What does a Single Run Tell About the Ensemble?

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

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.
\[ \text{DATA} \xrightarrow{\mathbb{W}} \text{Wavelet Coefficients} \]

\[ \text{Processed DATA} \xleftarrow{\mathbb{W}^{-1}} \text{Process the Coefficients} \]

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 → Orthogonal.

\( \mathbf{y} \) (time series) \( \xrightarrow{\text{W}} \mathbf{d} = [c_{J_0}, d_{J_0}, d_{J_0+1}, \ldots, d_{J-2}, d_{J-1}] \).

\( J_0, J_0 + 1, \ldots, J - 1 \) scales ranging from the coarsest to the finest.

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

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 forecasting/prediction by wavelets.

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 bootstrapping (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;
Test for number of modes.

Gaussian kernel density estimator with $h = 0.00333$ (border-line between 2- and 3-modal estimator).

$y \rightarrow y_1^*, y_2^*, \ldots, y_B^*$.

$B$ 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.
- Resample detail levels and average. Resembles Stochastic Resonance.
 Bootstrap and Multiscale Domains:  

More sophisticated method: **Skeleton BootWave**

- **Procedure:**
  1. Make the time/scale skeleton [say surviving coefficients of universal (or any under-fitting) thresholding method];
  2. fill-in the “meat” by resampling the discarded (thresholded) coefficients.
  3. Average for the estimator;
  4. 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 [$n = 1024, B = 10,000, \text{Symm 8}$]

<table>
<thead>
<tr>
<th></th>
<th>DWT</th>
<th>DWPT</th>
<th>actual</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR(1) $\phi = 0.9$</td>
<td>$0.016 \pm 0.001$</td>
<td>$0.015 \pm 0.001$</td>
<td>$0.014$</td>
</tr>
<tr>
<td>MA(1) $\theta = 0.98$</td>
<td>$0.026 \pm 0.001$</td>
<td>$0.024 \pm 0.001$</td>
<td>$0.022$</td>
</tr>
<tr>
<td>FD $d = 0.45$</td>
<td>$0.047 \pm 0.004$</td>
<td>$0.047 \pm 0.003$</td>
<td>$0.053$</td>
</tr>
</tbody>
</table>
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.
- $U, V, W, \text{ and } T$ 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 $\mathcal{G}(0.15)$. $\hat{H} = 0.3334$, $\sigma_{\hat{H}} = 0.0216$. 

![Graphs showing stationary bootstrap results](image-url)
“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**

- Spectral slope $U$
- Spectral slope $W$

Normalized histograms of $U$ and $W$.
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.
### $fbm(1/3)$

<table>
<thead>
<tr>
<th>lag</th>
<th>$\hat{\rho}$</th>
<th>$\sigma_{\hat{\rho}}$</th>
<th>$\hat{\rho}_W$</th>
<th>$\sigma_{\hat{\rho}_W}$</th>
<th>$\hat{\rho}_S$</th>
<th>$\sigma_{\hat{\rho}_S}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.0000</td>
<td>0.0000</td>
<td>1.0000</td>
<td>0.0000</td>
<td>1.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>1</td>
<td>0.9947</td>
<td>0.0036</td>
<td>0.9793</td>
<td>0.0029</td>
<td>0.9926</td>
<td>0.0008</td>
</tr>
<tr>
<td>2</td>
<td>0.9915</td>
<td>0.0057</td>
<td>0.9664</td>
<td>0.0045</td>
<td>0.9887</td>
<td>0.0012</td>
</tr>
<tr>
<td>3</td>
<td>0.9888</td>
<td>0.0075</td>
<td>0.9553</td>
<td>0.0058</td>
<td>0.9842</td>
<td>0.0016</td>
</tr>
<tr>
<td>4</td>
<td>0.9863</td>
<td>0.0091</td>
<td>0.9456</td>
<td>0.0074</td>
<td>0.9811</td>
<td>0.0020</td>
</tr>
<tr>
<td>5</td>
<td>0.9841</td>
<td>0.0106</td>
<td>0.9356</td>
<td>0.0088</td>
<td>0.9779</td>
<td>0.0024</td>
</tr>
</tbody>
</table>

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.