CNR-IMATI Milano

December 17, 2004

What does a Single Run Tell About the Ensemble?

Claudia Angelini, Daniela Cava, Gabriel Katul, and Brani Vidakovic

P L A N

1. Data in the Wavelet Domain
2. Bootstrap and Its Relatives
   - Efron's Example: Hidalgo Stamps
3. Percival's Wavestrap
4. Bootstrapping Turbulence
   - Duke Forest Measurements
   - Simulated Data
5. Conclusions

■ DATA IN THE WAVELET DOMAIN

- DWT "Rotates" Data.
- Dimension of Data Preserved.
- Entropy Reduced [Amenability to Shrinkage].
- Self-Similarity Visible.
- Scales/Frequencies Separated.
- Data Whitened.
- Also, Wavelet Transformations are Fast, Lossless, Versatile.

$$\boxed{\textbf{DATA}} \xrightarrow{\mathbb{W}} \boxed{\textbf{Wavelet Coefficients}}$$

$$\|$$

$$\boxed{\textbf{Processed DATA}} \xleftarrow{\mathbb{W}^{-1}} \boxed{\textbf{Process the Coefficients}}$$

Process $\equiv$

- Shrink
- Transform
- Simulate New
- Resample
- etc.

Multiscale Domains: 1

- Various Wavelet Transformations: Continuous, Orthogonal, Non-Decimated (Stationary), Wavelet Packets, Complex, etc.

- Focus on Discrete, and because of energy preservation $\rightarrow$ Orthogonal.

- $\mathbf{y}$ (time series) $\xrightarrow{\mathbb{W}}$ $\mathbf{d} = [c_{J_0}, d_{J_0}, d_{J_0+1} \dots d_{J-2}, d_{J-1}]$. $J_0, J_0+1, \dots, J-1$ scales ranging from the coarsest to the finest.

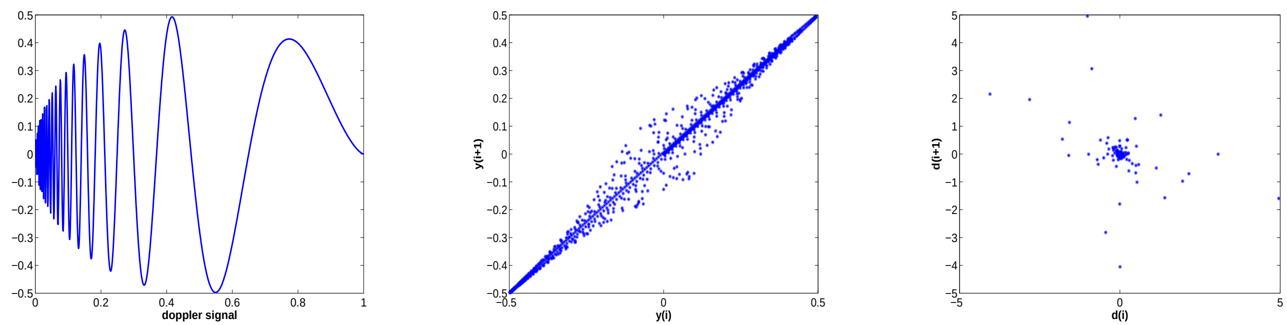
- $\ell(\mathbf{y}) = 2^J$, $\ell(d_{J-1}) = 2^{J-1}$ [it decimates],$\dots$, $\ell(d_{J_0}) = 2^{J_0}$, $\ell(c_{J_0}) = 2^{J_0}$; Thus, $\ell(\mathbf{d}) = 2^J$.

- ON Wavelet transformations are rotations in the 2^J dimensional space where data are points. Amazingly, the rotations bring many coordinate axes closer to the data point.

Multiscale Domains: 2

- Why Wavelet Domains? Wavelets "Disbalance" the data, Filter the data, Assess self-similarity, they are Fast, Versatile, and **Whitening.**

- Focus on Whitening Property. Well explored when inputs possess some stochastic structure. Inhibiting the forcasting/prediction by wavelets.



- References: Flandrin, 1992; Tewfik & Kim 1992; Walter, 1994; Johnstone & Silverman, 1997; Craigmile et al, 2000.

Bootstrap 1

- Bootstrap (Efron & Tibshirani, 93), Powerful and Controversial Technique. Bootstrap Originally designed for i.i.d. (independent, identically distributed) data.

- Analogous Technique: **Surrogate Data** (Theiler et al., 92). Mainly used for testing nonlinearity in observed time series.

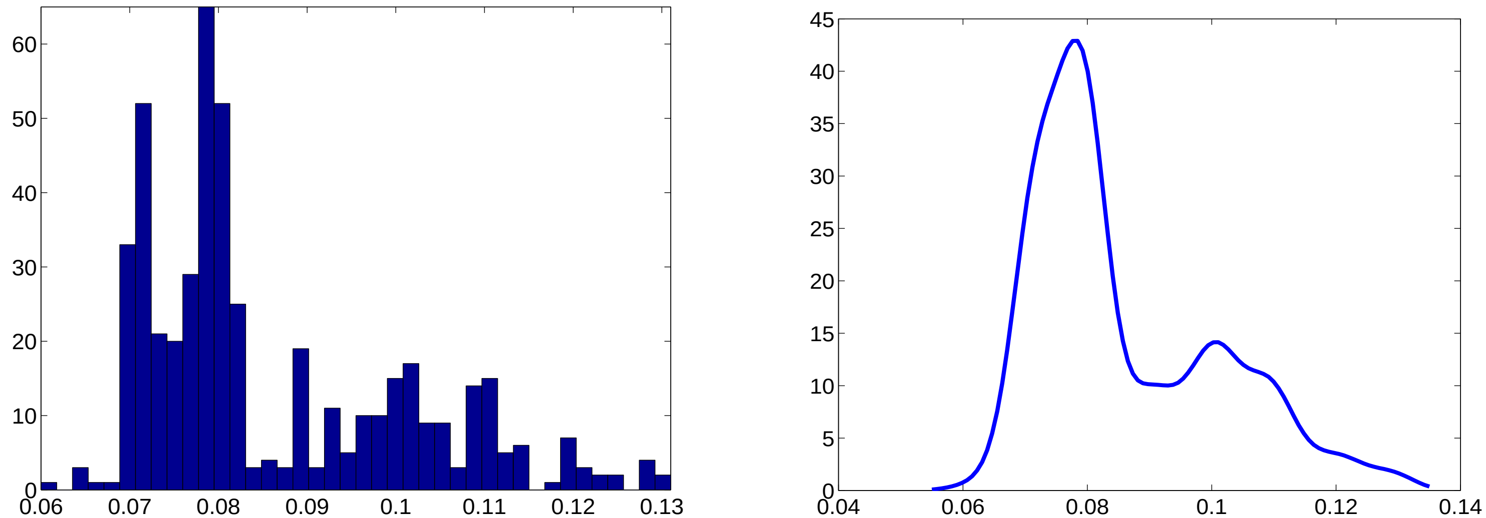
- When data are dependent, either do (i) block bootstraping (block jackknifing), (ii) whiten and bootstrap, or (ii) BOTH.

- Moving Block Bootstrap (MBB), (Künsch, 89); Stationary B (Politis & Romano, 94) Lin. Comb. Varying Size BB (Politis and Romano, 95); Matched BB; Circular B; Non-overlapping BB (NBB); Missing Data Bootstrap (Alfonso et al, 2003).

■ Bootstrap 2: Hidalgo Stamps

Mexico, Issue 1872; $n = 485$ ■ Izenman & Sommer, JASA 1988; Different papers used;

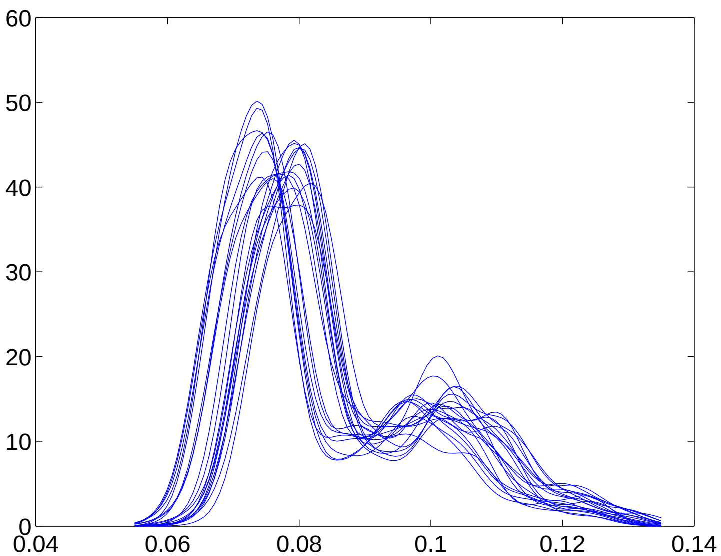
Bootstrap 3: Hidalgo Stamps

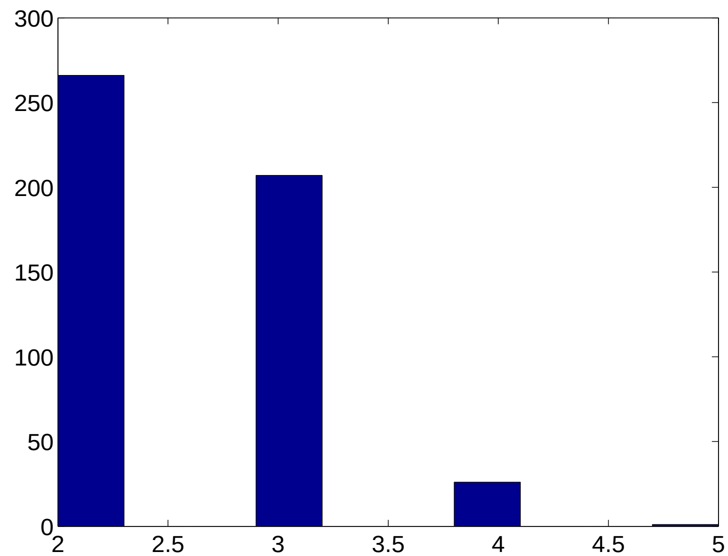


- Test for number of modes.
- Gaussian kernel density estimator with $h = 0.00333$ (border-line between 2- and 3-modal estimator).
- $y \longrightarrow y_1^*, y_2^*, \ldots, y_B^*$.
- B density estimators ($h = 0.00333$) based on y_i^*

Bootstrap 4: Hidalgo Stamps

- Count # of modes: n_1 (one), n_2 (two), etc.
- $H_0 : \#of modes = 2$ vs $H_1 : \#of modes > 2$





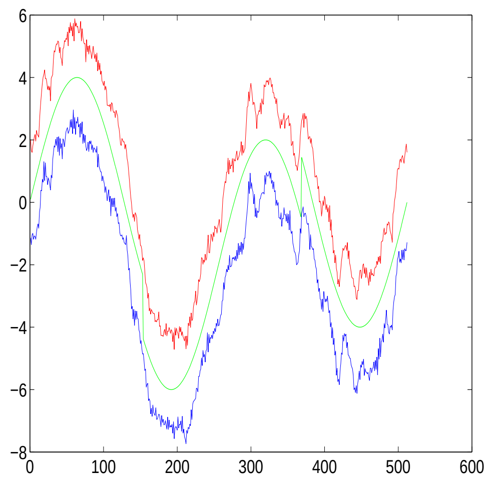
- ASV (empirical counterpart to p-value) is about 0.5.

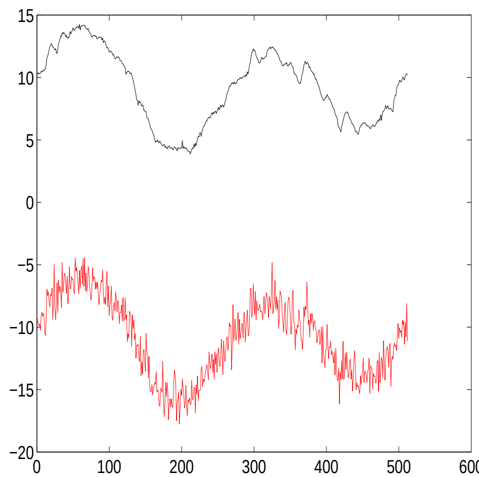
■ Bootstrap and Multiscale Domains: ♡ 1

■ Under-researched. Wavelets are Time/Scale representations and Bootstrap is scrambling Time information.

■ Bootstrap Shrinkage: Possible? Yes.   Any research? No.

■ Resample detail levels and average. Resembles *Stochastic Resonance.*





■ Bootstrap and Multiscale Domains:♡ 2

■ More sophisticated method: **Skeleton BootWave**

■ **Procedure:**

(i) Make the time/scale skeleton [say surviving coefficients of universal (or any under-fitting) thresholding method];

(ii) fill-in the “meat” by resampling the discarded (thresholded) coefficients.

(iii) Average for the estimator;

(iv) Use envelope for the error bands.

■ Competitive Shrinkage Technique when: (i) Skeleton is sparse; (ii) **B**- the number of bootstrap resamples is large.

Bootstrap and Multiscale Domains: 3

- Percival's Wavestrap (Percival et al., 2000)

- Adaptive Wavestrap via Wavelet Packets. Best basis is the most decorrelating basis. Criteria: whiteness of scales in WP. If a scale not white enough, go to its "children" and check their whiteness.

- (Percival et al, 2000), Estimating the variance of $\rho(1)$-one-lag autocorrelation $[n = 1024,\ B = 10,000,\ \text{Symm } 8]$

	DWT	DWPT	actual
AR(1) $\phi = 0.9$	0.016 ± 0.001	0.015 ± 0.001	0.014
MA(1) $\theta = 0.98$	0.026 ± 0.001	0.024 ± 0.001	0.022
FD $d = 0.45$	0.047 ± 0.004	0.047 ± 0.003	0.053

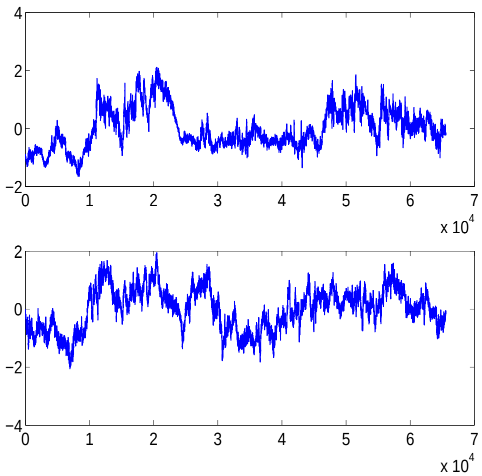
Bootstrapping Turbulence

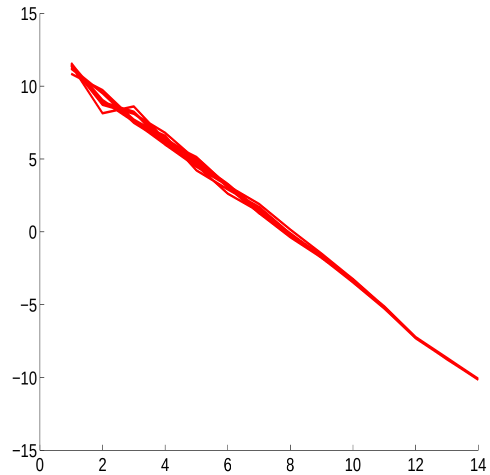
- Turbulence is not a "long memory" but it has long memory (in an antipersistent sense).
- Pointwise bootstrap whitens the runs.
- For estimating fluxes keep the bootstrapping policy the same.
- U, V, W, and T collected over a grass-covered forest clearing at Duke Forest near Durham, North Carolina. The measurements were collected on July, 12-16, 1995 at 5.2 m above the grass surface using a Gill triaxial sonic anemometer.

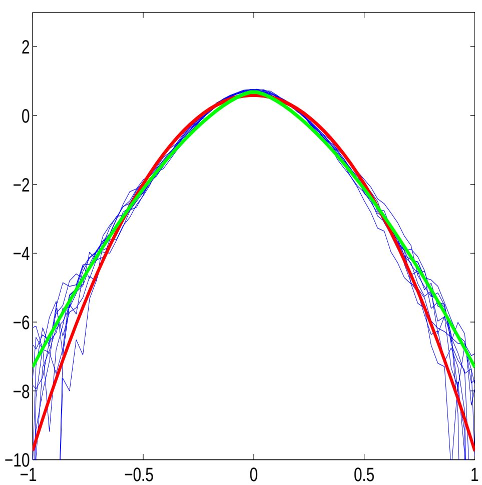
Stationary Bootstrap in the Wavelet Domain

- Stationary Bootstrap: Random level-block size. Block Length Geometric $\mathcal{G}(p_j)$. Asymptotically optimal choice $p_j = O(2^{-j/3})$. For turbulence experiments recommend $p = 1/30$. and 6-8 vanishing moments for the wavelet basis.

- Single Run U, Duke Forest measurements by Gaby Katul $G950713.04$. Block Length Geometric $\mathcal{G}(0.15)$. $\hat{H} = 0.3334$, $\sigma_{\hat{H}} = 0.0216$.

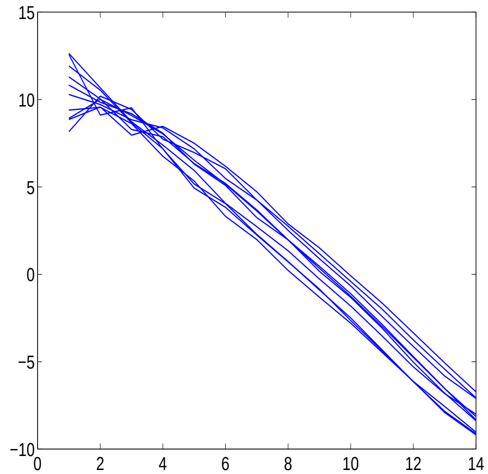


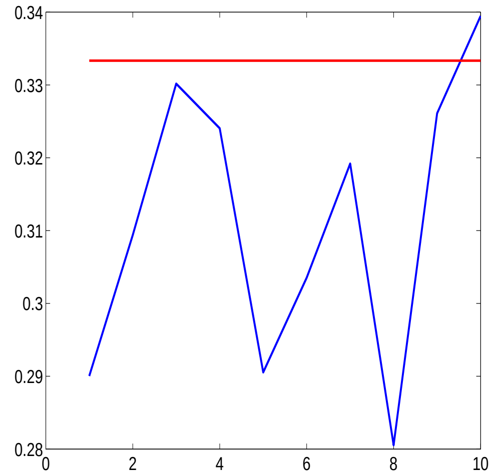




■ “Nature” Bootstrap

■ Ten Runs U, Duke Forest measurements by Gaby Katul (Unstable Conditions).

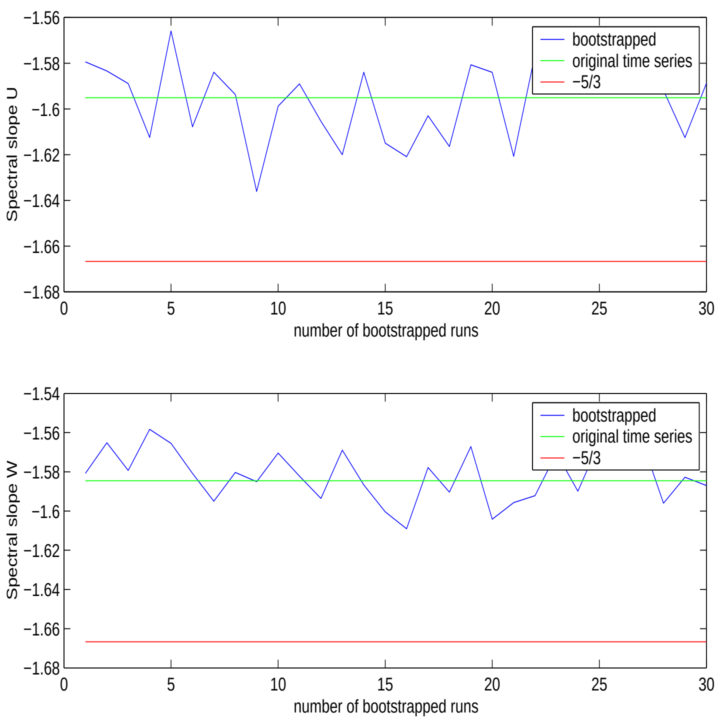


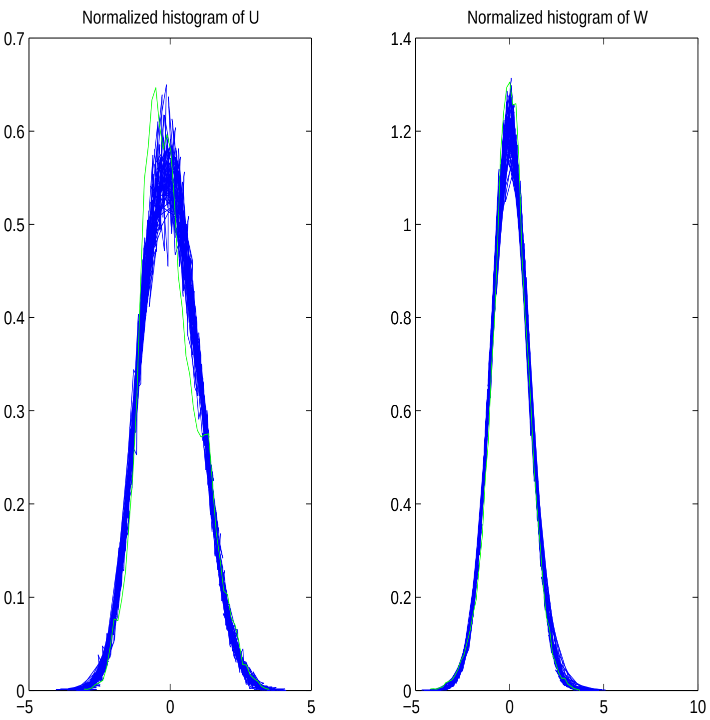


■ $\hat{H} = 0.3113$

■ $\sigma_{\hat{H}} = 0.0197.$

Figures from Italy





Simulated Data $fBm(1/3)$

- Mimics Second Order Properties of Turbulence.
- Trade-off Between Bias and Variance
- Wavestrap or Stationary Bootstrap with small expected block-size: Estimators biased (whitened) but variability preserved.
- Stationary bootstrap with large expected block-size: Estimators close to theoretical, but variance small.

$fbm(1/3)$

lag	0	1	2	3	4	5
$\hat{\rho}$	1.0000	0.9947	0.9915	0.9888	0.9863	0.9841
$\sigma_{\hat{\rho}}$	0	0.0036	0.0057	0.0075	0.0091	0.0106
$\hat{\rho}_W$	1.0000	0.9793	0.9664	0.9553	0.9456	0.9356
$\sigma_{\hat{\rho}_W}$	0	0.0029	0.0045	0.0058	0.0074	0.0088
$\hat{\rho}_S$	1.0000	0.9926	0.9887	0.9842	0.9811	0.9779
$\sigma_{\hat{\rho}_S}$	0	0.0008	0.0012	0.0016	0.0020	0.0024

length $= 2^{13}$,Symmlet 8, # of bootstrap replicates=400, Block Sizes Geometric(0.1).

An Additional Method: Waveknife

- Keep data length as power of 2 or a multiple of a power of 2.
- Erase block(s) and estimate the missing data from the remaining data, levelwise.
- Combine with “parametric”, model based bootstrap. When the input is fBm or turbulence measurements, levels well modeled by $ARMA(p, q)$ processes, $p, q < 5$.
- First numerical experiments indicate low variability.

CONCLUSIONS

- Stationary Bootstrapping useful technique in getting the estimators and their variability from a single vector of measurements.
- Methodology sensitive to selection of wavelet basis, more research needed.
- Bias-Variance trade-off.
- Potential in Wavelet-Shrinkage-type of problems.
- Bootstrapping multifractal spectra in Turbulence.