Using Extended Records of Streamflow to Simulate Low-Frequency Variability in Future Flows

Upmanu Lall (1,2), Francisco Assis Souza Filho (1), Hyun-Han Kwon (2), and Connie Woodhouse (3)

[1]International Research Institute for Climate and Society, Columbia University, Palisades, NY, 10964
[2]Department of Earth & Environmental Engineering, Columbia University, New York, NY, 10027
[3]Department of Geography and Regional Development, University of Arizona, Tucson, Arizona 85721

Motivating questions

- How can we best assess the risk of a drought with a duration greater than T (e.g., 5) years, a cumulative deficit in supply greater than D acre-feet, in a period of N (e.g. 100 years)?
- How can we assess the resilience and reliability of a reservoir system and the associated water allocation policies in the face of such droughts?

Problem:

Climate change simulations may not capture the variability and low-frequency characteristics that are an inherent part of natural variability and that may underlie future trends in climate.

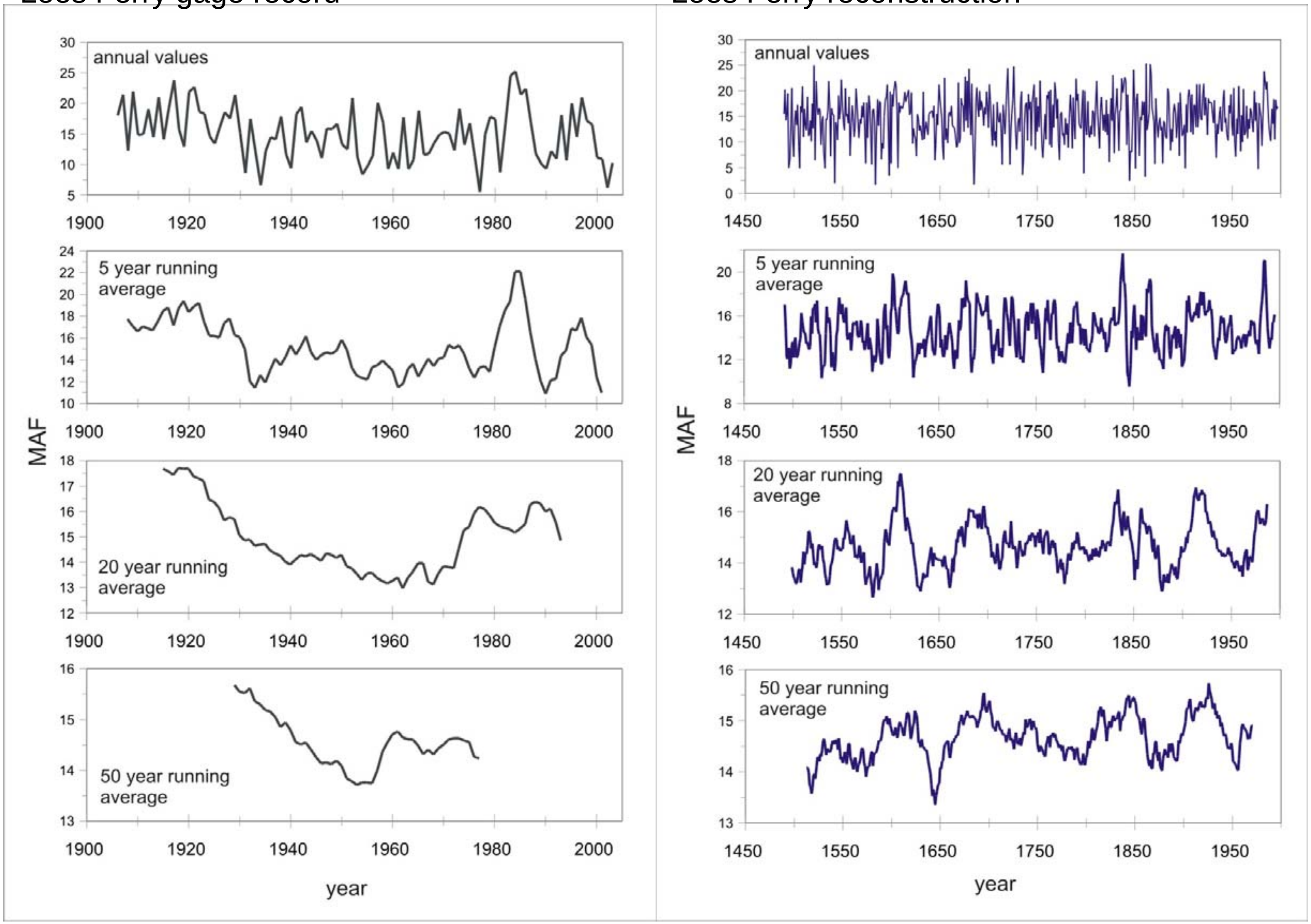
From extended records of hydroclimate from tree-ring based reconstructions, we know that these centuries-long records contain decadal and multidecadal scale variability driven by slowly varying ocean/atmospheric dynamics.

How can this information be incorporated into simulations of future hydroclimatic variability and hence used for assessing the risk of severe sustained drought?

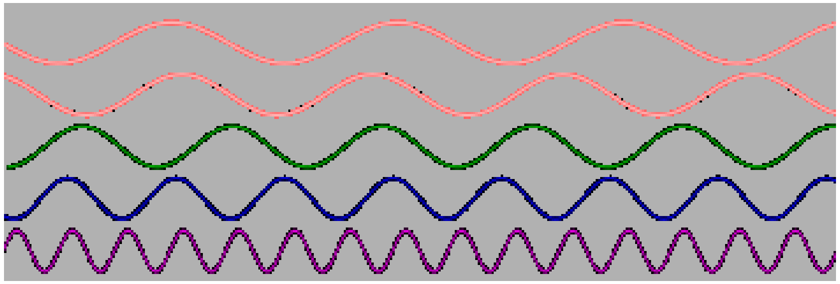
The reconstruction of Lees Ferry flow allows an assessment of decadal scale and longer characteristics, not possible with the shorter gage record.

Lees Ferry gage record

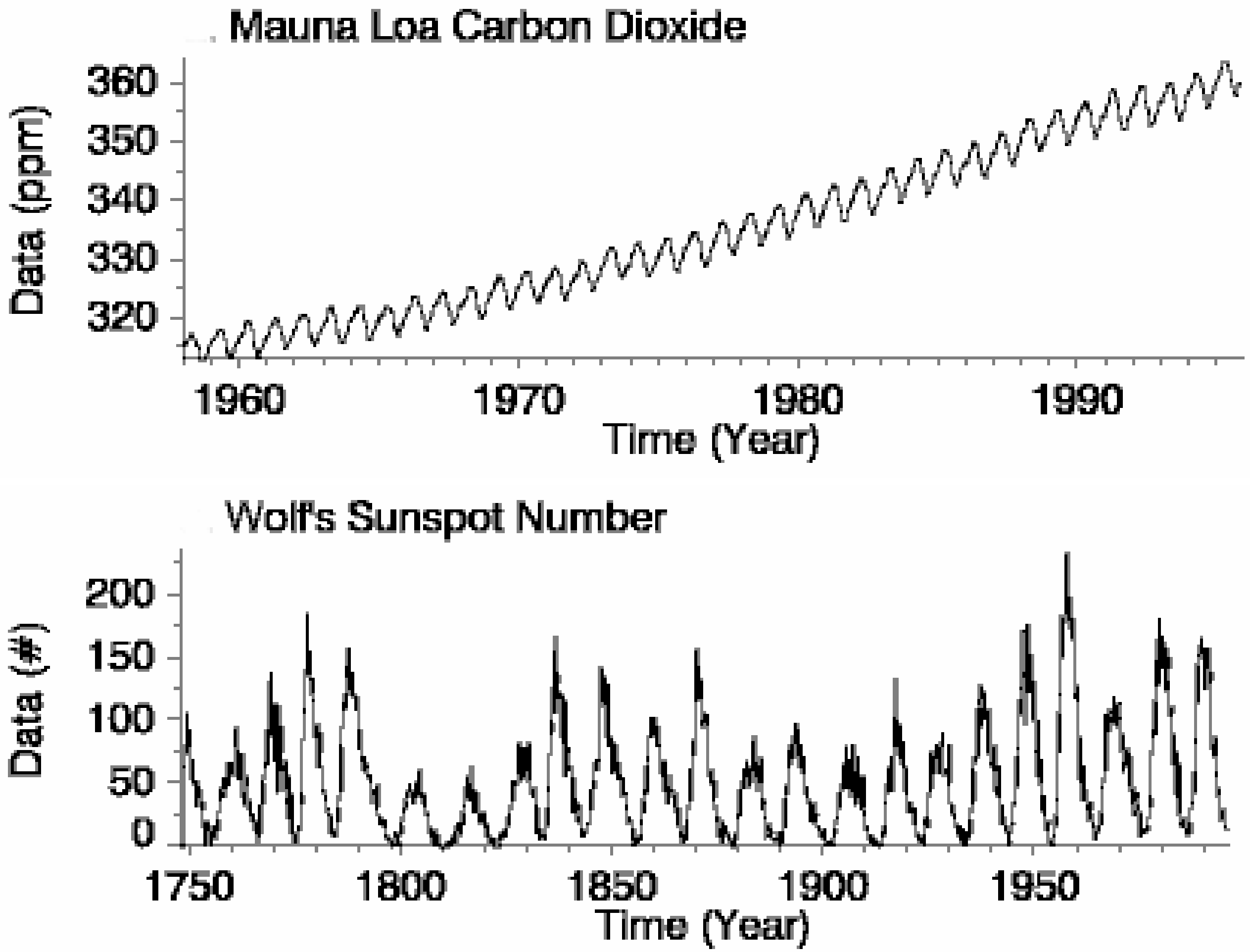
Lees Ferry reconstruction



Time series can contain variations that occur at regular intervals.

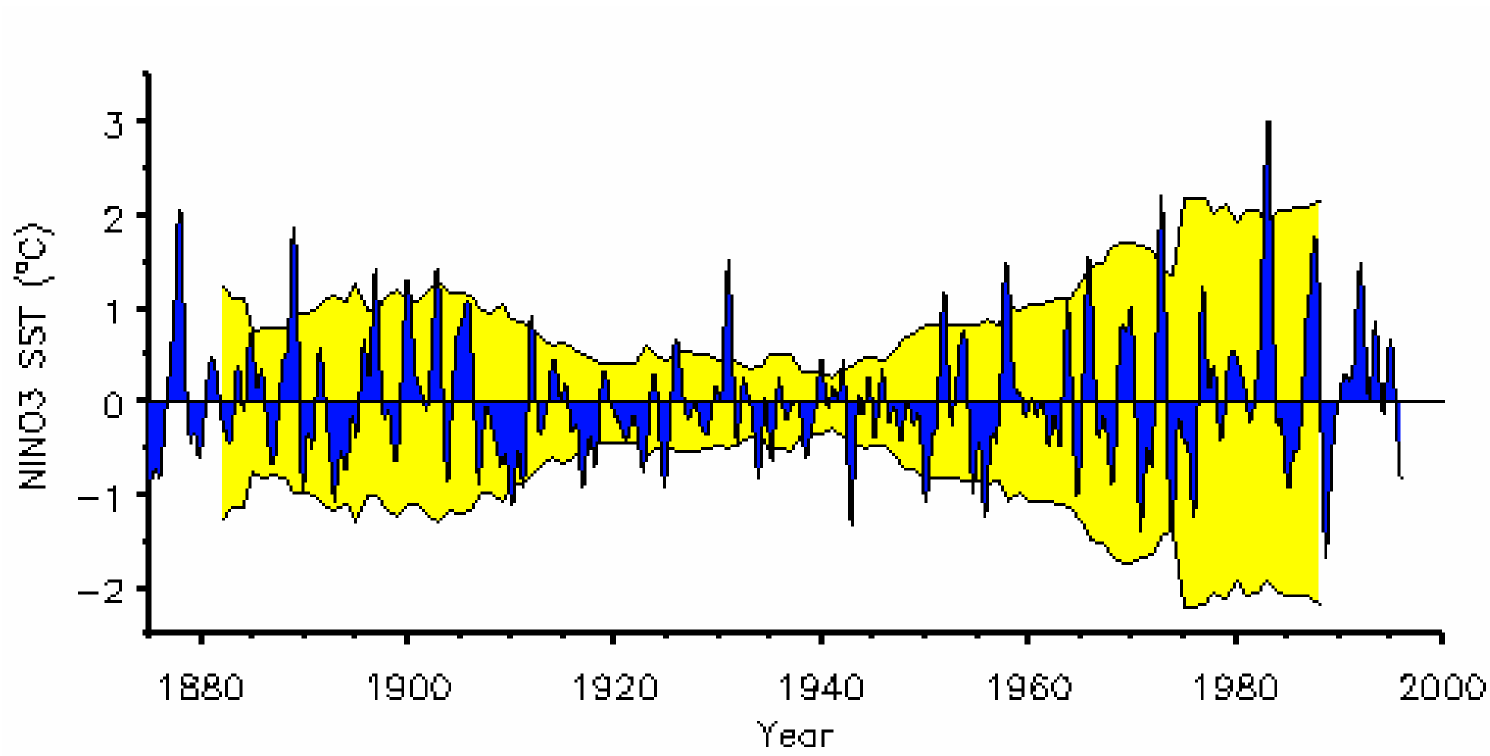


The intervals of variability, or periodicities, within a record may be dominated by a single frequency, as in sunspot numbers or annual CO2 values.



Torrence and Campo, http://atoc.colorado.edu/research/wavelets/

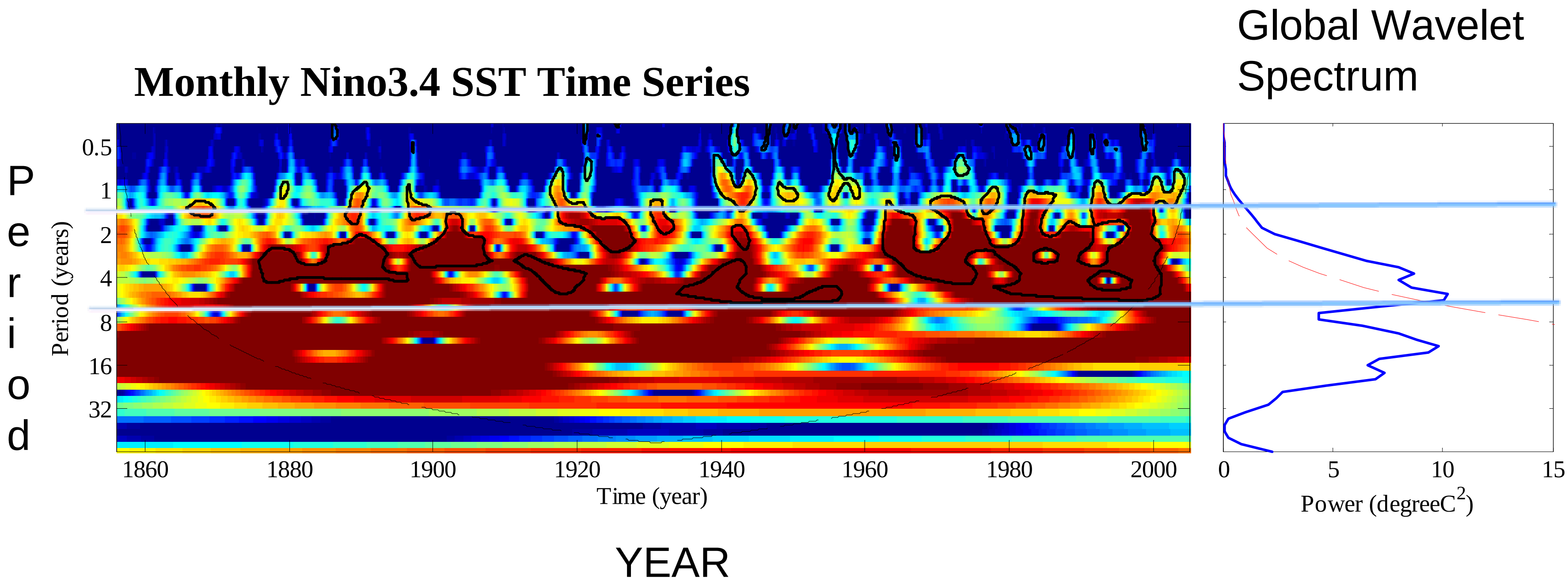
Time series may also contain variability at several different frequencies. ENSO has a periodicity of about 2-7 years, but the strength or amplitude of this periodicity varies over decadal time scales.



Sea surface temperatures averaged over the NINO3 region in the eastern Pacific (5°S-5°N, 90°W-150°W). Blue curve is low-pass filtered (>12 months) SST. Yellow background curve is running 15-year variance, plotted at mid-point of 15-year period. Curve has been reproduced upside-down to show "envelope" of variance.

Torrence and Campo, http://atoc.colorado.edu/research/wavelets/

Wavelet Spectrum of ENSO time series, the NINO3 index



Traditional Approach

- Short climate or hydrologic records are “extended” by generating psuedo samples with similar statistical properties
  - Parameteric (e.g. Autoregressive) or Non-Parameteric (e.g., resampling) methods
- Statistics preserved are usually mean, standard deviation and some short memory statistics like lag1 correlation
- The Colorado River record and other climate records show many long term departures from the mean that reflect long memory and regime like organization, that is not captured by these statistics used for judging the performance of traditional models

Overview – developing a new approach

• Multi-century Colorado River streamflow reconstructions show preferred regimes with recurrence structure or cycles.
• These time series characteristics, particularly those operating at decadal and multidecadal time scales, may not be evident in the length-limited gage records.
• Quantifying these low-frequency characteristics is critical, because they will underlie climate change, and be a primary determinant of the attributes of severe sustained drought.
• These low frequency components, once identified using the historical and the reconstructed paleo-data can be used as the basis for simulated flow in the future.
• Climate change models would need to show the ability to generate similar regional hydrologic variability in their "control" simulations to be credible for use with future anthropogenically forced simulations.

Wavelet**A**uto**R**egressive**M**ovingaverage **Approach:**

1.“Signal” Identification and isolation:

- Assemble a hybrid paleo-historical time series ensemble
    - 1900 to 2004 measured flows (+/- measurement error)
    - 1550 to 1899 reconstructed flow (+/- reconstruction error)
- Sample a streamflow sequence from the ensemble
    - 1550 to 2004 sequence
- Perform wavelet analysis on this sequence and identify frequency bands or cycles where a lot of variability is concentrated over the entire period of record, i.e, the probability that this amount of variability occurred purely due to chance is low.
- Repeat the previous 2 steps (sample+ wavelet) 100 times to sharpen the identification of the key cycles
- Reconstruct each cycle identified as a “signal”. Unlike regular spectral analysis the cycles identified will have phase and amplitude that varies with time.

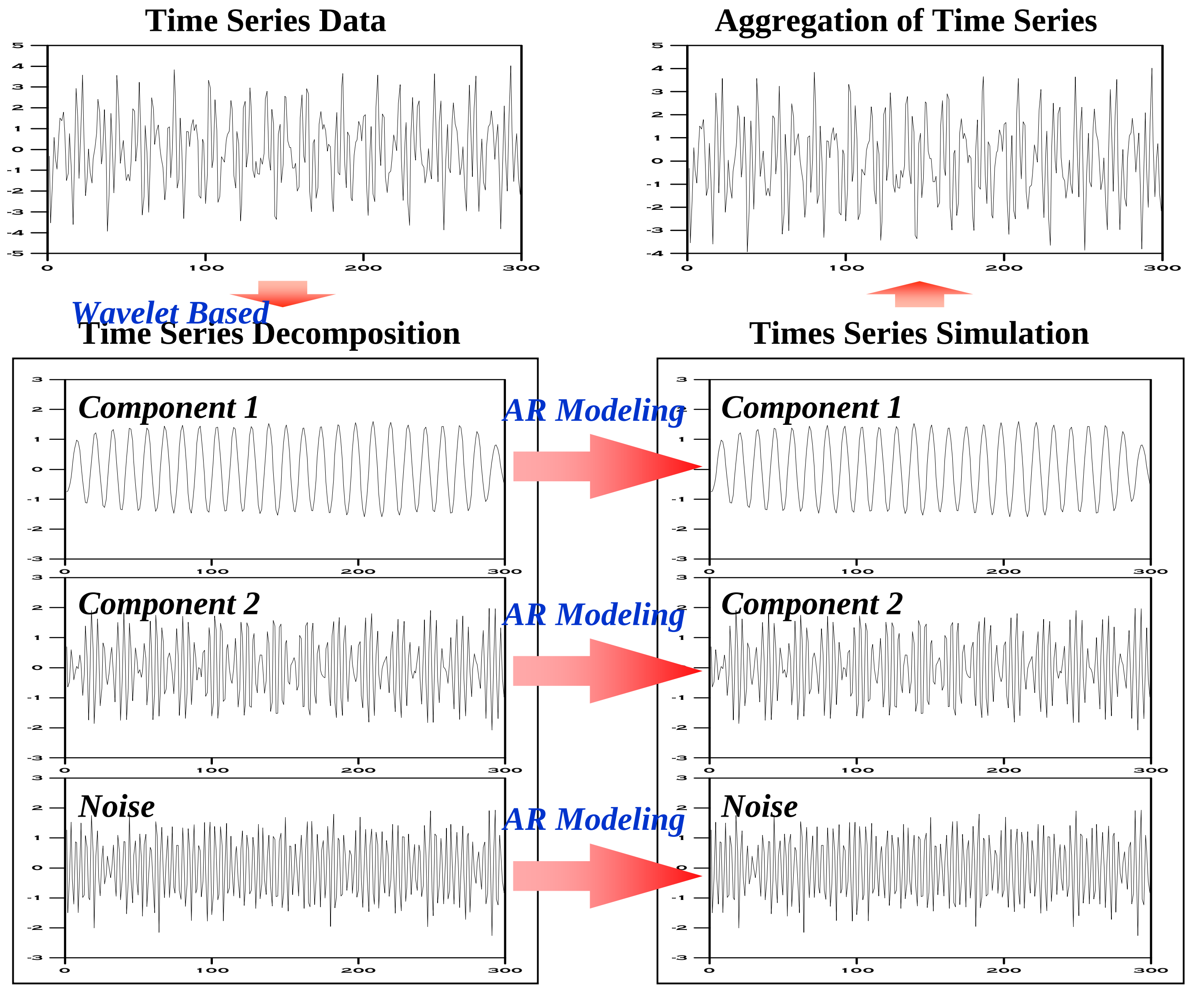
Wavelet**A**uto**R**egressive**M**ovingaverage **Approach:**

2. Develop a Simulation Model to generate new time series

- For each signal fit an ARMA (autoregressive moving average model) to capture the memory at the time scale of that signal
- For the “noise” or variability left after extracting the signals, fit an ARMA model to capture the short term memory in the system
- Simulate ARMA sequences for each signal and for the “Noise” and add them back up to generate a new simulation of the desired length (50, 100, 200, 500 years)

3. Compute the frequency of each drought severity and duration for each of the series, and develop the desired probability of a drought of a certain severity and/or duration

WARM Simulation



Wavelet Transform to WARM

$$Q_t = RC_{1,t} + RC_{2,t} + RC_{3,t} + \varepsilon_t$$

$RC_{i,t}$ is the i^{th} Reconstructed component or “signal” from the wavelet analysis of the original time series Q_t

Each RC and the noise term are modeled separately using an appropriate Autoregressive model with different number of terms p

$$Q_t = \sum_{k=1}^{K} AR(R_{kt}; p_k) + AR(\varepsilon_t; p)$$

Where

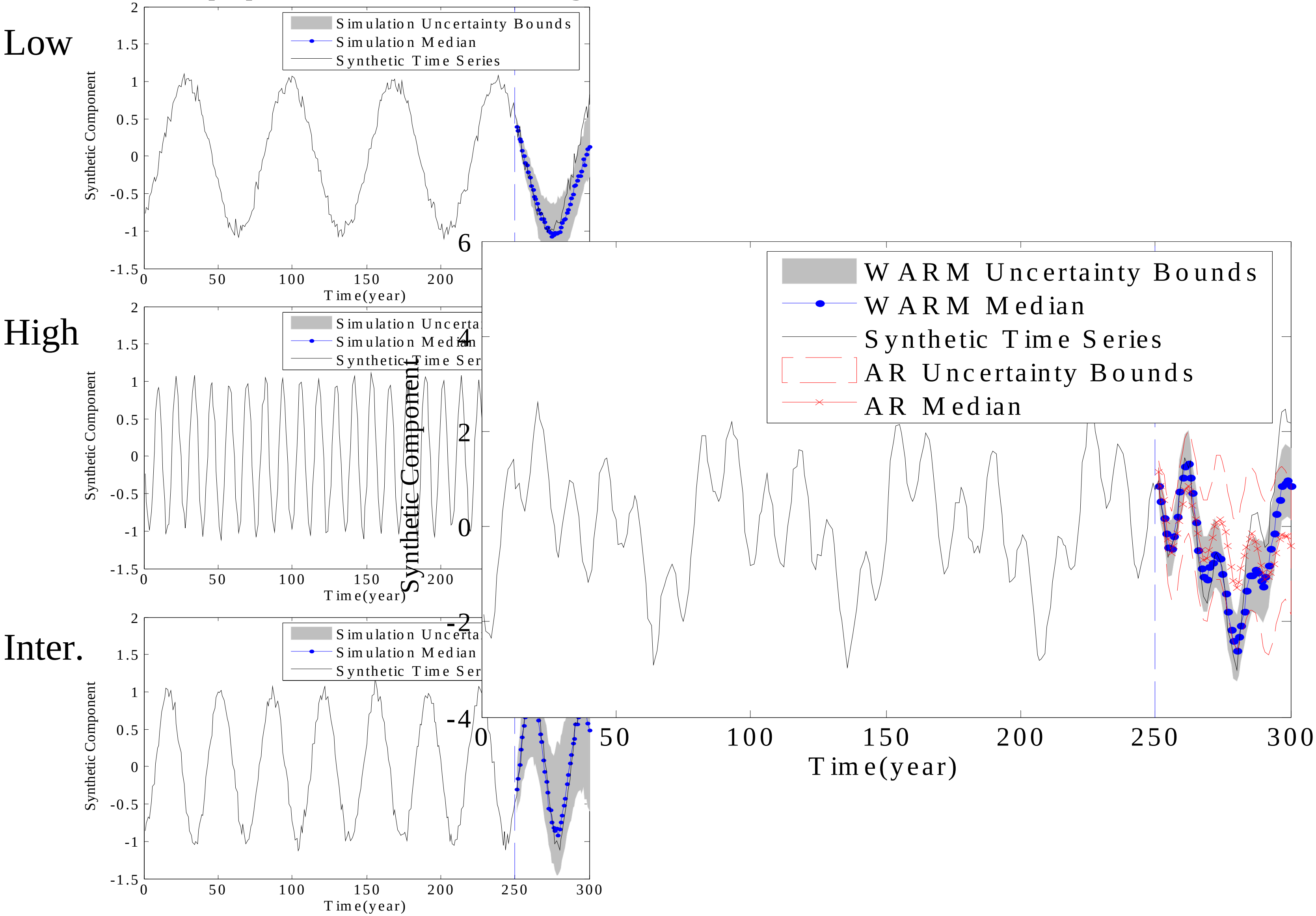
$$RC_1 = \alpha_{1,1} RC_{1,t-1} + \alpha_{1,2} RC_{1,t-2} + \alpha_{1,3} RC_{1,t-3} + \alpha_{1,4} RC_{1,t-4} + \nu_{1,t}$$

$$RC_1 = \alpha_{2,1} RC_{2,t-1} + \alpha_{2,2} RC_{2,t-2} + \alpha_{2,3} RC_{2,t-3} + \nu_{2,t}$$

$$RC_1 = \alpha_{3,1} RC_{3,t-1} + \alpha_{3,2} RC_{3,t-2} + \alpha_{3,3} RC_{3,t-3} + \nu_{3,t}$$

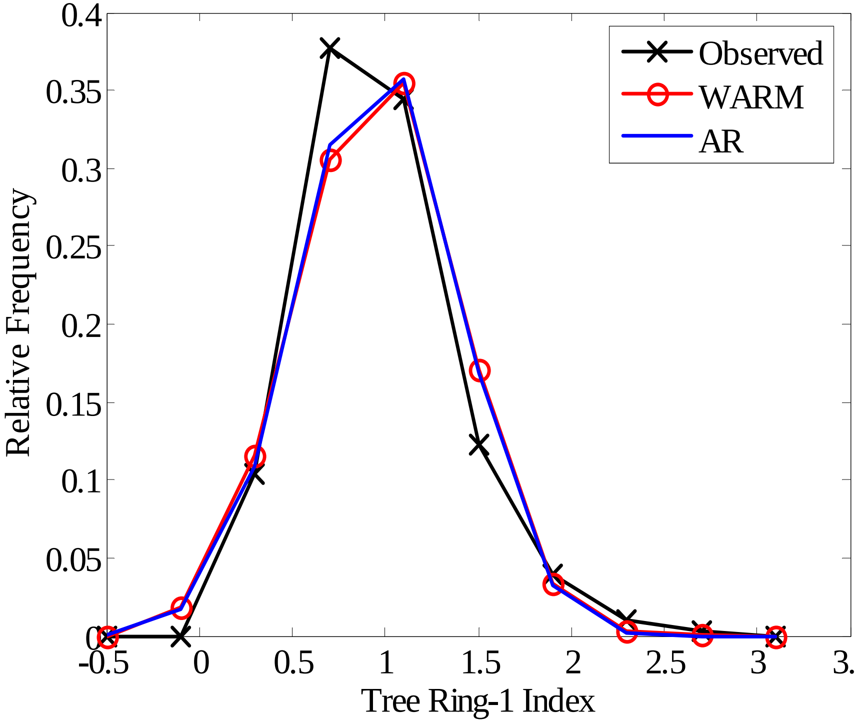
$$\varepsilon_t = \alpha_{\varepsilon,1} RC_{\varepsilon,t-1} + \alpha_{\varepsilon,2} RC_{\varepsilon,t-2} + \ldots + \alpha_{\varepsilon,6} RC_{\varepsilon,t-6} + \nu_{\varepsilon,t}$$

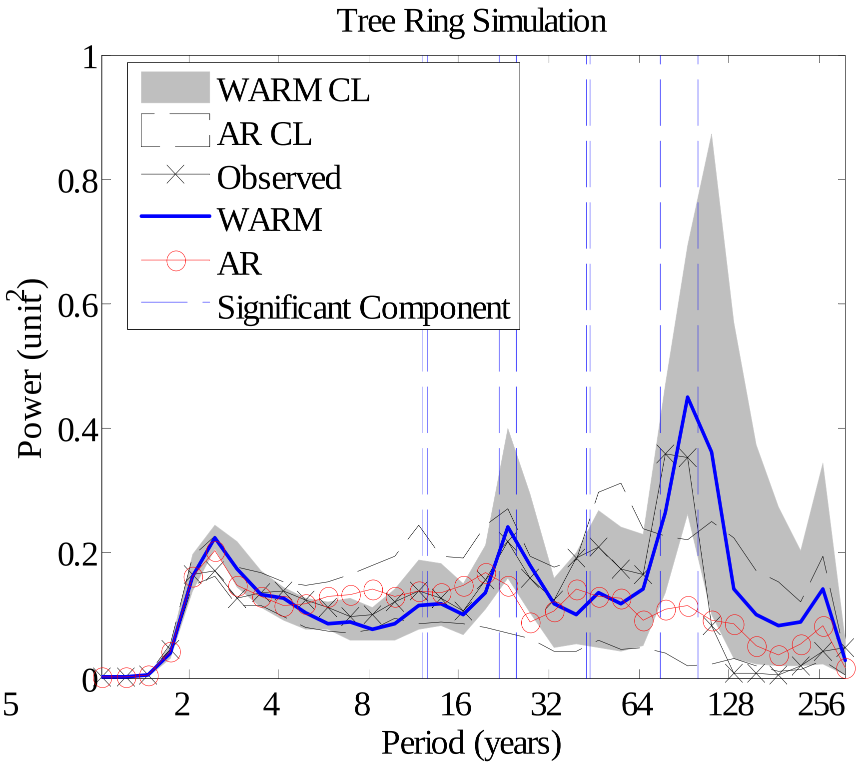
Application: Synthetic Time Series



WARM Simulation of Tree Ring Index in Florida

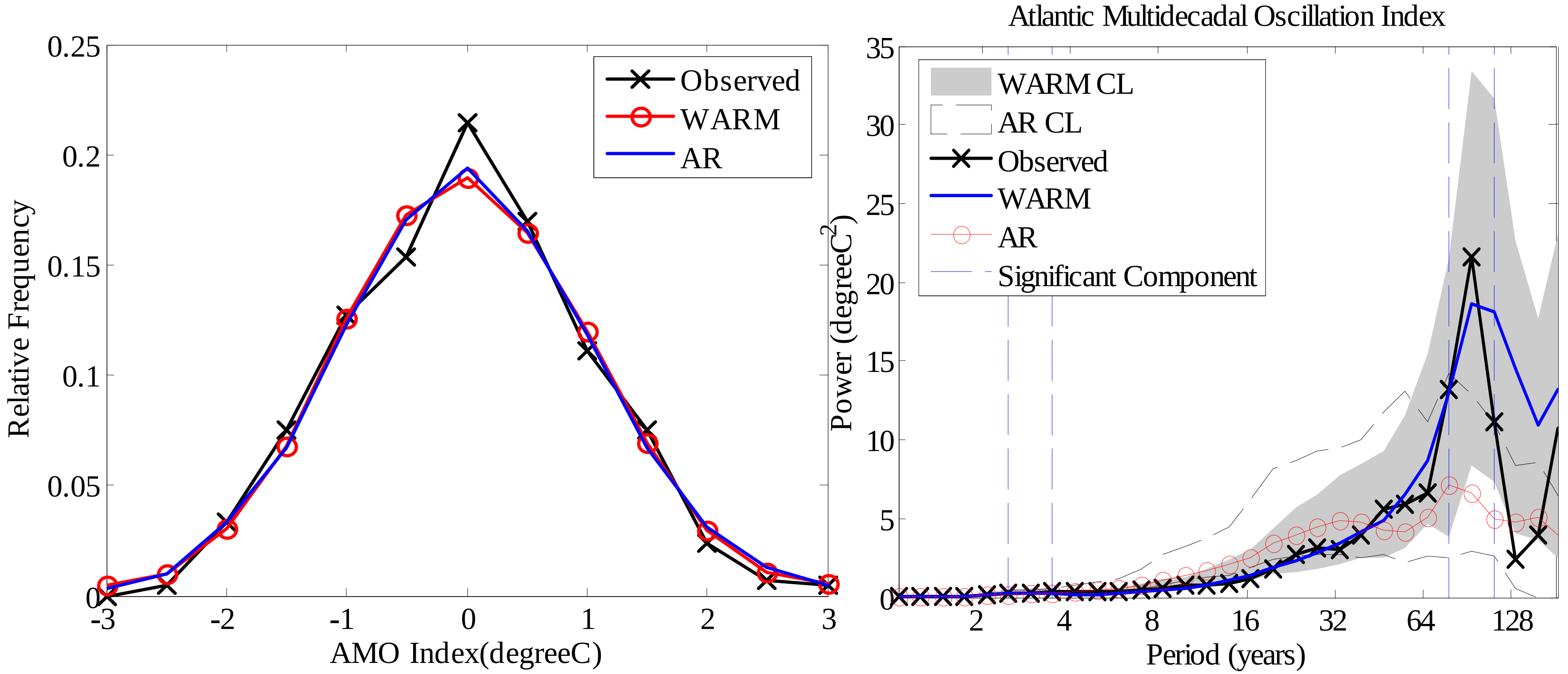
- 384 years Observed data are used.
- We note that the entire 10 year, 20 year, 40 year 60 year band has a GWP level higher than the significance level.
- Both the AR model and WARM model generally preserve the marginal distribution
- The WARM simulations are considerably better at reproducing the decadal and multidecadal spectral signatures in the original observations



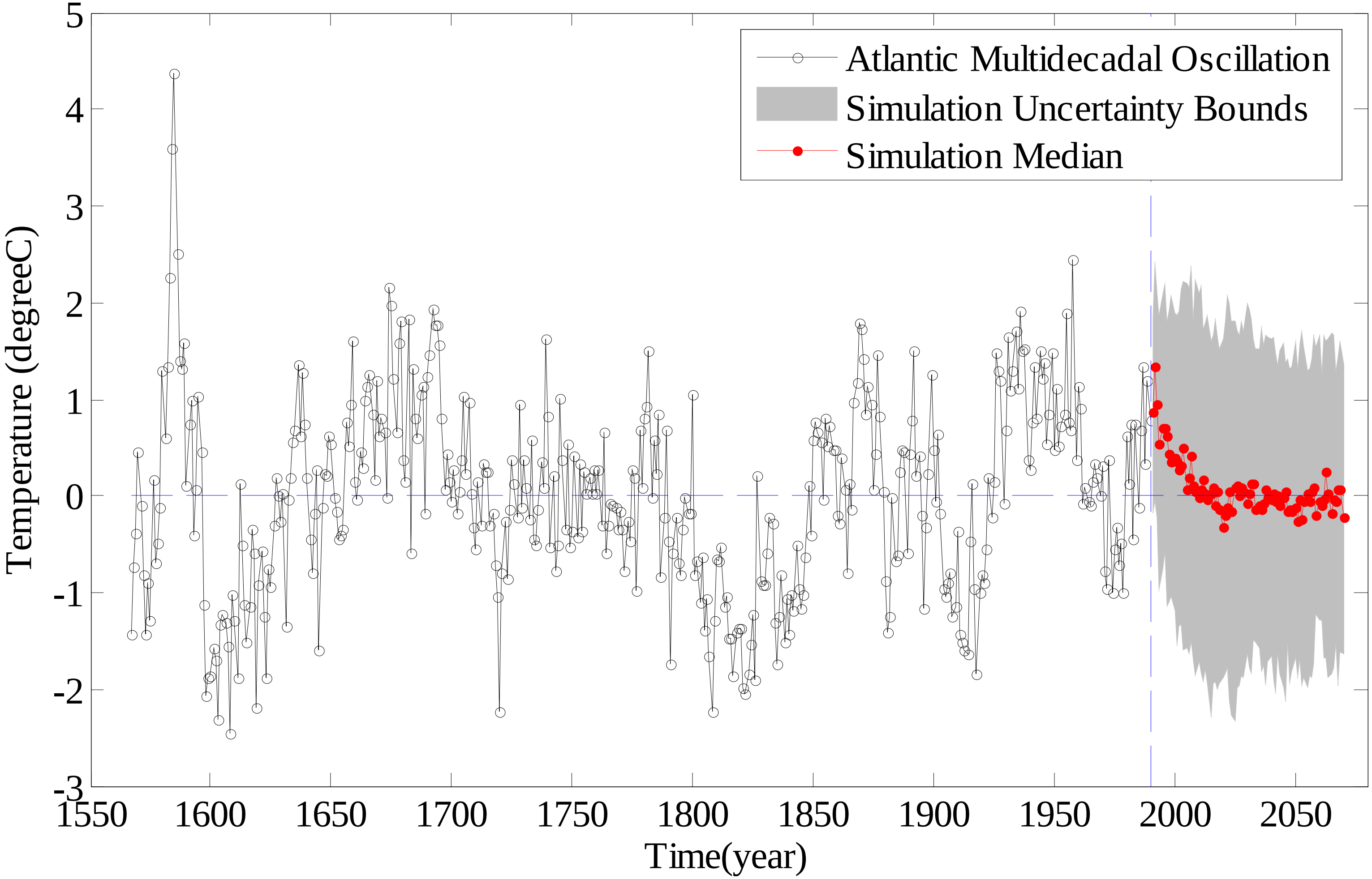


WARM Simulation of Paleo. AMO

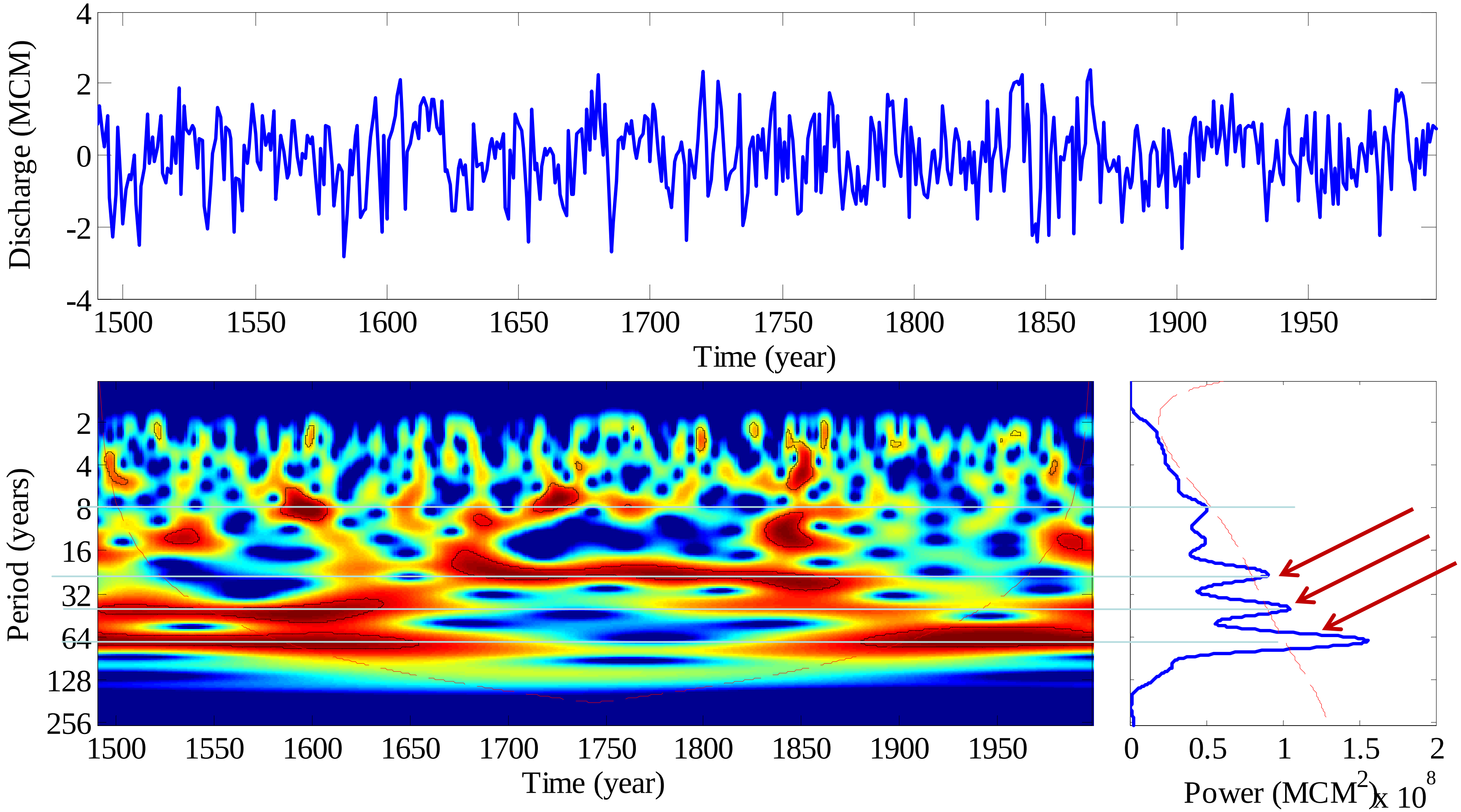
- 424 years Paleo. data are used.
- We note that the 3 year & 100 year band has a GWP level higher than the significance level.
- Both the AR model and WARM model generally preserve the marginal distribution
- AR spectrum tends to be rather flat over the frequency range considered, while the WARM simulations suggest a much sharper peak in the extreme low frequency end similar to that in the observed spectrum.



WARM Simulation of Paleo. AMO

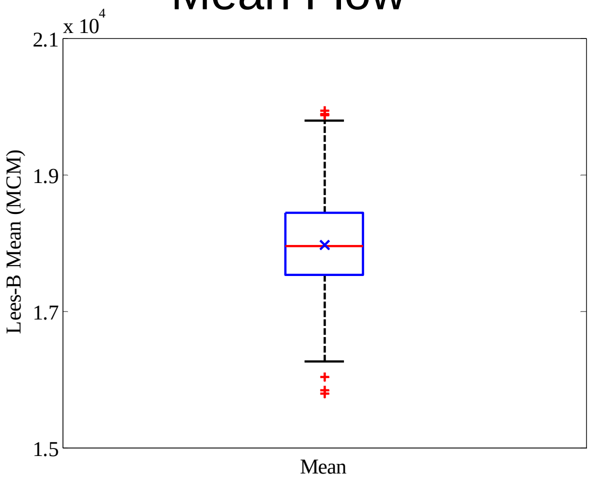


Lees-B FLOW, reconstruction, wavelet, and global wavelet

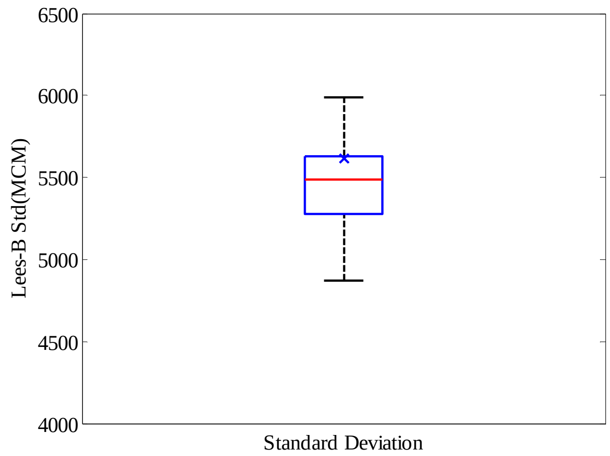


Results; modeling Lees-B

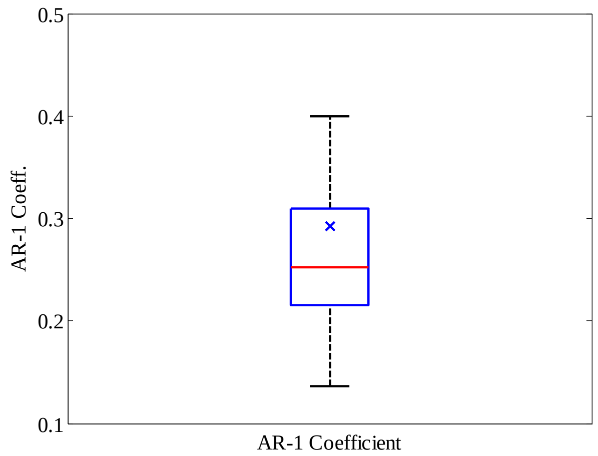
Mean Flow



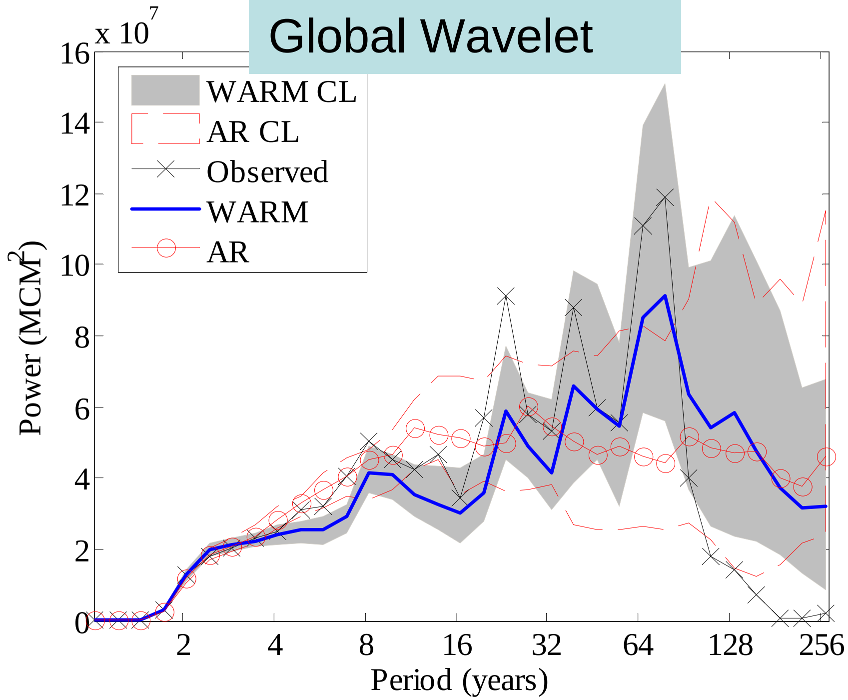
Std Dev of Flow



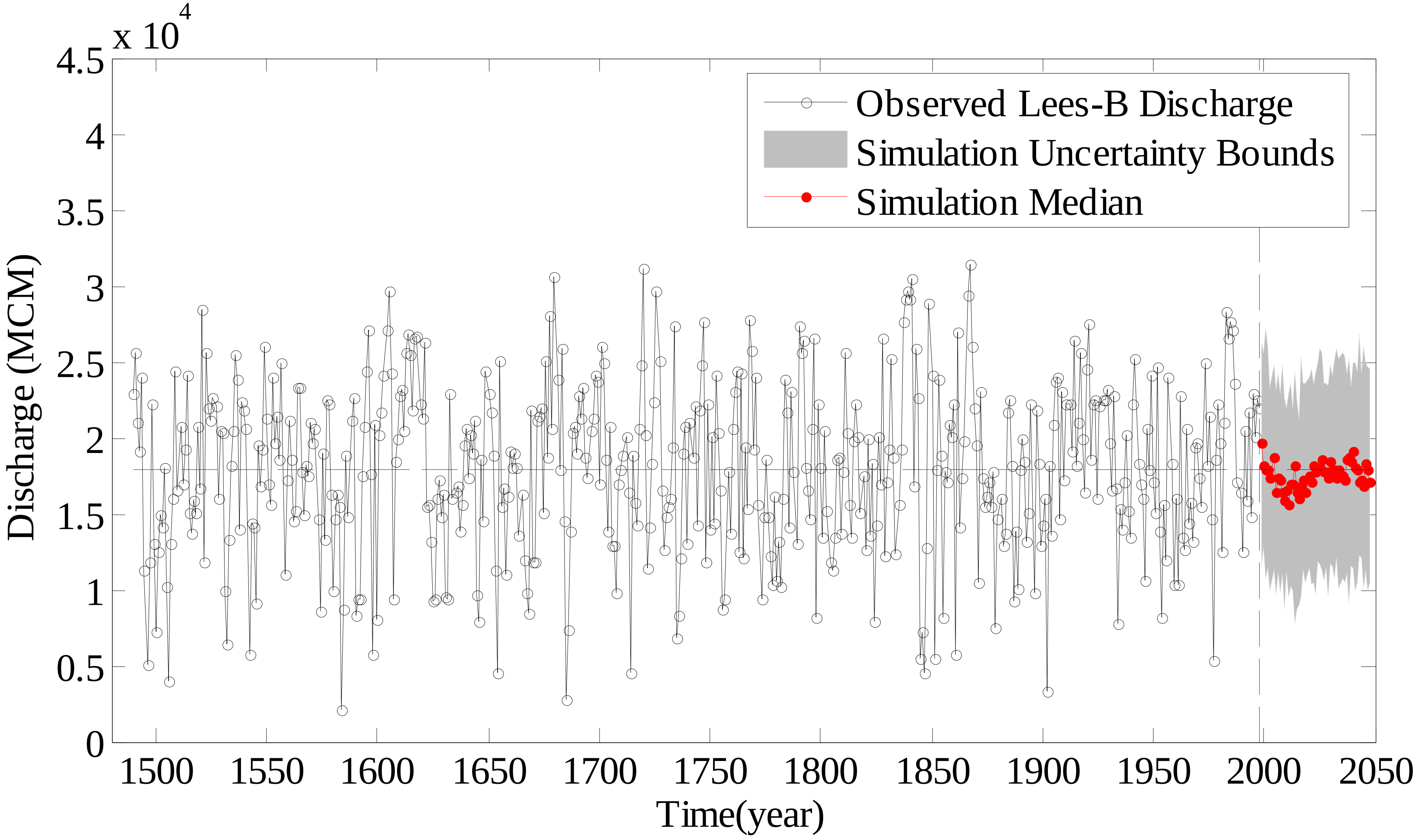
Lag 1 Correlation



Global Wavelet



Results; simulation



Applications and future work

• Analyses of the probability of a drought of specified severity, duration or both have been done using this technique for some sites in the Delaware river basin. Reservoir operating policy evaluations using these simulations are planned

•Work is in progress to identify long tree ring records in S. Florida and to use them to build simulations of seasonal and annual rainfall for S. Florida. These simulations will then be disaggregated to daily rainfall simulations using a Non-Homogeneous Markov Model that simulates rainfall considering some latent climate regimes that are informed by the signals identified by wavelet analysis – the simulations of each RC.

• A considerable amount of testing has been done with synthetic data, whose attributes are known, and testing with some known climate series for short term and long term forecasts and simulation is being done

• The possibility of using non-linear rather than linear models for simulation is being explored. However, so far most phenomenan are well reproduced by ARMA models applied to wavelet components

• We are also exploring how to decide the optimal number of wavelet components to keep to reduce the uncertainty of forecast/simulation without using too many parameters