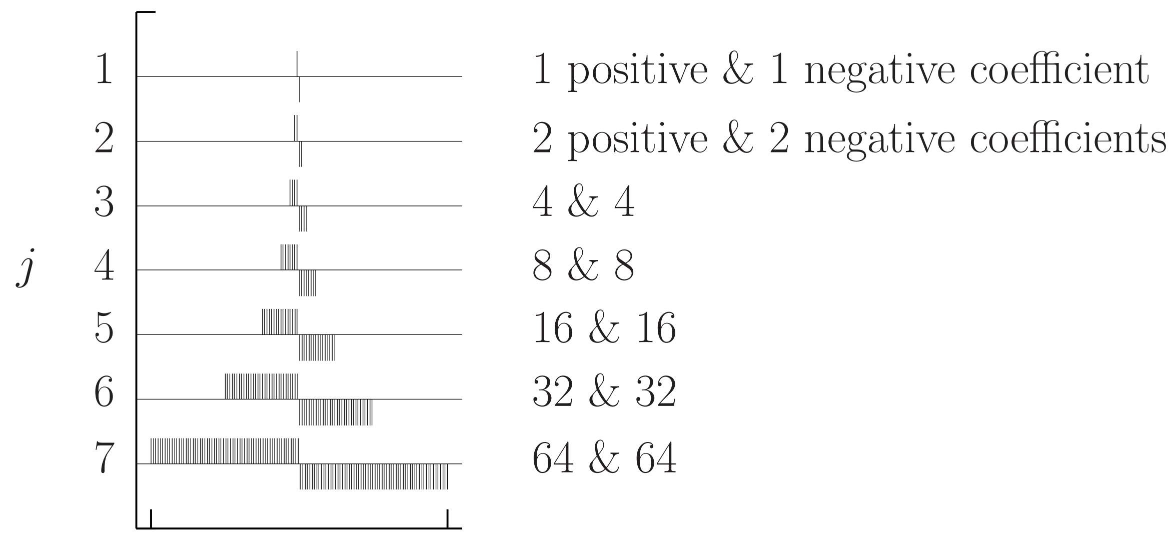
- index j specifies associated scale as $\tau_j \equiv 2^{j-1}$, $j = 1, 2, \ldots$; i.e., scales are powers of two (1, 2, 4, 8, …)

Daubechies Wavelet Filters

- analysis of variance requires filter $\tilde{h}_{1,l}$ of unit scale to satisfy certain conditions
- will use Daubechies wavelet filters with L_1 coefficients, for which
  - $\sum_{l=0}^{L_1-1} \tilde{h}_{1,l} = 0$
  - $\sum_{l=0}^{L_1-1} \tilde{h}_{1,l}^2 = 1/2$
  - $\sum_{l=0}^{L_1-1} \tilde{h}_{1,l}\tilde{h}_{1,l+2k} = 0$ for nonzero integers k
- $\tilde{h}_{j,l}$'s for $j > 1$ are 'stretched out' versions of $\tilde{h}_{1,l}$
- L_1 must be even integer $(2, 4, 6, \ldots)$
- when $L_1 = 2$, filter is known as the Haar wavelet filter

Example: Haar Wavelet Filters

- Haar wavelet filters $\tilde{h}_{j,l}$ for scales indexed by $j = 1, \ldots, 7$

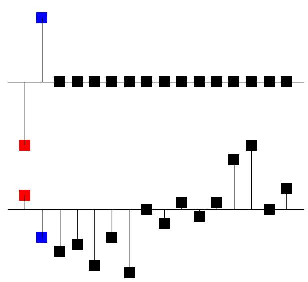


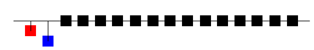
Haar Wavelet Coefficients: I

- consider how $\widetilde{W}_{1,1} = \sum_l \tilde{h}_{1,l} X_{1-l \bmod N}$ is formed ($N = 16$):

$\tilde{h}_{1,l}$

X_t

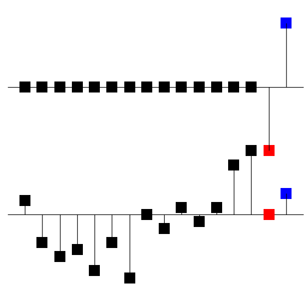


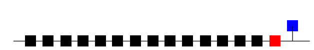
product  sum $\propto \overline{X}_1(1) - \overline{X}_0(1)$

- similar interpretation for $\widetilde{W}_{1,15} = \sum_l \tilde{h}_{1,l} X_{15-l \bmod N}$:

$\tilde{h}_{1,l}$

X_t



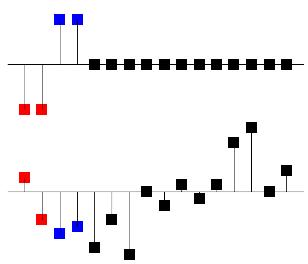
product  sum $\propto \overline{X}_{15}(1) - \overline{X}_{14}(1)$

Haar Wavelet Coefficients: II

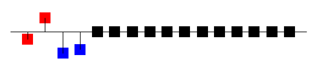
- now consider form of $\widetilde{W}_{2,3} = \sum_l \tilde{h}_{2,l} X_{3-l \bmod N}$:

$h_{2,l}$

X_t



product



sum $\propto \overline{X}_3(2) - \overline{X}_1(2)$

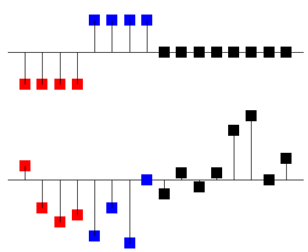
- similar interpretation for $\widetilde{W}_{2,4}, \widetilde{W}_{2,5}, \ldots, \widetilde{W}_{2,15}$
- note: $\widetilde{W}_{2,0}, \widetilde{W}_{2,1}$ and $\widetilde{W}_{2,2}$ aren't proportional to differences of adjacent averages (called 'boundary' coefficients)

Haar Wavelet Coefficients: III

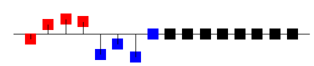
- $\widetilde{W}_{3,7} = \sum_l \tilde{h}_{3,l} X_{7-l \bmod N}$ takes the following form:

$\tilde{h}_{3,l}$

X_t



product

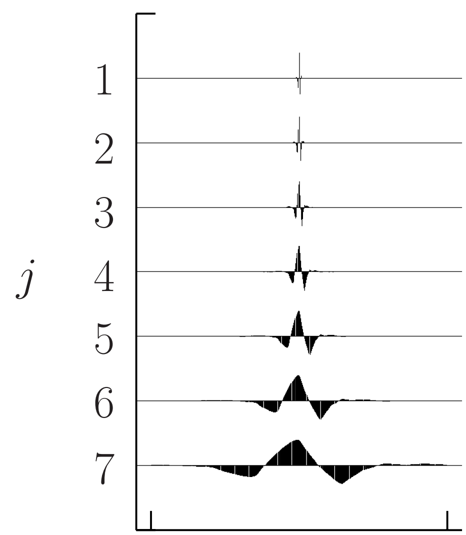


sum $\propto \overline{X}_7(4) - \overline{X}_3(4)$

- Haar wavelet coefficients $\widetilde{W}_{j,t}$ for scale $\tau_j = 2^{j-1}$ proportional to $\overline{X}_t(\tau_j) - \overline{X}_{t-\tau_j}(\tau_j)$. i.e., to change in adjacent τ_j averages
  - change measured by simple first difference
  - average is localized sample mean
  - if $\widetilde{W}^2_{j,t}$ small, not much variation over scale τ_j
  - if $\widetilde{W}^2_{j,t}$ large, lot of variation over scale τ_j

Second Example: LA(8) Wavelet Filters

- as example of another wavelet filter, consider the Daubechies 'least asymmetric' filter of width 8 (denoted as LA(8))



- LA(8) wavelet coefficients proportional to difference between central weighted average and 2 surrounding weighted averages

Empirical Wavelet Variance

- define empirical wavelet variance for scale τ_j as

$$\tilde{\nu}_X^2(\tau_j) \equiv \frac{1}{N} \sum_{t=0}^{N-1} \widetilde{W}_{j,t}^2$$

- if $N = 2^J$, obtain analysis (decomposition) of sample variance:

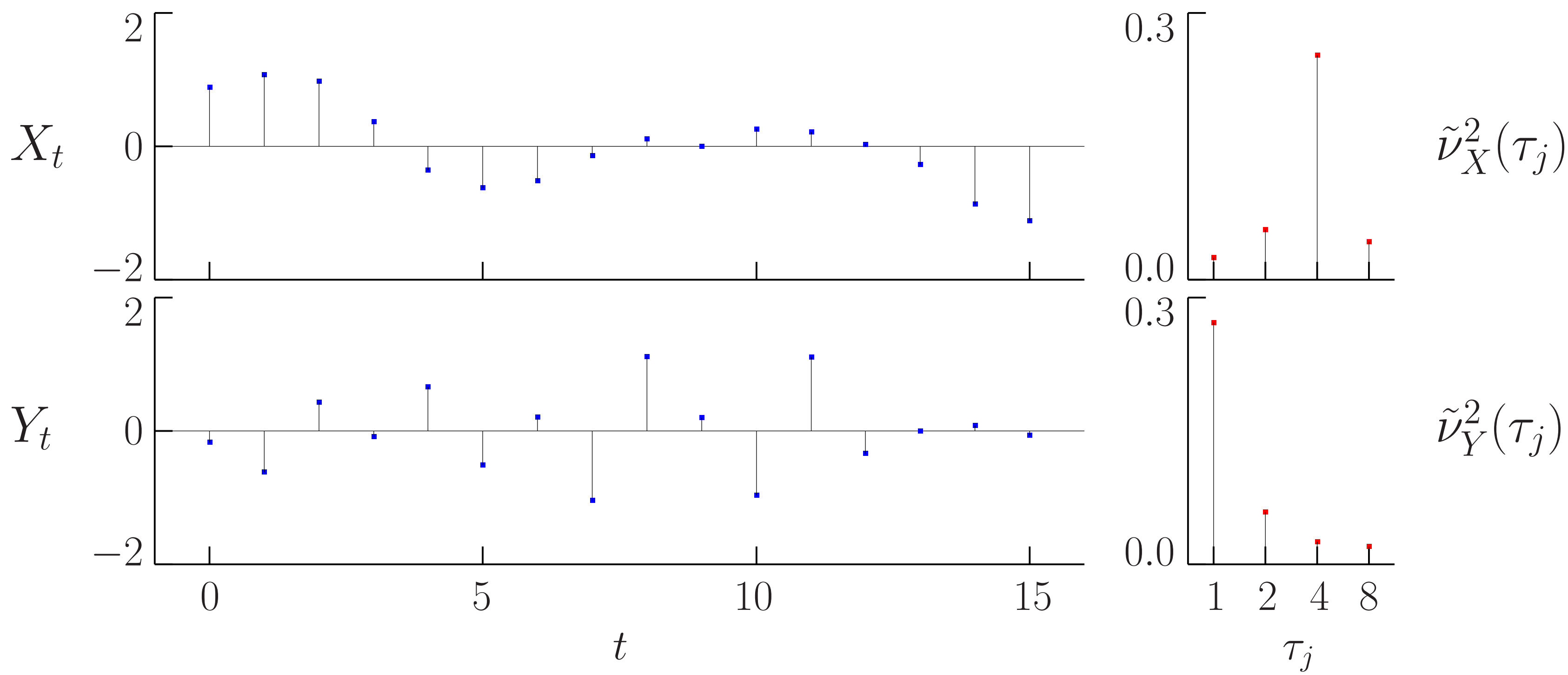
$$\hat{\sigma}_X^2 = \frac{1}{N} \sum_{t=0}^{N-1} \left(X_t - \overline{X}\right)^2 = \sum_{j=1}^{J} \tilde{\nu}_X^2(\tau_j)$$

(if N not a power of 2, can still obtain an analysis of variance to a given level J_0, but have component due to 'scaling' filter)

- interpretation: $\tilde{\nu}_X^2(\tau_j)$ is portion of $\hat{\sigma}_X^2$ due to changes in averages over scale τ_j; i.e., 'scale by scale' analysis of variance

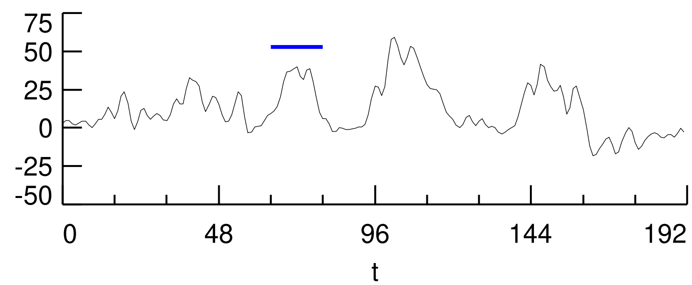
Example of Empirical Wavelet Variance

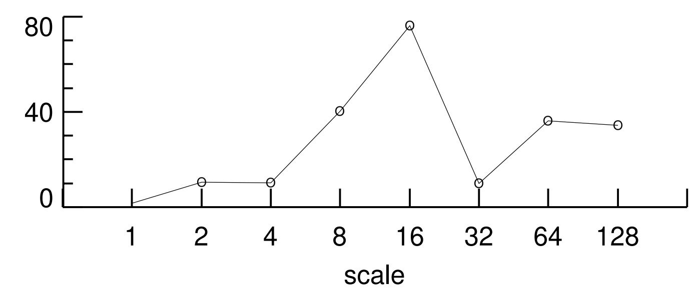
- wavelet variances for time series X_t and Y_t of length $N = 16$, each with zero sample mean and same sample variance



Second Example of Empirical Wavelet Variance

- top: subtidal sea level series X_t (blue line shows scale of 16)





- bottom: empirical wavelet variances $\tilde{\nu}_X^2(\tau_j)$
- note: each $\widetilde{W}_{j,t}$ associated with a portion of X_t, so $\widetilde{W}_{j,t}^2$ versus t offers time-based decomposition of $\tilde{\nu}_X^2(\tau_j)$

Theoretical Wavelet Variance: I

- now assume X_t is a real-valued random variable (RV)
- let X_t, $t \in \mathbb{Z}$ denote a stochastic process, i.e., collection of RVs indexed by 'time' t (here $\mathbb{Z}$ denotes the set of all integers)
- filter X_t to create new stochastic process:

$$\overline{W}_{j,t} \equiv \sum_{l=0}^{L_j-1} \tilde{h}_{j,l} X_{t-l}, \quad t \in \mathbb{Z},$$

which should be contrasted with

$$\widetilde{W}_{j,t} \equiv \sum_{l=0}^{L_j-1} \tilde{h}_{j,l} X_{t-l \bmod N}, \quad t = 0, 1, \ldots, N-1$$

Theoretical Wavelet Variance: II

- if Y is any RV, let $E\{Y\}$ denote its expectation
- let $\operatorname{var}\{Y\}$ denote its variance: $\operatorname{var}\{Y\} \equiv E\{(Y - E\{Y\})^2\}$
- definition of time dependent wavelet variance:

$$\nu^2_{X,t}(\tau_j) \equiv \operatorname{var}\{\overline{W}_{j,t}\},$$

  with conditions on X_t so that $\operatorname{var}\{\overline{W}_{j,t}\}$ exists and is finite
- $\nu^2_{X,t}(\tau_j)$ depends on τ_j and t
- will focus on time independent wavelet variance

$$\nu^2_X(\tau_j) \equiv \operatorname{var}\{\overline{W}_{j,t}\}$$

  (can adapt theory to handle time varying situation)
- $\nu^2_X(\tau_j)$ well-defined for stationary & related processes, so let's review concept of stationarity

Definition of a Stationary Process

- if U and V are two RVs, denote their covariance by
$$\operatorname{cov}\{U, V\} = E\{(U - E\{U\})(V - E\{V\})\}$$
- stochastic process X_t called stationary if
  - $E\{X_t\} = \mu_X$ for all t, i.e., constant independent of t
  - $\operatorname{cov}\{X_t, X_{t+\tau}\} = s_{X,\tau}$, i.e., depends on lag τ, but not t
- $s_{X,\tau}$, $\tau \in \mathbb{Z}$, is autocovariance sequence (ACVS)
- $s_{X,0} = \operatorname{cov}\{X_t, X_t\} = \operatorname{var}\{X_t\}$; i.e., variance same for all t

Example of a Stationary Process: White Noise

- simplest example of a stationary process is 'white noise'
- process X_t said to be white noise if
  - it has a constant mean $E\{X_t\} = \mu_X$
  - it has a constant variance $\operatorname{var}\{X_t\} = \sigma_X^2$
  - $\operatorname{cov}\{X_t, X_{t+\tau}\} = 0$ for all t and nonzero τ; i.e., distinct RVs in the process are uncorrelated
- ACVS for white noise takes a very simple form:

$$s_{X,\tau} = \operatorname{cov}\{X_t, X_{t+\tau}\} = \begin{cases} \sigma_X^2, & \tau = 0; \\ 0, & \text{otherwise.} \end{cases}$$

Wavelet Variance for Stationary Processes

- for stationary processes, wavelet variance decomposes var $\{X_t\}$:

$$\sum_{j=1}^{\infty} \nu_X^2(\tau_j) = \text{var}\,\{X_t\}$$

  (above result similar to one for sample variance)

- $\nu_X^2(\tau_j)$ is thus contribution to var $\{X_t\}$ due to scale τ_j
- example: for a white noise process, have

$$\nu_X^2(\tau_j) = \frac{\text{var}\,\{X_t\}}{2^j} = \frac{\text{var}\,\{X_t\}}{2\tau_j},$$

  so largest contribution to var $\{X_t\}$ is at smallest scale τ_1

- note: $\nu_X(\tau_j)$ has same units as X_t, which is important for interpretability

Generalization to Certain Nonstationary Processes

- if L_1 is properly chosen, $\nu_X^2(\tau_j)$ well-defined for processes with stationary backward differences
- first order backward difference of X_t is process defined by

$$X_t^{(1)} = X_t - X_{t-1}$$

- second order backward difference of X_t is process defined by

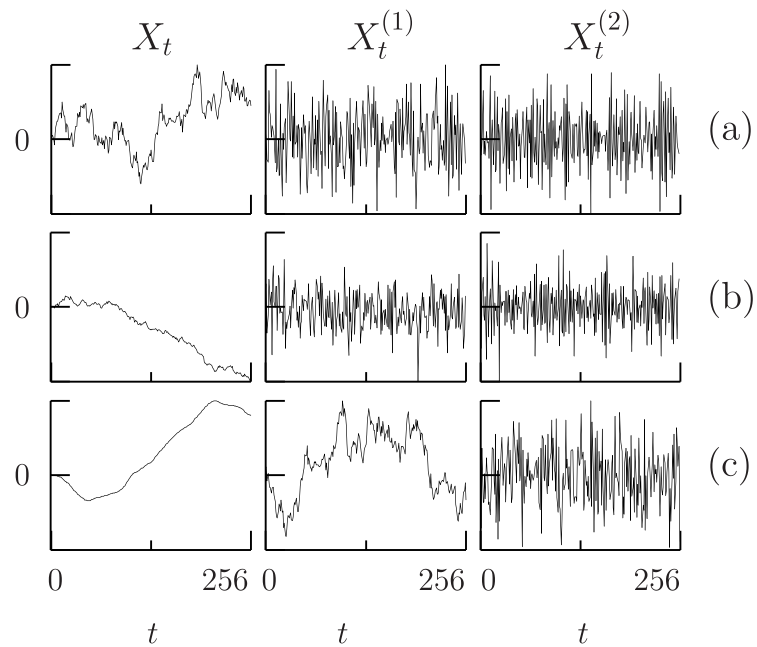
$$X_t^{(2)} = X_t^{(1)} - X_{t-1}^{(1)} = X_t - 2X_{t-1} + X_{t-2}$$

- X_t has dth order stationary backward differences if

$$Y_t \equiv \sum_{k=0}^{d} \binom{d}{k} (-1)^k X_{t-k}$$

forms a stationary process (d is a nonnegative integer)

Examples of Processes with Stationary Increments



- 1st column shows, from top to bottom, realizations from

(a) random walk: $X_t = \sum_{u=1}^{t} \epsilon_t$, & ϵ_t is zero mean white noise

(b) like (a), but now ϵ_t has mean of -0.2

(c) random run: $X_t = \sum_{u=1}^{t} Y_t$, where Y_t is a random walk

- 2nd & 3rd columns show 1st & 2nd differences $X_t^{(1)}$ and $X_t^{(2)}$

Wavelet Variance for Processes with Stationary Backward Differences

- suppose X_t nonstationary with dth order stationary differences
- if $L_1 \geq 2d$, then $\nu_X^2(\tau_j)$ is well-defined & finite for all τ_j, but now we have

$$\sum_{j=1}^{\infty} \nu_X^2(\tau_j) = \infty$$

- example: for a random walk process $X_t = \sum_{u=1}^{t} \epsilon_t$, have

$$\nu_X^2(\tau_j) = \frac{\operatorname{var}\{\epsilon_t\}}{6}\left(\tau_j + \frac{1}{2\tau_j}\right)$$

with Haar wavelet, so $\nu_X^2(\tau_j)$ increases as j increases

Fractionally Differenced (FD) Processes: I

- as an example, consider wavelet variance for FD processes (Granger & Joyeux, 1980; Hosking, 1981)
- FD processes determined by 2 parameters $-\infty < \delta < \infty$ & $\sigma_\epsilon^2 > 0$ (relatively unimportant)
- let FD(δ) refer to FD process with parameter δ
- if $\delta < 1/2$, FD process X_t is stationary, and, in particular,
  - reduces to white noise if $\delta = 0$
  - has 'long memory' if $\delta > 0$
  - is 'antipersistent' if $\delta < 0$ (i.e., $\operatorname{cov}\{X_t, X_{t+1}\} < 0$)

Fractionally Differenced (FD) Processes: II

- if $\delta \geq 1/2$, FD process X_t is nonstationary with dth order stationary backward differences Y_t
  - here $d = \lfloor \delta + 1/2 \rfloor$, where $\lfloor x \rfloor$ is integer part of x
  - Y_t is stationary FD($\delta - d$) process
- if $\delta = 1$, FD process is the same as a random walk process
- at large scales, have

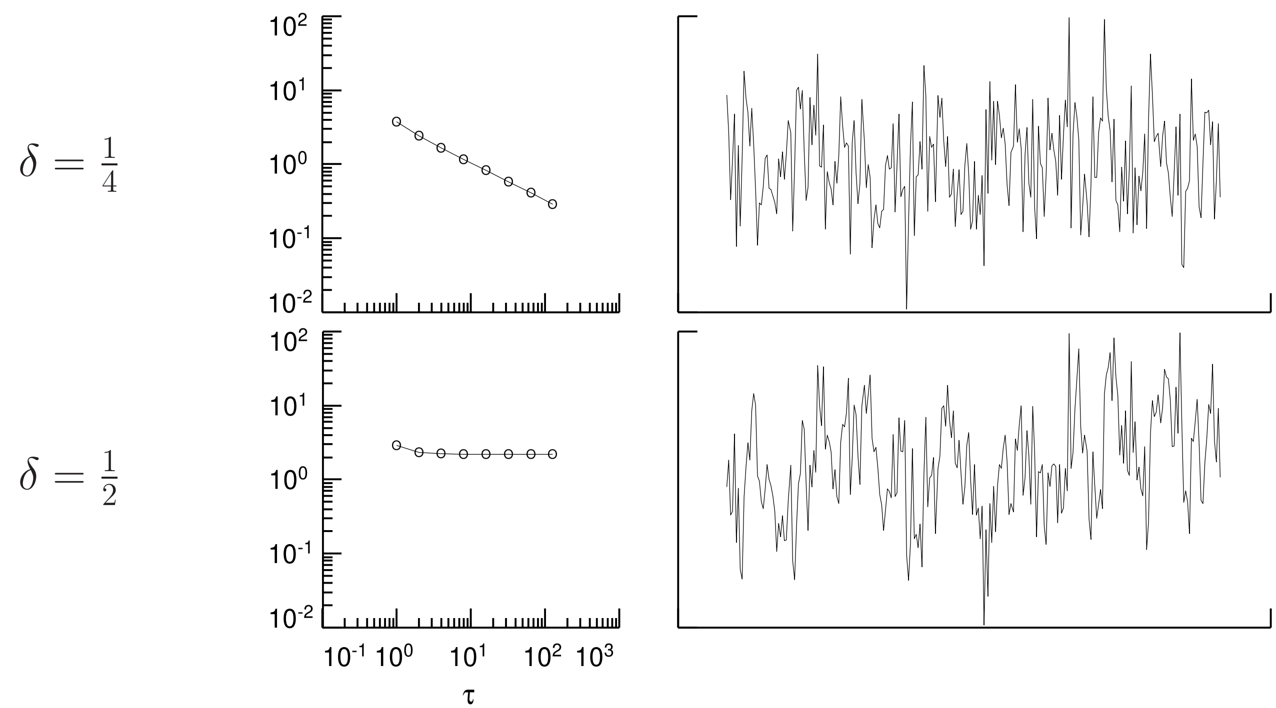
$$\nu_X^2(\tau_j) \approx C\tau_j^{2\delta-1}$$

- thus

$$\log\left(\nu_X^2(\tau_j)\right) \approx \log\left(C\right) + (2\delta - 1)\log\left(\tau_j\right),$$

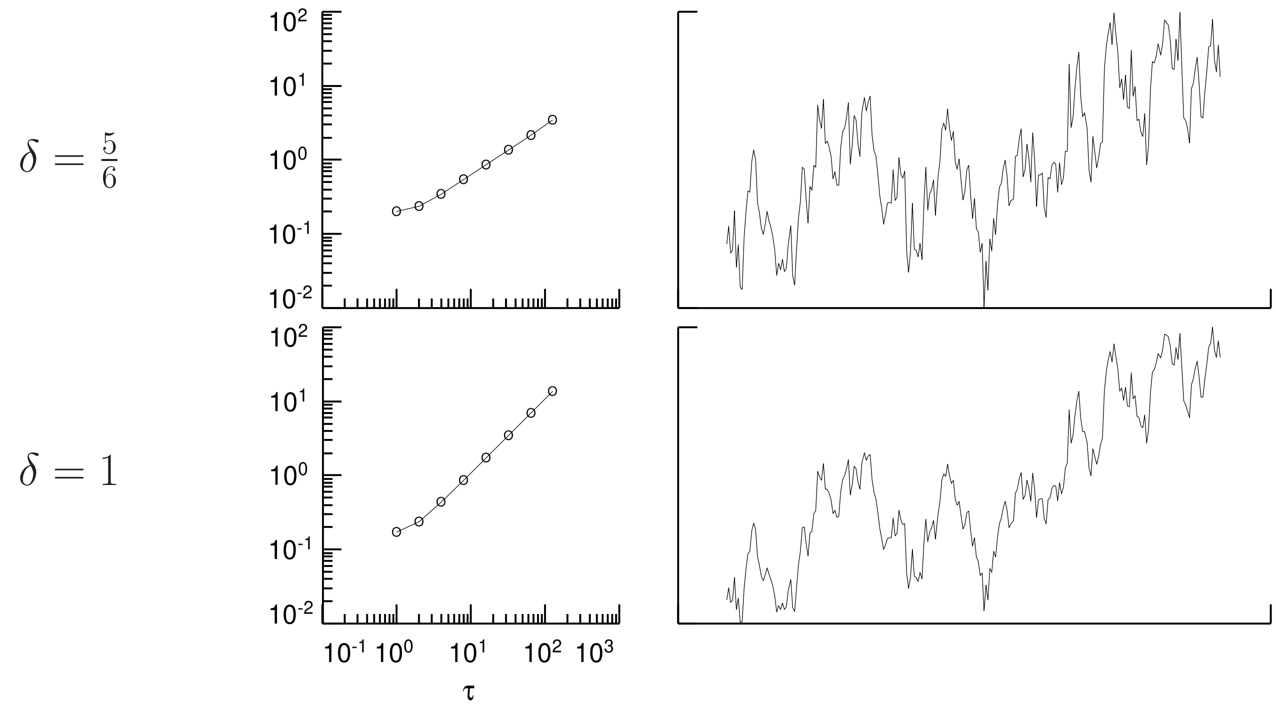
so a log/log plot of $\nu_X^2(\tau_j)$ vs. τ_j looks approximately linear with slope $2\delta - 1$ for τ_j large enough

LA(8) Wavelet Variance for 2 FD Processes



- left-hand column: $\nu_X^2(\tau_j)$ versus τ_j based upon LA(8) wavelet
- right-hand: realization of length $N = 256$ from each FD process (created via circulant embedding – details in Craigmile, 2003)

LA(8) Wavelet Variance for 2 More FD Processes



- $\delta = \frac{5}{6}$ is Kolmogorov turbulence; $\delta = 1$ is random walk
- note: positive slope indicates nonstationarity, while negative slope indicates stationarity

Unbiased Estimator of Wavelet Variance: I

- given a realization of $X_0, X_1, \ldots, X_{N-1}$ from a process with dth order stationary differences, want to estimate $\nu_X^2(\tau_j)$
- for wavelet filter such that $L_1 \geq 2d$ and $E\{\overline{W}_{j,t}\} = 0$, have

$$\nu_X^2(\tau_j) = \operatorname{var}\{\overline{W}_{j,t}\} = E\{\overline{W}_{j,t}^2\}$$

- can base estimator on squares of

$$\widetilde{W}_{j,t} \equiv \sum_{l=0}^{L_j-1} \tilde{h}_{j,l} X_{t-l \bmod N}, \qquad t = 0, 1, \ldots, N-1$$

- recall that

$$\overline{W}_{j,t} \equiv \sum_{l=0}^{L_j-1} \tilde{h}_{j,l} X_{t-l}, \qquad t \in \mathbb{Z}$$

Unbiased Estimator of Wavelet Variance: II

- comparing

$$\widetilde{W}_{j,t} = \sum_{l=0}^{L_j-1} \tilde{h}_{j,l} X_{t-l \bmod N} \text{ with } \overline{W}_{j,t} \equiv \sum_{l=0}^{L_j-1} \tilde{h}_{j,l} X_{t-l}$$

says that $\widetilde{W}_{j,t} = \overline{W}_{j,t}$ if 'mod N' not needed; this happens when $L_j - 1 \le t < N$

- if $N - L_j \ge 0$, unbiased estimator of $\nu_X^2(\tau_j)$ is

$$\hat{\nu}_X^2(\tau_j) \equiv \frac{1}{N - L_j + 1} \sum_{t=L_j-1}^{N-1} \widetilde{W}_{j,t}^2 = \frac{1}{M_j} \sum_{t=L_j-1}^{N-1} \overline{W}_{j,t}^2,$$

where $M_j \equiv N - L_j + 1$

Statistical Properties of $\hat{\nu}_X^2(\tau_j)$ (Gaussian)

- suppose $\{\overline{W}_{j,t}\}$ Gaussian with mean zero & ACVS $s_{j,\tau}$ (note: filtering tends to yield normality)
- suppose square summability condition holds:

$$A_j \equiv \sum_{\tau=-\infty}^{\infty} s_{j,\tau}^2 < \infty.$$

- can show $\hat{\nu}_X^2(\tau_j)$ asymptotically normal with mean $\nu_X^2(\tau_j)$ & large sample variance $2A_j/M_j$
- A_j finite if ACVS damps quickly to 0
- if A_j infinite, can usually correct by increasing L_1
- conclusion: square integrability easy to satisfy
- Monte Carlo studies: large sample theory good if $M_j \geq 128$

Estimation of A_j

- in practical applications, need to estimate

$$A_j = \sum_{\tau=-\infty}^{\infty} s_{j,\tau}^2$$

- for large M_j, an approximately unbiased estimator is

$$\hat{A}_j \equiv \frac{\hat{s}_{j,0}^2}{2} + \sum_{\tau=1}^{M_j-1} \hat{s}_{j,\tau}^2,$$

where

$$\hat{s}_{j,\tau} \equiv \frac{1}{M_j} \sum_{t=L_j-1}^{N-1-|\tau|} \widetilde{W}_{j,t} \widetilde{W}_{j,t+|\tau|}$$

- Monte Carlo results: $\hat{A}_j$ reasonably good for $M_j \geq 128$

Confidence Intervals (CIs) for $\nu_X^2(\tau_j)$

- for finite M_j, Gaussian-based CIs problematic: lower limit of CI can very well be negative
- can avoid by basing CIs on the assumption that

$$\hat{\nu}_X^2(\tau_j) = \frac{1}{M_j} \sum_{t=L_j-1}^{N-1} \widetilde{W}_{j,t}^2$$

has the same distrubution as $a\chi_\eta^2$, i.e., a constant times a chi-square RV with η equivalent degrees of freedom (EDOF)

- moment matching yields

$$\eta = \frac{2\left(E\{\hat{\nu}_X^2(\tau_j)\}\right)^2}{\mathrm{var}\,\{\hat{\nu}_X^2(\tau_j)\}}$$

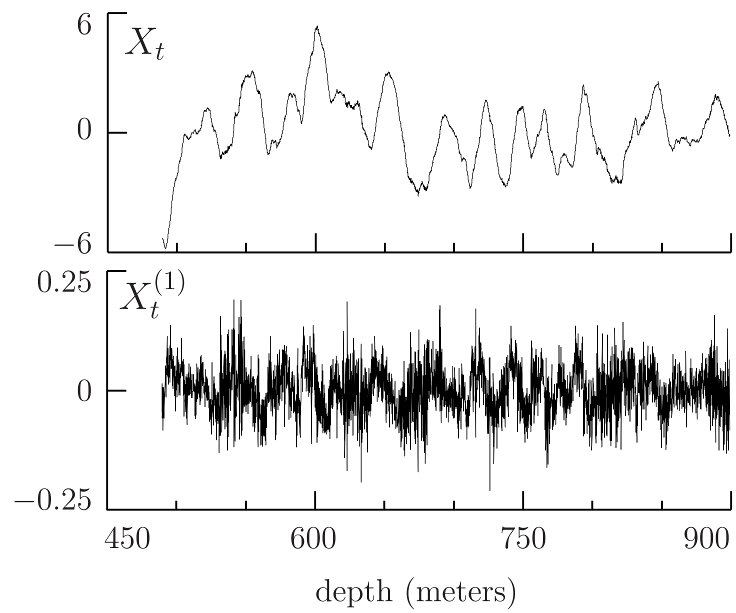
Three Ways to Set η

1. use large sample theory with appropriate estimates:
$$\hat{\eta}_1 = \frac{M_j \hat{\nu}_X^4(\tau_j)}{\hat{A}_j}$$

2. assume nominal shape for spectral density function of X_t:
$$S_X(f) = hC(f), \text{ where } C(f) \text{ is known, but } h \text{ is not;}$$
though questionable, get acceptable CIs using
$$\eta_2 = \frac{2\left(\sum_{k=1}^{\lfloor (M_j-1)/2 \rfloor} C_j(f_k)\right)^2}{\sum_{k=1}^{\lfloor (M_j-1)/2 \rfloor} C_j^2(f_k)}$$

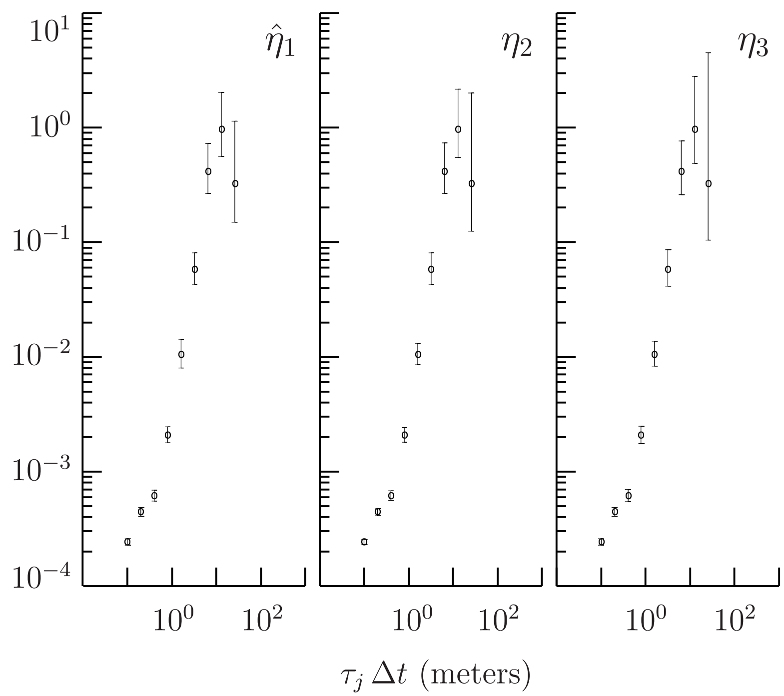
3. make an assumption about the effect of wavelet filter on X_t to obtain simple (but effective!) approximation
$$\eta_3 = \max\{M_j/2^j, 1\}$$

Example: Vertical Shear in the Ocean: I



- top plot: vectical shear measurements X_t
- bottom: backward differences $X_t^{(1)}$

Example: Vertical Shear in the Ocean: II



- wavelet variance estimates based upon Daubechies wavelet with $L_1 = 6$, along with 95% confidence intervals for true wavelet variance with EDOFs determined by $\hat{\eta}_1$ estimated from data, η_2 using a nominal model for $S_X(\cdot)$ and $\eta_3 = \max\{M_j/2^j, 1\}$

Statistical Properties of $\hat{\nu}_X^2(\tau_j)$ (Non-Gaussian)

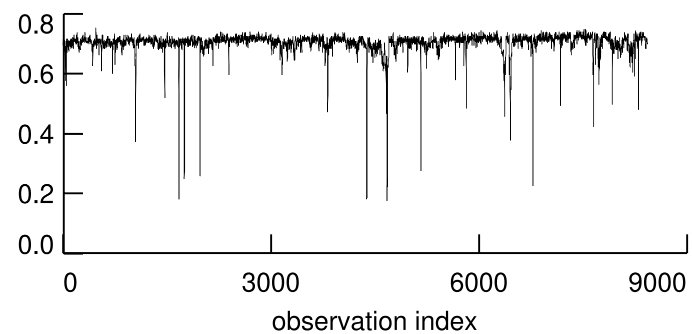
- assume $\{\overline{W}_{j,t}\}$ strictly stationary process satisfying
  - $E\{\overline{W}_{j,t}\} = 0$ and $E\{|\overline{W}_{j,t}|^{4+2\delta}\} < \infty$ for some $\delta > 0$
  - mixing condition $\alpha_{\overline{W}_{j,t}} = O(1/n^{\chi})$, where

$$\alpha_{\overline{W}_{j,t}} \equiv \sup_{A \in \mathcal{M}_{-\infty}^{0},\, B \in \mathcal{M}_t^{\infty}} |\mathbf{P}(A \cap B) - \mathbf{P}(A)\mathbf{P}(B)|,$$

  $\mathcal{M}_m^n$ is σ-algebra for $\overline{W}_{j,m}, \ldots, \overline{W}_{j,n}$ and $\chi > (2+\delta)/\delta$

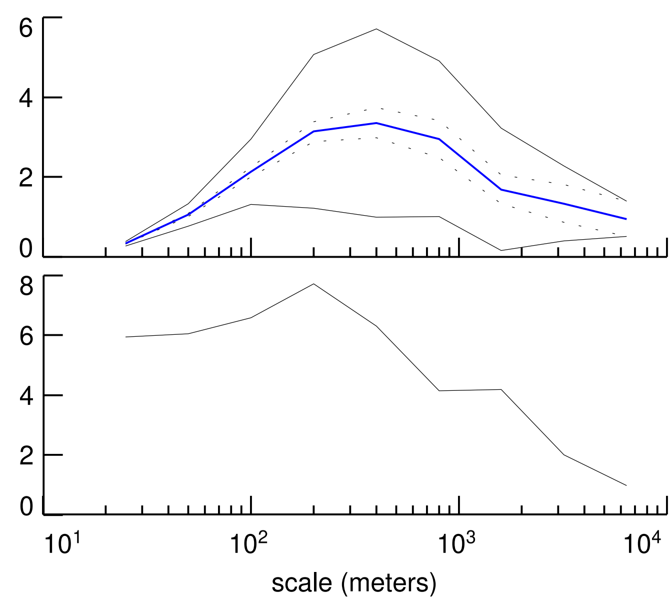
- let $Z_{j,t} \equiv \overline{W}_{j,t}^2$ have spectral density function (SDF) $S_{Z_j}(\cdot)$ such that $0 < S_{Z_j}(0) < \infty$
- $\hat{\nu}_X^2(\tau_j)$ asymptotically normal with mean $\nu_X^2(\tau_j)$ & large sample variance $S_{Z_j}(0)/M_j$ (can be estimated using standard SDF estimators such as multitaper or autoregressive estimators)

Example: Surface Albedo of Spring Pack Ice: I



- data from Beaufort Sea ($N = 8428$, sampled every 25 meters)

Example: Surface Albedo of Spring Pack Ice: II



- upper plot: estimated LA(8) wavelet variance (blue curve), along with upper and lower 90% confidence intervals based upon Gaussian (thin dotted curves) and non-Gaussian theory (thin solid)
- lower plot: ratio of estimated non-Gaussian versus Gaussian large sample standard deviations

Wavelet Variance Analysis of Time Series with Time-Varying Statistical Properties

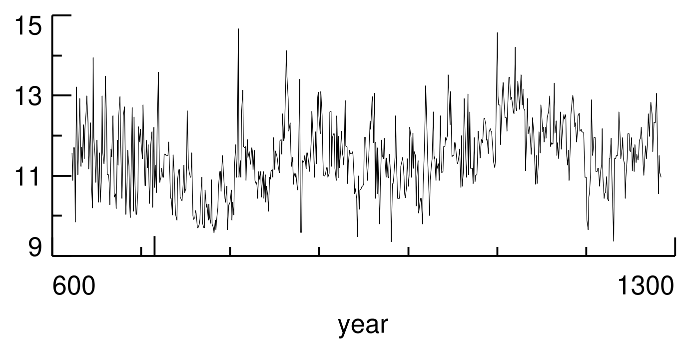
- each wavelet coefficient $\widetilde{W}_{j,t}$ formed using portion of X_t
- suppose X_t associated with actual time $t_0 + t\,\Delta t$
  - t_0 is actual time of first observation X_0
  - Δt is spacing between adjacent observations
- suppose $\tilde{h}_{j,l}$ is least asymmetric Daubechies wavelet
- can associate $\widetilde{W}_{j,t}$ with an interval of width $2\tau_j\,\Delta t$ centered at

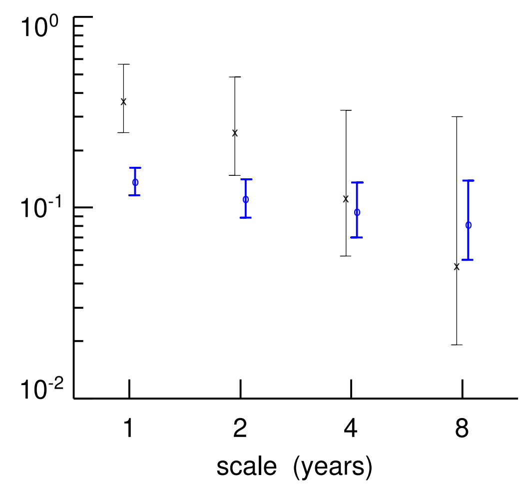
$$t_0 + (2^j(t+1) - 1 - |\nu_j^{(H)}| \bmod N)\,\Delta t,$$

where, e.g., $|\nu_j^{(H)}| = [7(2^j - 1) + 1]/2$ for LA(8) wavelet

- can thus form 'localized' wavelet variance analysis (implicitly assumes stationarity or stationary increments locally)

Example: Annual Minima of Nile River





- left plot: annual minima of Nile River
- bottom: Haar $\hat{\nu}_X^2(\tau_j)$ before ($\mathtt{x}$'s) and after ($\mathtt{o}$'s) year 715.5, with 95% confidence intervals based upon $\chi^2_{\eta_3}$ approximation

Some Extensions and Ongoing Work

- wavelet cross-covariance and cross-correlation (see references)
- biased estimators of wavelet variance
- unbiased estimator of wavelet variance for 'gappy' time series
- extension of notion and estimators to random fields

Summary

- wavelet variance gives scale-based analysis of variance
- statistical theory worked out for
  - Gaussian processes with stationary backward differences
  - non-Gaussian processes satisfying a mixing condition
- applications include analysis of
  - genome sequences
  - frequency fluctuations in atomic clocks
  - changes in variance of soil properties
  - accumulation of snow fields in polar regions
  - turbulence in atmosphere and ocean
  - regular and semiregular variables stars
- thanks for invitation to speak!!!

References

- fractionally differenced processes
  - C. W. J. Granger and R. Joyeux (1980), 'An Introduction to Long-Memory Time Series Models and Fractional Differencing,' *Journal of Time Series Analysis*, **1**, pp. 15–29
  - J. R. M. Hosking (1981), 'Fractional Differencing,' *Biometrika*, **68**, pp. 165–76.
- simulation via circulant embedding
  - P. F. Craigmile (2003), 'Simulating a Class of Stationary Gaussian Processes using the Davies–Harte Algorithm, with Application to Long Memory Processes,' *Journal of Time Series Analysis*, **24**, pp. 505–11.
- surface albedo of spring pack ice from Beaufort Sea
  - A. Serroukh, A. T. Walden and D. B. Percival (2000), 'Statistical Properties and Uses of the Wavelet Variance Estimator for the Scale Analysis of Time Series,' *Journal of the American Statistical Association*, **95**, pp. 184–96
- wavelet cross-covariance and cross-correlation
  - B. J Whitcher, P. Guttorp and D. B. Percival (2000), 'Wavelet Analysis of Covariance with Application to Atmospheric Time Series,' *Journal of Geophysical Research*, **105**, D11, pp. 14,941–62.
  - A. Serroukh and A. T. Walden (2000), 'Wavelet Scale Analysis of Bivariate Time Series I: Motivation and Estimation,' *Journal of Nonparametric Statistics*, **13**, pp. 1–36.
  - A. Serroukh and A. T. Walden (2000), 'Wavelet Scale Analysis of Bivariate Time Series II: Statistical Properties for Linear Processes,' *Journal of Nonparametric Statistics*, **13**, pp. 37–56.
- general reference
  - D. B. Percival and A. T. Walden (2000), *Wavelet Methods for Time Series Analysis*, Cambridge, UK, Cambridge University Press