Preserving long-term Variability in Simulation of Multisite Streamflow Extremes.

Abstract Number:
H13T-2050

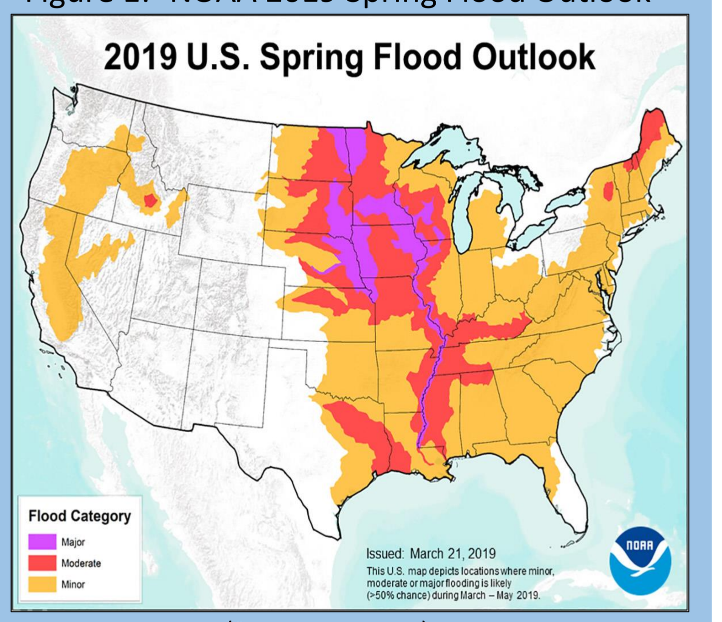
Yash Amonkar, James Doss-Gollin, Upmanu Lall
Department of Earth and Environmental Engineering, Columbia Water Center, Columbia University.

Columbia Water Center
Earth Institute | Columbia University

Introduction

- Floods led to more than half a million deaths worldwide from 1980-2009[1]. Clustered flood occurrences across large regions are particularly catastrophic. This has been observed in the Mid-western USA since the 1930s.
- Moseley (**1939**)[2]: "…the centers of greatest storminess over the United States and Canada shift in both latitude and longitude. After moving east for a number of years the location of these centers returns abruptly to a position much farther west. It may be that after ninety years the storm centers return to very nearly the same place which they had at the beginning."
- Brooks(**1937**)[3]"The conditions causing the excessive rainfall of up from 13 inches in southwestern Indiana to 12.7 in north- central Ohio in March 1913, were almost identical with those of January 1937"
- Record breaking floods across the Mid-west in 2019, surpassing the last great floods of 1993 seem to reflect a similar pattern. This has been the wettest 12 months in the history of United States since 1895[4], and the longest flood duration since 1927 on the Mississippi

Figure 1:- NOAA 2019 Spring Flood Outlook



(Source: noaa.gov)

Research Questions:-

1. Are low-frequency variations present in streamflow/flooding extremes ?
2. What can we infer as to space-time clustering of floods from historical streamflow data from the Ohio River Basin?
3. Are there statistical models that can represent the observed low frequency variations ?

1 - Doocy, S., Daniels, A., Murray, S., & Kirsch, T. D. (2013). The human impact of floods: a historical review of events 1980-2009 and systematic literature review. *PLoS currents*, *5. 2*- Moseley, E. L. (1939). Long time forecasts of Ohio River floods. 3- Long time forecasts of Ohio River floods. 3- Brooks, C. F., & Thiessen, A. H. (1937). The meteorology of great floods in the eastern United States. *Geographical Review*, *27*(2), 269-290. 4- https://www.ncdc.noaa.gov/cag/national/time-series

Data Sources

- Streamflow Gauges - USGS/Data Retrieval Package in R[1].
- Climate Indices[2] .ENSO – Nino 3.4 – HadISST. PDO – JISAO. AMO – HadSST. NAO – Jones et al.

1. Hirsch, R.M., and De Cicco, L.A., 2015, User guide to Exploration and Graphics for RivEr Trends (EGRET) and dataRetrieval: R packages for hydrologic data (version 2.0, February 2015): U.S. Geological Survey Techniques and Methods book 4, chap. A10, 93 p., http://dx.doi.org/10.3133/tm4A10
2. KNMI Climate Explorer - https://climexp.knmi.nl/

Methods

Streamflow

- Daily Data from 28 stations for the period 1937-2017 were used and converted to annual data using site-specific annual maximum as the extremity metric.
- Streamflow gauges were selected which had a drainage area greater than 1450 sq. Miles.

PC-Wavelet Clustering → Space Time Clustering

- Principal Component Analysis → Dimension Reduction
- Wavelet Decomposition → Extracting signals over noise.

Another option, computed but not presented in this poster is to invert the chronology and look at 'time' first(wavelet decomposition) followed by spatial domain(PCA). This method leads to explicit-hard clustering.

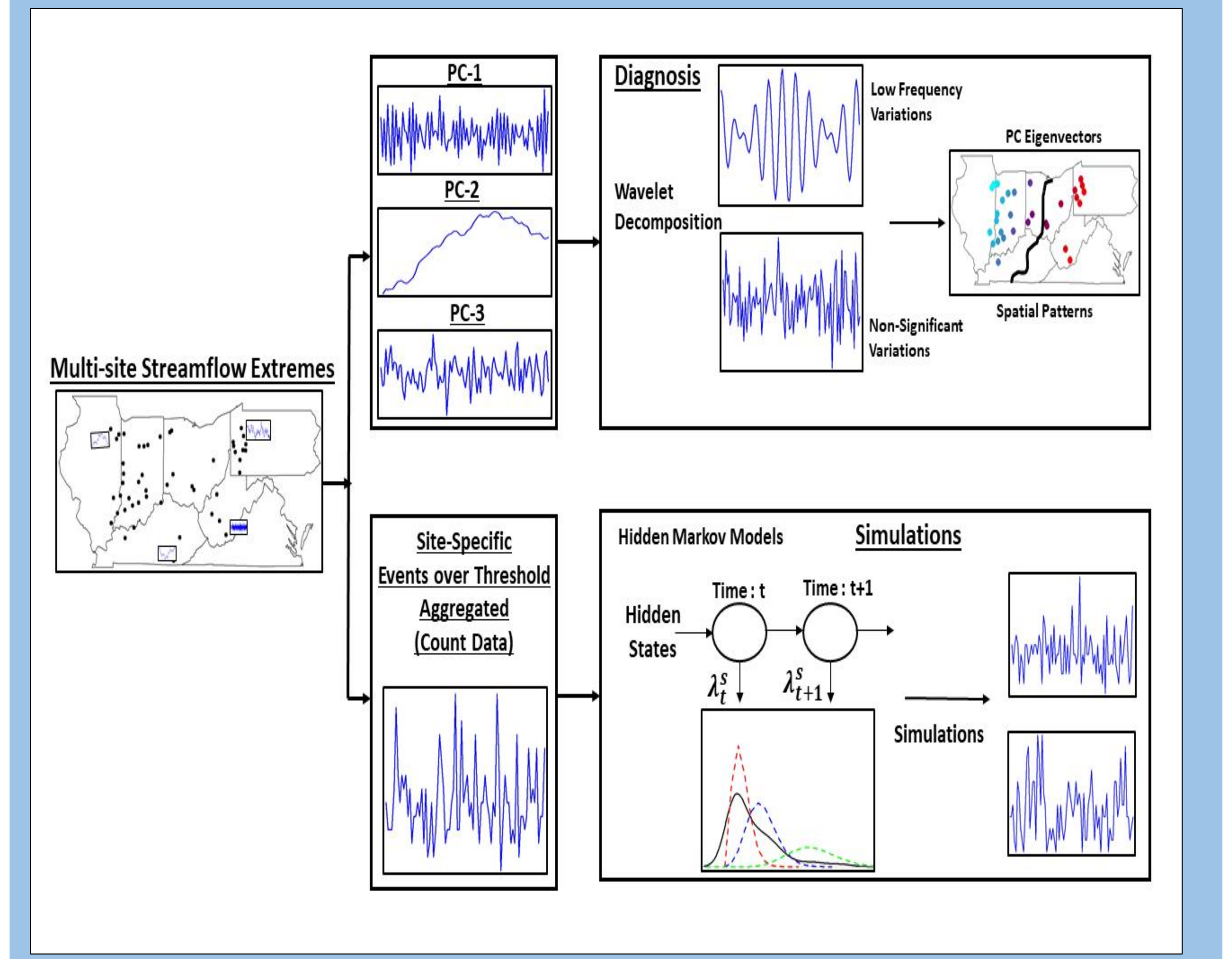


Figure 2: Schematic of overall methodology.

Hidden Markov Model → Aggregated Persistence Structure

- Convert the site annual maximum streamflow to exceedances based on a high value e.g. 90[th] percentile.
- Aggregate it across the sites, fit with a Poisson emission distribution.
- Simulate additional runs based on the computed transition matrix and emission parameter rates.

Results

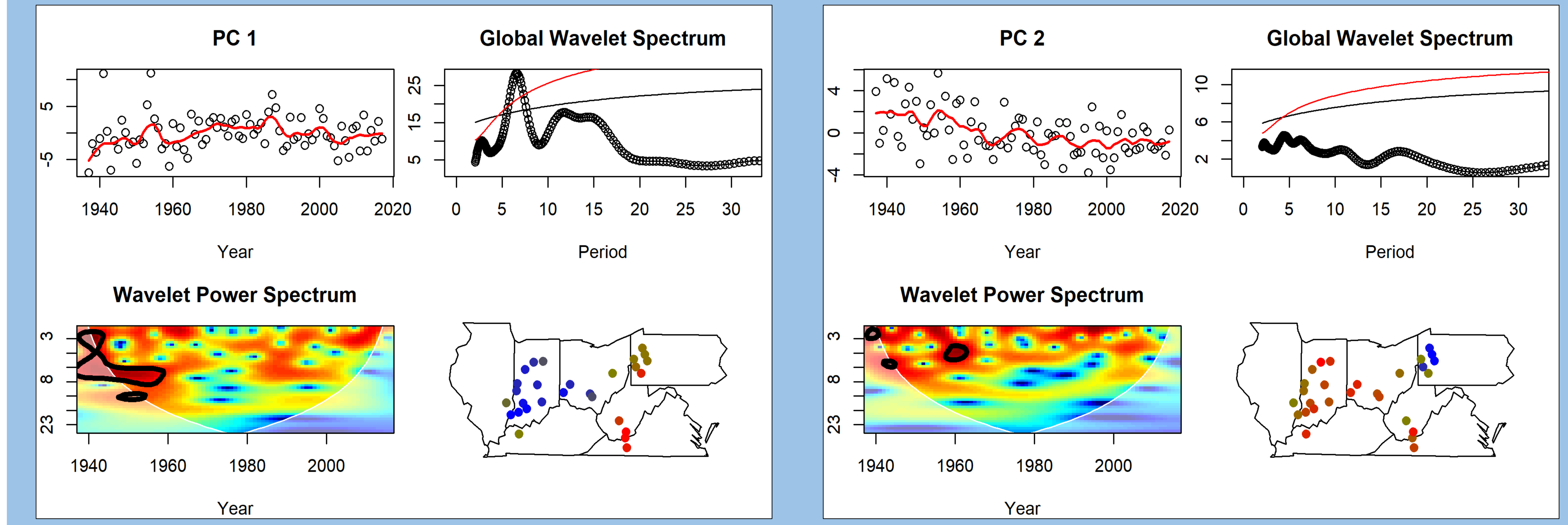


Figure 3 :- Wavelet Analysis on the Principal Components (a) PC-1 (b)PC-2. TOP-LEFT –PC with a loess line. TOP-RIGHT – Global Wavelet Spectrum of the PC. BOTTOM-LEFT – Power Spectrum of the Wavelet (Regions bounded in the black line are statistically significant at the 90% level). BOTTOM-RIGHT – Absolute values of the loadings(eigenvectors) color code Red to Blue.

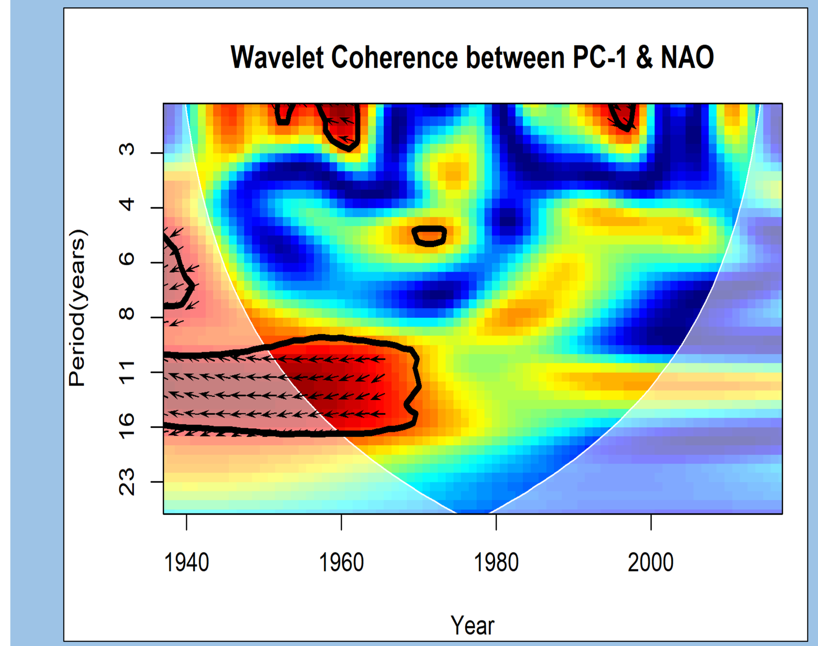


Figure 4: Wavelet Coherence between PC-1 and North Atlantic Oscillation

The spatial pattern of the 1[st] PC is based in the western region of Ohio River Basin, accounting for ~40% of total variance and has a characteristic 7-yr low frequency cycle.

PC-1 is significantly correlated with NAO and the interaction between [ENSO-PDO], pointing to presence of teleconnections in the region. The relationship between NAO and PC-1 was strong in the earlier decades.

The 2[nd] PC is based in the eastern part of Ohio River Basin, explains about 17% of the variance and has a secular trend.

Component	Correlation with PC-1
ENSO	0.21
NAO	-0.31***
PDO	0.16
ENSO-NAO	0.01
ENSO-PDO	0.27**
NAO-PDO	0.01

Table 1: Correlation between PC-1 and climate indices.

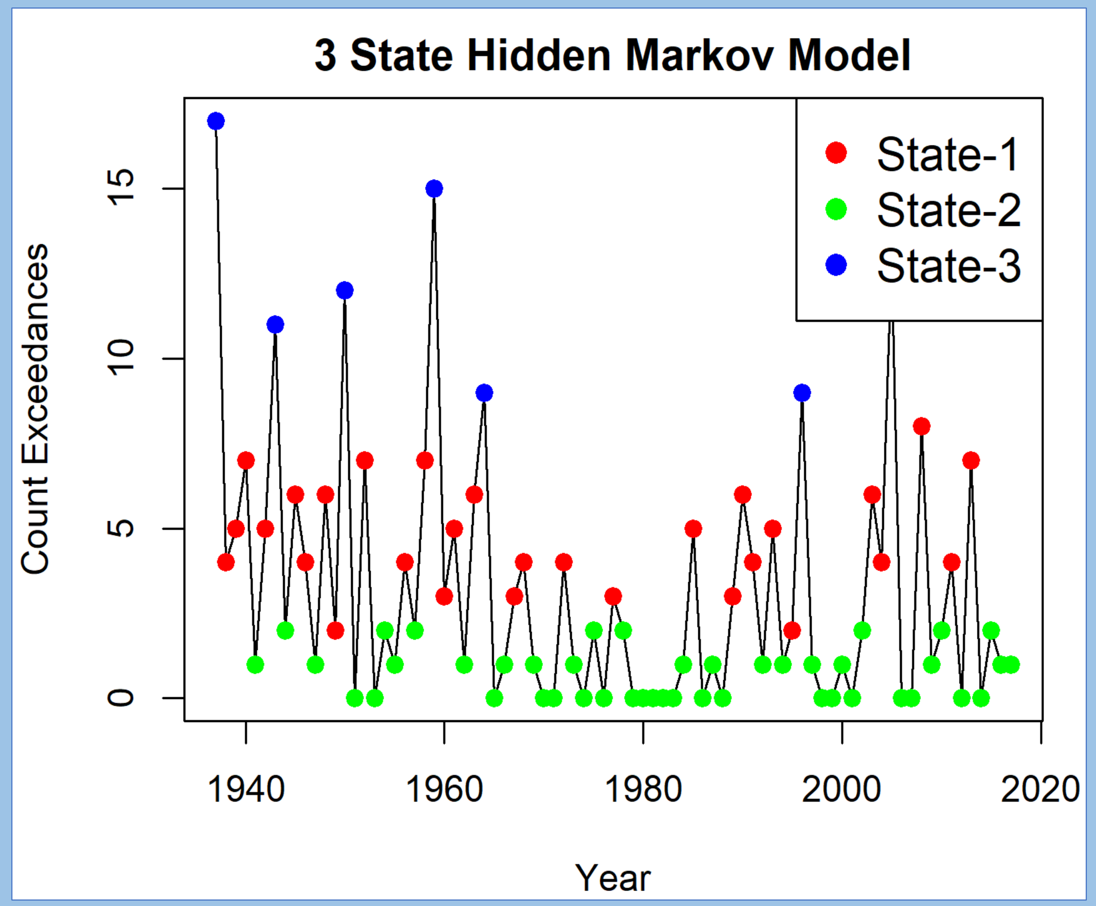
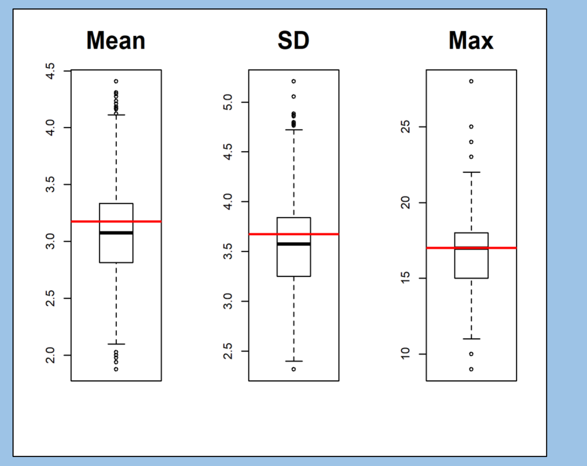
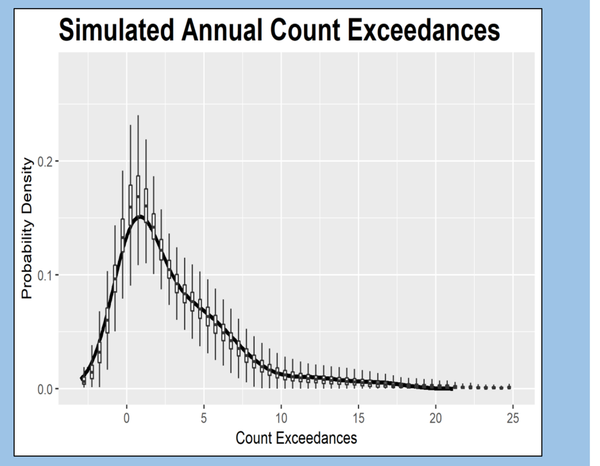
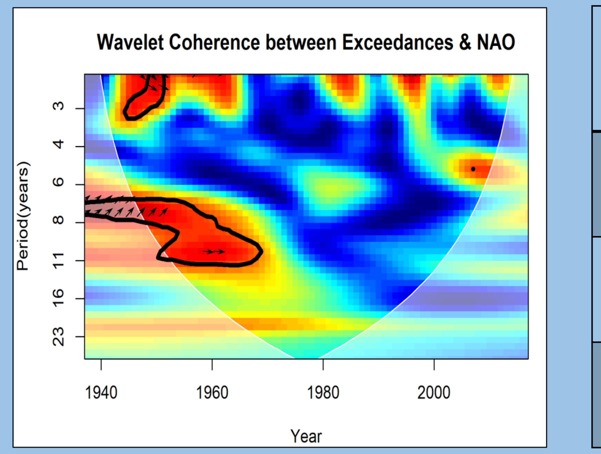


Fig 5:- Count Threshold Exceedances with the hidden states.







to/from	S1	S2	S3
S1	0.36	0.47	0.17
S2	0.46	0.54	0
S3	0.32	0.68	0

Fig 6:- Simulation Skill for (a)- Mean, Standard Deviation and Maximum replication skill. (b) PDF (c) Wavelet Coherence between exceedances and NAO (d) Transition Matrix

There is no persistence in the extreme wet state and no sudden jump from the dry to wet state, providing a measure of predictability for the extreme streamflow events. The Poisson rate parameters associated with the three hidden states are:-

$\lambda =$ 0.7 events/year → Dry State,
4.56 events/year → Intermediate State
12.34 events/year → Wet State.

Conclusions

- We have presented a general methodology for streamflow risk management accounting for low-frequency variations and spatio-temporal clustering on a river basin scale. The Ohio River Basin shows significant clustering in the regional streamflow regime with ramifications in the understanding of the local climatology and for the design of critical infrastructure in the region.
- The Hidden Markov Models help capture the nature of truly extreme events over large scales. The fat tails in the aggregated spatial domain across the ORB is captured and well simulated by a Poisson emission based Hidden Markov Model.
- Large scale climate drivers potentially induce space and time clustering in the occurrence of flooding in this region, even without anthropogenic climate change, with significant regional impact on losses and supply chains.

Acknowledgements/Contact Information

The authors thank the researchers at the Columbia Water Center for their suggestions and comments.
Correspondence – yva2000@columbia.edu