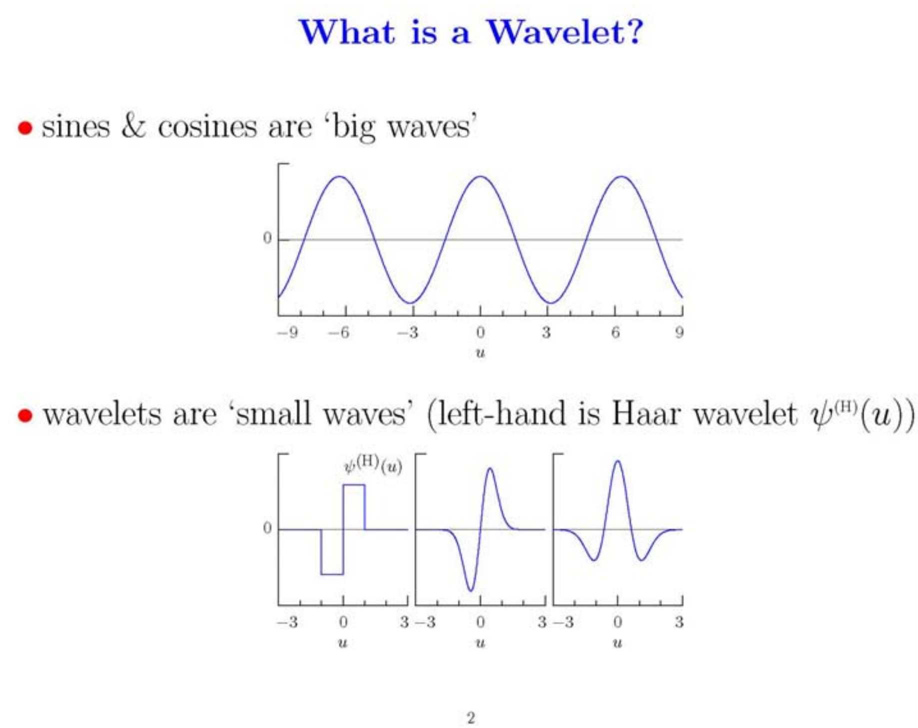
WAVELET ANALYSIS OF CLOCK NOISE

Don Percival
Applied Physics Laboratory
University of Washington
Seattle, WA 98195-5640, USA

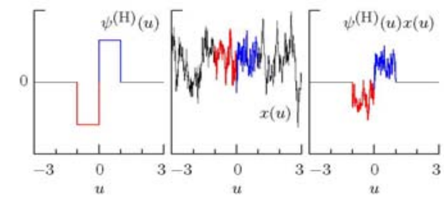
Abstract

A wavelet is a relatively new mathematical tool that can be used to analyze data arising in PTTI applications. This paper presents a basic introduction to wavelets and wavelet analysis. We also look briefly at some specific uses for wavelets with PTTI data, including (1) tracking data whose statistical properties are evolving over time, (2) decomposing the sample variance for a set of data into components that are attributable to variations over different scales, and (3) decorrelating highly correlated data.



What is Wavelet Analysis?: I

- multiply wavelet & time series $x(u)$ together & integrate:

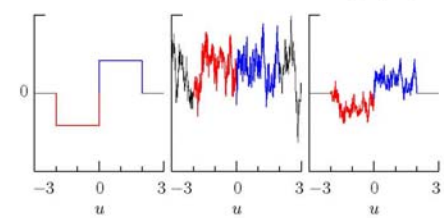


- $\int_{-\infty}^{\infty} \psi^{(\mathrm{H})}(u)x(u)\,du = W(1,0)$ is proportional to difference between averages of $x(u)$ over intervals $[-1,0]$ and $[0,1]$
- defines wavelet coefficient $W(1,0)$ for
  - scale 1 (width of each interval)
  - time 0 (center of combined intervals)

3

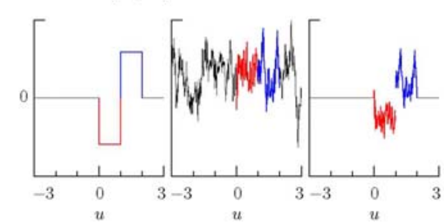
What is Wavelet Analysis?: II

- stretch or shrink wavelet to define $W(\tau,0)$ for other scales τ:



yields $W(2,0)$

- relocate to define $W(\tau,t)$ for other times t:



yields $W(1,1)$

4

What is Wavelet Analysis?: III

- $W(\tau, t)$ over all scales $\tau > 0$ and all times t called continuous wavelet transform (CWT) for $x(u)$
- CWT analyzes $x(u)$ into components that are
  - associated with a scale and a time
  - physically related to a difference of averages
- similar interpretation for other wavelets $\psi(u)$
- $W(\tau, t)$ equivalent to $x(u)$ since, given CWT, can recover $x(u)$:

$$x(u) = \frac{1}{C_\psi} \int_0^\infty \frac{1}{\tau^2} \left[\int_{-\infty}^\infty W(\tau, t) \frac{1}{\sqrt{\tau}} \psi\left(\frac{u-t}{\tau}\right) \, dt \right] d\tau,$$

where C_ψ is a constant depending on specific wavelet $\psi(u)$

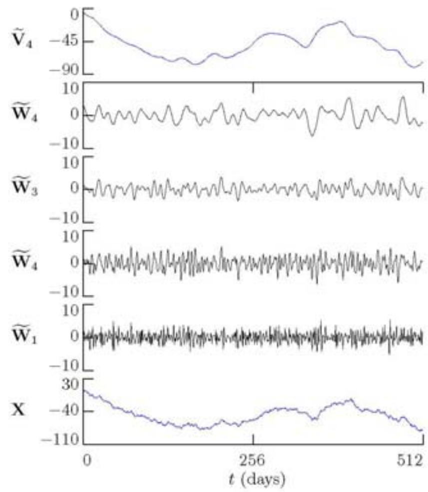
5

Maximal Overlap Discrete Wavelet Transform

- let $\mathbf{X} = [X_0, X_1, \ldots, X_{N-1}]^T$ be observed time series
- can formulate MODWT of $\mathbf{X}$ as vectors $\widetilde{\mathbf{W}}_1, \ldots, \widetilde{\mathbf{W}}_{J_0}$ & $\widetilde{\mathbf{V}}_{J_0}$, each of dimension N (number of levels J_0 chosen by user)
- $\widetilde{\mathbf{W}}_j$ contains wavelet coefficients, $j = 1, \ldots, J_0$
  - associated with differences in averages over scale $\tau_j = 2^{j-1}$
  - closely related to $W(\tau_j, t)$ over restricted set of times
- $\widetilde{\mathbf{V}}_{J_0}$ contains scaling coefficients
  - associated with averages over scale $2\tau_{J_0} = 2^{J_0}$
  - summarizes $W(\tau, t)$ over scales $\tau > \tau_{J_0}$
- $\mathbf{X}$ & MODWT equivalent: given MODWT, can recover $\mathbf{X}$

6

Example: MODWT Coefficients for Clock 55



- can use to track variations across time at a given scale

7

Wavelet-Based Analysis of Variance: I

- consider 'energy' in time series: $\|\mathbf{X}\|^2 = \mathbf{X}^T\mathbf{X} = \sum_{t=0}^{N-1} X_t^2$
- energy preserved in MODWT coefficients:

$$\|\mathbf{X}\|^2 = \sum_{j=1}^{J_0} \|\widetilde{\mathbf{W}}_j\|^2 + \|\widetilde{\mathbf{V}}_{J_0}\|^2$$

- leads to analysis of sample variance:

$$\hat{\sigma}_X^2 = \frac{1}{N}\sum_{t=0}^{N-1}(X_t - \hat{\mu}_X)^2 = \frac{1}{N}\Big(\sum_{j=1}^{J_0} \|\widetilde{\mathbf{W}}_j\|^2 + \|\widetilde{\mathbf{V}}_{J_0}\|^2\Big) - \hat{\mu}_X^2,$$

where $\hat{\mu}_X = \frac{1}{N}\sum_{t=0}^{N-1} X_t$ is sample mean

8

Wavelet-Based Analysis of Variance: II

- if $\mathbf{X}$ realization of process with stationary increments, $\|\widetilde{\mathbf{W}}_j\|^2/N$ is estimator of wavelet variance $\nu_X^2(\tau_j)$
- wavelet variance analyzes process variance σ_X^2 across scales τ_j:

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

(note: σ_X^2 can be infinite for certain processes)

- special case: Haar wavelet variance with fractional frequency deviates $\overline{Y}_t$ essentially same as Allan variance $\sigma_{\overline{Y}}^2(2, \tau_j)$ since

$$\nu_{\overline{Y}}^2(\tau_j) = \tfrac{1}{2}\sigma_{\overline{Y}}^2(2, \tau_j)$$

- Q: 'old wine in a new bottle,' or something new?

9

What Wavelets Bring to the Table

- $2\|\widetilde{\mathbf{W}}_j\|^2/N$ gives a previously unknown estimator for $\sigma_{\overline{Y}}^2(2, \tau_j)$
- with addition of 'reflection' boundary conditions, estimator is an improvement over existing estimators (smaller mean square error; Greenhall, Howe & Percival, 1999)
- non-Haar wavelets provide interesting generalizations
  - still provide exact decompositions of sample variance
  - can handle wider range of power laws
  - can handle polynomial trends of certain orders
  - competitive with modified Allan variance
- unified theory provides methods for getting confidence intervals that do not require *a prior* assumption of noise type

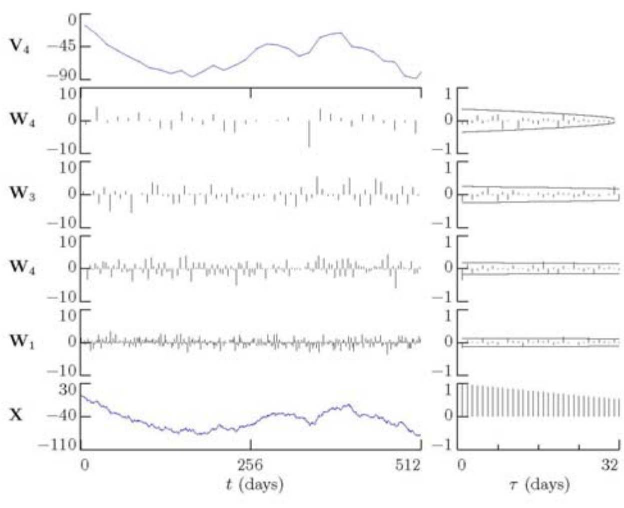
10

Discrete Wavelet Transform (DWT)

- obtain by subsampling and rescaling MODWT
- yields vectors $\mathbf{W}_1, \ldots, \mathbf{W}_{J_0}$ & $\mathbf{V}_{J_0}$
  - $\mathbf{W}_j$ has $N/2^j$ wavelet coefficients
  - $\mathbf{V}_{J_0}$ has $N/2^{J_0}$ scaling coefficients
- total # of DWT coefficients is N, i.e., dimension of $\mathbf{X}$
- $\mathbf{X}$ & DWT equivalent: given DWT, can recover $\mathbf{X}$
- DWT acts as a decorrelating transform

11

Example: DWT Coefficients for Clock 55 X_t



- have approximate within-scale & between-scale decorrelation (non-Haar wavelets offer better between-scale approximation)

12

Uses for DWT Decorrelating Property

- estimation of parameters for statistical models
  - consider modeling $\mathbf{X}$ as process with spectrum
$$S_X(f) = C|f|^\alpha$$
  - power-law model depends on parameters C and α
  - consider estimating C and α via maximum likelihood (ML)
  - exact ML estimators difficult to obtain
  - DWT yields simple, but effective, approximate ML estimator
- testing for homogeneity of $\mathbf{X}$ at scale τ_j across time
- assessing variability in certain statistics via bootstrapping
- fast simulation of time series

13

Other Potential Uses for Wavelets in PTTI

- multiresolution analysis (based on wavelet synthesis of $\mathbf{X}$)
- detection of singularities (maximum modulus of CWT)
- data compression
- signal extraction (wavelet shrinkage)

14

Summary

- wavelets give insight into frequency instability characterization
  - emphasize role of exact analysis of process/sample variance
  - provide estimators with reduced mean square error
- wavelets lead to easily computed approximate maximum likelihood estimators for parameters of power-law processes
- many other potential uses
  - article #20,654 is waiting to be written!
- thanks to conference organizers for invitation to speak!

15

REFERENCES

[1] C. A. Greenhall, D. A. Howe, and D. B. Percival, 1999, *"Total Variance, an Estimator of Long-term Frequency Stability,"* **IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, UFFC-46,** 1183-1191.

[2] D. B. Percival, 2003, *"Stochastic Models and Statistical Analysis for Clock Noise,"* **Metrologia, 40,** S289-S304.

[3] D. B. Percival and A. T. Walden, 2000, **Wavelet Methods for Time Series Analysis** (Cambridge University Press).

QUESTIONS AND ANSWERS

MARC WEISS (National Institute of Standards and Technology): How does the wavelet analysis compare in terms of the long term? There is a problem with getting a long-term Allan variance in terms of confidence and bias. Does the wavelet variance, if you have a variance via wavelet analysis, improve the problem?

DON PERCIVAL: Yes, it does. In the work I did with Dave Howe and Chuck Greenhall on TOTVAR, what you are doing is getting a bias estimate of the wavelet of the Allan variance using this scheme right here. But it has associated with it a smaller variance. In other words, the total mean squared area can show that it decreased. So it gives the superior estimate out the very, very longest time scales. That is probably the big advantage.

WEISS: So it is essentially TOTVAR

PERCIVAL: Yes, with a Haar thing, it ends being TOTVAR. So you can regard TOTVAR as a wavelet-based estimator of the Allan variance. So it is just a new estimator of the Allan variance.

FRANÇOIS VERNOTTE (Observatoire de Besançon): In the literature, there are many different wavelet chips. How is it possible to choose the right one for a given problem?

PERCIVAL: It is not that difficult. The Daubechie family of wavelets is a series of filters of different lengths. So the Haar is length two. There is a length four, length six, and length eight. What I found to be a very effective thing is just to do an analysis in which you look at what you get for the different wavelet lengths. What you will often find is that sometimes the Haar is not adequate, because it has certain leakage properties. That is why the Allan variance has problems at times. When you go to these higher order things, things will stabilize and you will find that using the length four wavelet and the length six wavelet gives you about the same thing. So that means you can back off and just use the length four one. So there are some very simple little techniques you can use in order to pick an appropriate wavelet from amongst the Daubechie family of wavelets.

DENNIS McCARTHY (U.S. Naval Observatory): Just a quick question. One of the concerns for people in this community is you have fixed finite-length data, and you would like to determine what is the stochastic data and what is sort of the underlying trend. Have you thought about a way the wavelets might be used to separate the deterministic part from the noise part?

PERCIVAL: Yes. If the trend can be modeled as a low-order polynomial, what happens is that polynomial will not show up at all in the wavelet coefficients. You can in a certain sense handle polynomial trends very easily. The Haar wavelet has a first difference in it. The next wavelet up has a second difference and the next wavelet up has a third difference. So a third difference would kill off a quadratic term, which would get rid of them totally. By going to the higher-order wavelet, you can actually automatically handle at least polynomial trends and get out coefficients which are impervious to that trend background.

McCARTHY: A quick comment. The wavelet business here has been why it has received widespread use in the geodetic and geophysical community. Something you did not mention was looking at the spectral content of time series as a function of time. That has proved to be extremely useful to establish the spectra of various physical processes as a function of time, which can be done with this.

PERCIVAL: Right, exactly. So that little example I gave where the clock seemed to be increasing in variability toward the end is exactly what you are talking about. Because the scale eight things would relate to a certain band of frequencies; you can use that to track the frequency variations across time. So, yes, that is a very good point.