Hydroclimate Drivers and Atmospheric Teleconnections of Long Duration Floods: An Application to Large Reservoirs in the Missouri River Basin

Nasser Najibi^{1,2,3}, Naresh Devineni^{1,2,3} and Mengqian LU⁴

¹Department of Civil Engineering, City University of New York (City College), New York, NY 10031, USA

²Center for Water Resources and Environmental Research (City Water Center), City University of New York (City College), New York, NY 10031, USA

³NOAA-Cooperative Remote Sensing Science and Technology Center (NOAA-CREST), City University of New York, New York, NY 10031, USA

⁴Department of Civil and Environmental Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong SAR

Corresponding authors: Nasser Najibi (nnajibi@ccny.cuny.edu) and Naresh Devineni (ndevineni@ccny.cuny.edu)

†Current/Permanent Address:

160 Convent Avenue, Steinman Hall 106, New York, NY 10031, USA

Abstract

A comprehensive framework is developed to assess the flood types, their spatiotemporal characteristics and causes based on the rainfall statistics, antecedent flow conditions, and atmospheric teleconnections. The Missouri River Basin (MRB) is used as a case study for the application of the framework. Floods are defined using the multivariate characteristics of annual peak, volume, duration, and timing. The temporal clustering of flood durations is assessed using a hierarchical clustering analysis, and low-frequency modes are identified using wavelet decomposition. This is followed by an identification of the synoptic scale atmospheric processes and an analysis of storm tracks that entered the basin and their moisture releases. Atmospheric teleconnections are distinctively persistent and well developed for long duration flood events. Long duration floods are triggered by high antecedent flow conditions which are in turn caused by high moisture release from the tracks. For short duration floods, these are insignificant and random across the MRB's in the recent half-century. The relative importance of hydroclimatic drivers (rainfall duration, rainfall intensity and antecedent flow conditions) in explaining the variance in flood duration and volume is discussed using an empirical log-linear regression model. The implication of analyzing the duration and volume of the floods in the context of flood frequency analysis for dams is also presented. The results demonstrate that the existing notion of the flood risk assessment and consequent reservoir operations based on the instantaneous peak flow rate at a stream gage needs to be revisited, especially for those flood events caused by persistent rainfall events, high antecedent flow conditions and synoptic scale atmospheric teleconnections.

Keywords: Long Duration Floods, Flood Characteristics, Antecedent Streamflow, Persistent Rainfall, Atmospheric Teleconnection.

1. Introduction

Recent mega-floods in Midwest USA, Thailand, Pakistan, Queensland, India, China, and Europe have placed risk assessment for floods at the forefront. In some cases, such as Thailand and the Mississippi River, the efficacy of the flood control projects and their operation have been called into question [Ziegler *et al.*, 2012; Promchote *et al.*, 2015]. Areas not previously considered a major risk had industrial infrastructure inundated by flooding, leading to substantial global supply chain effects in addition to the direct loss of use of assets [The World Bank, 2011; Haraguchi and Lall, 2015]. While some of these floods of interest (e.g. Mumbai 2005) were attributed to a single intense rainfall event, several were associated with multiple, recurrent events that led to floods of durations of 30 to 170 days. Past flood risk analyses did not formally consider the risk of such long duration flood events [Ward *et al.*, 2016]. There is little to no literature on how to estimate and link the probability of persistent rainfall over 30 to 120 days that leads to high antecedent moisture conditions in the region, as a basis for projecting the risk of mega-floods in a region.

The single event floods conform to the traditional view of flood risk analysis, where a single extreme event (e.g. the tropical cyclone-induced rain) appears to occur randomly, and predictability may be limited to a few hours to a day. On the other hand, the events related to persistent and recurrent rainfall may correspond to the persistence of particular global climate patterns. For example, Nakamura *et al.* [2013] and Robertson *et al.* [2011] recently demonstrated that the 2011 flood in the Ohio River Basin is due to repeated waves (recurrent advection) of

tropical moisture every 5 to 7 days leading to high antecedent moisture conditions and subsequent river overflows. Similarly, *Smith et al.* [2013] showed that the June 2008 Iowa flood was produced by a sequence of organized thunderstorm systems over a period of two weeks. The repeated tropical cyclones and rainfall in Thailand and Queensland in 2011, and the 2015 floods in Chennai, India provide more examples of persistence in the moisture delivery system.

The high degree of spatiotemporal variability in the meteorological processes that result in large area inundation and regional floods and pose challenges to reservoir management is also of interest in this context. The recent 2011 Missouri River Basin (MRB) flood is an example. According to the U.S. National Weather Service (NWS) Office of Climate, Water and Weather Services, an anomalously high snowfall, compared to conditions typical of the late 20th century (130% of average for April 1st) occurred during the winter of 2011 in the Rocky Mountains and Northern Plains of the MRB. An anomalously cooler spring that delayed the snowmelt followed this event. The rapid snowmelt during the late spring coincided after that with the record-setting rainfall events in May and early June 2011 over Montana and western North Dakota [NWS-NOAA, 2011]. Eventually, the combination of these events caused a record river and reservoir levels as well as extensive flooding over the Missouri and Souris River Basins (along the states of Montana to Missouri) from May through August 2011. This flood event resulted in extensive damage to almost one-third of the homes located in the states of Missouri, North Dakota and Kansas [NWS-NOAA, 2011]. Winter snowpack conditions in the mountainous headwaters are strongly linked to factors affecting the position of the jet stream and the Pacific North American (PNA) pattern of flow [Dettinger *et al.*, 1998; Brown and Comrie, 2004; Woodhouse *et al.*, 2010]. The phase and strength of El Niño Southern Oscillation (ENSO) and Northern Pacific decadal variability (e.g. PDO) also influence winter snowpack in MRB. Hoerling *et al.* [2013] in their climate assessment report showed how a previous La Nina climate pattern that aided the shift of position and strength of the Jet Stream set the stage for the 2011 MRB flood event. Recently, Archfield *et al.* [2016], in their work on trend analysis for flood frequency, magnitude, duration and volumes, demonstrated that among several climate indices, ENSO has statistically significant correlations to flood duration and volume at lag times of 0 to 6 months for approximately 25% of the 345 streamflow stations across the United States.

A multitude of such recent developments motivated us to study and develop a climate-informed flood risk assessment framework. It is important to explicitly understand the dependence of the likelihood or frequency and intensity of extreme regional floods on a causal chain of ocean-atmosphere processes whose slow variation and regime-like changes translate into significant and persistent changes in the probability of major floods in the large river basins. Mapping of these factors into a dynamic risk framework is necessary for establishing a process by which flood risk for large basins could be systematically updated reflecting changing climate conditions, or as part of the natural cycles of climate variation. We define a flood event, not just through the annual peak flow, but also through attributes such as flood volume, duration, and time of occurrence, i.e., in a multivariate context. It will help us to understand better, the effect of each attribute on flood control and damage mitigations strategies [de Moel *et al.*, 2015].

In this study, we attempt to develop an exploratory data analysis based inference system for flood risk assessment using regional climate information and atmospheric teleconnections. First, we develop multivariate flood attributes and classify their spatial variability using geographic characteristics, and temporal variability using the hierarchical clustering [Hartigan, 1975] approach. Since there may exist a systematic structure in the variability, we employ wavelet

decomposition to understand the dominant modes of variability in flood duration, i.e. the low-frequency variability that could eventually be attributed to climate mechanisms. Depending on the flood event type, different rainfall inducing mechanisms (e.g. tropical storm, local convection, frontal system, recurrent tropical waves) may be involved with characteristic spatial scales and statistical properties. Hence, we identify the flood types and map their corresponding specific atmospheric circulation patterns using compositing of the NCEP/NCAR reanalysis data. One can then develop stochastic models that can reproduce these attributes with appropriate intensity-duration-frequency and spatial expression and provide a basis for conditioning basin hydrologic attributes for flood risk assessment. We choose the MRB for the implementation of the framework given it is one of the longest rivers draining approximately one-sixth of the contiguous U.S. [Galat *et al.*, 2005]. We also present the case for developing the flood frequency analysis in a multivariate context.

Section 2 presents the data and the MRB context. In Section 3, we introduce the methodology of the statistical inference system to identify the spatiotemporal properties of floods in MRB and discuss the results. In Section 4, we present the case for multivariate flood frequency analysis and provide the implications for managing the flood control pool. Finally, in Section 5 we present the summary and concluding remarks.

2. Data and Missouri River Basin Context

The Missouri River Basin (MRB) as an essential part of Mississippi River System encompasses around one-sixth of the U.S. (1,371,000 km²) including the state of Montana, North Dakota, South Dakota, Nebraska, Iowa, Kansas, Missouri and parts of Colorado, Minnesota, and Wyoming. The basin is stretching from the Rocky Mountains in the west to the Mississippi River Valley in the east and from the southern extreme of western Canada to the border of the Arkansas River watershed (diametric extent of Longitude: -111°, -90° and Latitude: 48°, 38°, respectively). The Missouri River in the MRB is the longest river in the US and the second longest in North America at 4,180 km. The headwaters are in the Rocky Mountains, where snowmelt is the largest source of water. The river then flows east across the Great Plains to its confluence with the Mississippi River. Spring-to-early-summer rains are the dominant source of moisture across this region. Figure 1(a) presents the detailed information about the MRB geographical extent, topographical properties, and the river network, together with co-located dams, streamflow stations, and nearest rainfall gauges.

Fig 1. (a) Missouri River Basin and selected large dams with co-located USGS streamflow and nearest GHCN rainfall gages stations; (b) Comparison of dam heights where the Dam No. is sorted according to the decreasing altitude; and (c) Purposes of the dams (e.g. flood control, hydroelectric, etc.).

There is considerable geographic variation in the hydroclimatic processes in the basin, ranging from extensively snowmelt-driven topographically organized systems in the upper Northwest corner (high-altitude dams) to spring-summer precipitation and snowmelt processes (low-altitude dams) in other parts. The geology ranges from the Rocky Mountains to the highly incised badlands of South Dakota, the sand hills of Nebraska, and flatlands of Iowa and Kansas. The MRB is part of a larger mid-latitude region that is projected to experience warmer and wetter cool season conditions (the source of greatest runoff) by the end of the 21st century, relative to the last decades of the 20th century [IPCC, 2007]. The character of the streamflow and floods can vary substantially over the basin, and an assessment of the spatial and temporal coincidence of the floods is of interest, especially in the context of adaptive water systems management. By addressing these questions, we aim to understand better, the range of variability in floods in the MRB, the forcing mechanisms of that variability, and the processes that alter hydroclimatic relationships at the large watershed scale.

2.1. Streamflow and reservoirs data

The MRB water management division operated by the Northwest division of the United States Army Corp. of Engineers (USACE) hosts a suite of information regarding the major reservoirs in the basin. We carefully identified thirteen of the large dams in MRB along the main stem and in the headwaters that are directly relevant to operational decisions. It has been shown in the past that the annual peak floods and annual mean flow for various recurrence intervals show scaling relationships with drainage area of the catchment [Thomas and Benson, 1970, Lima and Lall, 2009]. By choosing large dams, we are implicitly trying to understand how the flood peaks in large drainage area catchments relate to the corresponding flood durations and volumes. We have also identified their co-located stream gages. The specifics of the co-located USGS stream gauges on the major tributaries corresponding to inflow into each reservoir are also provided in Table 1.

Table 1. Specifications and geographical locations of selected dams, USGS streamflow and GHCN rainfall stations.

No.	Dam			USGS Streamflow Station			GHCN Rainfall Station		
ID	Name	Latitude (deg.)	Longitude (deg.)	ID	Latitude (deg.)	Longitude (deg.)	ID	Latitude (deg.)	Longitude (deg.)
1	Hebgen	44.865	-111.347	06038500	44.867	-111.338	USC00244038	44.867	-111.339
2	Toston	46.120	-111.408	06054500	46.135	-111.420	USC00248324	46.331	-111.538
3	Holter	46.992	-112.006	06066500	46.984	-112.011	USC00244241	46.991	-112.012
4	Morony	47.582	-111.057	06078200	47.435	-111.388	USC00248430	47.547	-111.326
5	Tiber	48.322	-111.098	06101500	48.301	-111.081	USC00248233	48.310	-111.088
6	Fort Peck	48.003	-106.416	06132000	48.036	-106.356	USC00243176	48.012	-106.412
7	Boysen	43.417	-108.178	06259000	43.417	-108.178	USC00481000	43.405	-108.163
8	Yellowtail	45.307	-107.958	06287000	45.317	-107.919	USC00249240	45.313	-107.938
9	Oahe	44.452	-100.399	06441500	44.318	-100.384	USC00393217	44.657	-99.672
10	Cottonwood Creek	46.298	-98.268	06470500	46.355	-98.305	USC00325479	46.813	-99.509
11	Glendo	42.483	-104.950	06652800	42.457	-104.948	USC00053005	42.576	-105.086
12	Lake Babcock-North Columbus	41.467	-97.367	06774000	41.368	-97.495	USC00251240	41.333	-97.565

13	Smithville	39.399	-94.555	06821150	39.388	-94.579	USC00237963	40.247	-93.716
----	------------	--------	---------	----------	--------	---------	-------------	--------	---------

The purpose and height of dams vary across the basin (Figure 1(b, c)). We ensured that all the stations have a common data record from 1966 to 2014. Relevant information on the name, reservoir location, date of operation, height, storage and surface capacity, maximum and normal capacity and drainage area for them are available from the National Inventory of Dams (NID) (<http://nid.usace.army.mil/>) [NID, 2005] and National Water Information System (NWIS) of U.S. Geological Survey (USGS) (<http://waterdata.usgs.gov/nwis/>). Table 2 presents this information for the thirteen selected reservoirs.

Table 2. Specifications of selected large dams located in the Missouri River Basin main stem.

No.	Dam Name	State	Reservoir	Height (ft)	Storage capacity (acre.ft)	Maximum Capacity (acre.ft)	Normal Capacity (acre.ft)	Surface Capacity (acre.ft)	Drainage Area (mi ²)	Year
1	Hebgen	MT	Madison River	88	325000	525000	273000	384800	904	1915
2	Toston	MT	Toston Reservoir	56	3000	32362	3000	23600	14641	1940
3	Holter	MT	Holter Lake	124	243000	265000	245000	265000	16924	1918
4	Morony	MT	Missouri River	59	3000	13000	7800	13000	20605	1930
5	Tiber	MT	Tiber Reservoir	211	6081	1555898	967320	1337000	4944	1956
6	Fort Peck	MT	Fort Peck lake	250	18463000	18910000	15200000	18910000	56487	1937
7	Boysen	WY	Boysen Reservoir	220	892226	1473000	802000	819800	7701	1952
8	Yellowtail	MT	Bighorn Lake	525	958	1375000	873000	1375000	19672	1966
9	Oahe	SD	Lake Oahe	245	23500000	23300000	18900000	23300000	3147	1966
10	Cottonwood Creek	ND	Cottonwood Creek	54	11400	11400	7540	4160	4390	1922
11	Glendo	WY	North Platte River	170	1170505	1124000	795200	798400	15548	1958
12	Lake Babcock-North Columbus	NE	Loup Canal	32	20000	20000	16000	5270	59300	1937
13	Smithville	MO	Little Platte River	101	246500	246500	144600	144600	213	1965

2.2. GHCN rainfall and NCEP/NCAR reanalysis data

The rainfall data is obtained from the Global Historical Climatology Network (GHCN) [Menne *et al.*, 2012; NOAA-NCDC, 2015] processed in National Climatic Data Center (NCDC) of National Oceanic and Atmospheric Administration (NOAA) (<https://www.ncdc.noaa.gov/>). The GHCN rainfall gauges are also co-located or geographically close to the USGS streamflow stations and the selected reservoirs. Data on the atmospheric circulation variables that capture climate forcing driving the regional hydroclimatology are obtained from NOAA's Climate Diagnostics Center (<http://www.cdc.noaa.gov/>).

We used the anomalies of Surface Air Temperature (SAT), Precipitation Rate (PR), Precipitable Water Content (PWC), Wind Vectors (WV), Sea Level Pressure (SLP) and 500mb Geopotential Height (GPH) available at 2.5° by 2.5° resolution from NCEP-NCAR Reanalysis project [Kalnay *et al.*, 1996; Kistler *et al.*, 2001]. The reanalysis data assimilation system includes the NCEP global spectral model with 28 sigma vertical levels. There are over 80 different variables

including precipitation, temperature, geopotential height, relative humidity, meridional and zonal wind components at a 2.5^0 by 2.5^0 spatial resolution. The Type A variables (upper air temperature, rotational wind, and geopotential height) within the Reanalysis dataset are most reliable products [see more details in *Kalnay et al.*, 1996].

3. Statistical Inference for Climate-Informed Flood Risk

Figure 2 presents the conceptual framework for the climate-informed flood risk assessment. We follow an inverse modeling approach to explicitly relate the likelihood of the floods on causal links of regional climatological and atmospheric processes. It includes event selection, relating it to river basin's physical characteristics such as topography, identifying the spatial and temporal clustering of flood attributes, detecting modes of the low-frequency variability in the data, relating floods to antecedent rainfall and flow conditions and synoptic circulation patterns. The essential parts of the methodology are shown in Figure 2(b).

Fig 2. (a) The climate-informed framework proposed in this study for flood mechanisms, risk and characteristics assessment; (b) The step-by-step tasks involve spatiotemporally analyzing the flood attributes (peak, volume, duration and timing) and relating them to precursor rainfall characteristics, and climate and atmospheric variables.

Each component is elaborated as follows:

3.1. Determining the flood attributes: Duration (D), Timing (T), Annual Peak (P) and exceedance Volume (V)

The Annual Maximum Flow (AMF), or the annual peak (P) is first identified for each water year (October 1 – September 30) from the daily streamflow data for each station. The day of the year corresponding to the annual peak is recorded as the peak flow timing (T). The total number of days (within a window of $k = \pm 30$ days around the peak flow timing) when the daily streamflow exceeds a chosen threshold Q^* is computed as the flood duration (D) each year. The cumulative flow during flood duration days is calculated as the flood volume (V) per year. Hence, the flood attributes are represented as follows:

$$P_{t,j} = \max(Q_{t,j}^i)$$

i=days in the water year (October 1 - September 30);

t=1966: 2014;

j=1:13 streamflow stations

(1)

$$T_{t,j} = i \text{ when } P_{t,j} = Q_{t,j}^i$$
(2)

$$D_{t,j} = \sum_{i=T_{t,j}-k}^{T_{t,j}+k} \delta_{t,j}^i ; k = \pm 30 \text{ days}$$
(3)

$$V_{t,j} = \sum_{i=T_{t,j}-k}^{T_{t,j}+k} \delta_{t,j}^i * Q_{t,j}^i$$
(4)

$$\text{where; } \delta_{t,j}^i = \begin{cases} 1 & \text{if } Q_{t,j}^i > Q_j^* \\ 0 & \text{if } Q_{t,j}^i < Q_j^* \end{cases}$$
(5)

We choose the 90th percentile, Q_{90} , of the daily streamflow as the threshold (Q^*). This threshold based on the daily streamflow, approximately corresponds to a return period between 1 to 2 years for the selected stations. *Dalyrymle*, [1960], *Waylen and Woo*, [1983] and *Irvine and Waylen*, [1986] recommended the usage of an average return period of 1.15 years or between 1.2 to 2 years for threshold exceedance problems. *Lang et al.* [1999] also suggested various tests for selecting the threshold, mainly to choose events that are independent and identical. In our study, we use the threshold only to choose the flood days around the peak flow each year; hence across the years the events are assumed to be independent and identically distributed.

Figure 3 shows a simple schematic for computing **P**, **T**, **D** and **V** based on discharge time series. It also shows preceding rainfall events (rainfall duration and intensity) and the nearest rainfall wet spell within the same window. By computing the flood attributes over a radius of k days centered on the peak (i.e. window=[$T-k$, $T+k$]), we mimic the occurrence of floods that could relate to saturated soil conditions produced during another event occurring a short time earlier. In other words, the manifestation of recurrent rainfall events as floods can well be captured. We tested for the sensitivity of flood duration and volume to the choice of window k , and found that 97% of the events have flood duration less than 60 days. In addition to these flood attributes, the initial flow fraction is calculated from the flow 30 days in advance of the flood peak.

$$Q_{f,t,j} = \frac{Q_{t,j}^{T-k}}{Q_j^*} \quad (6)$$

$Q_{f,t,j}$ is the initial flow fraction representing the antecedent flow conditions (f stands for fraction), $Q_{t,j}^{T-k}$ denotes the discharge 30 days before the occurrence of the flood peak (k denotes the number of days preceding the peak which takes place at time T). Q_j^* is the threshold Q_{90} as defined in Equation 5. $Q_{f,t,j}$ greater than 1 indicates that the initial flow at the beginning of the flood is greater than the threshold – an incipient flood condition. $Q_{f,t,j}$ close to 0 indicates that the initial flow at the beginning of the flood is much less than the threshold – an empty river condition. The time series (1966 – 2014) of the flood attributes and rainfall components are thus computed for all the selected streamflow stations and rainfall gages in the MRB.

Fig 3. Schematic illustration of flood attributes derived from daily streamflow; P (peak) is the annual maximum flow each year, T (timing) is the time of occurrence of flood each year, V and D are the volume and duration of the flood around the peak based on the 90th percentile of the daily flow.

The Missouri River stretches on an extensive territory with significantly varying topographic features along the basin. Among the 13 reservoirs we selected, the topography ranges from an elevation of 6448.47 ft for the Hebgen Dam (Dam # 1 with the highest altitude) to an elevation of 778.38 ft for the Smithville Dam (Dam # 13 with the lowest altitude). Figure 4(a) presents the boxplot of the flood duration time series for each dam arranged in descending order of elevation. A perusal of the boxplots shows that there is a clear separation of flood duration for the high and low elevation dams. The scatterplot of dam elevation and median flood duration (Figure 4(b)) clearly shows two distinct groups. Based on this, we separated the low-altitude dams from the high-altitude dams using a threshold elevation of 2500 ft. The average number of flood days for

the low altitude dams (elevation < 2500 ft) is approximately 15 days (with average of median flood duration around 9 days). The Fort Peck Dam, Lake North-Columbus Dam, Oahe Dam, Cottonwood Creek Dam and the Smithville Dam fall in this category. The average number of flood days for the high altitude dams (elevation > 2500 ft) is approximately 25 days (with average of median flood duration around 27 days). The Hebgen Dam, Boysen Dam, Glendo Dam, Toston Dam, Holter Dam, Morony Dam, Yellowtail Dam, and the Tiber Dam fall in this category (see Figure 4(b)).

Fig 4. (a) Spatial classification of high and low-altitude dams based on the altitudes and flood duration, (the boxplots for each dam represent the 25th, 50th and 75th percentiles in the middle with whiskers extending to 1st and 99th percentile of the flood duration data) (b) Geographic locations of low-altitude dams (blue) and high-altitude dams (red) along the MRB; (c) Kernel density estimates of flood duration and timing (e.g. DOY of 1 indicates January 1) in high and low-altitude dams.

Figure 4(c) shows the comparison of the probability distribution (using kernel density estimation [Bowman and Azzalini, 1997]) of the flood duration and flood timing (day of the year of annual maximum flow) for the low altitude dams and high altitude dams. While the low altitude dams have a heavy-tailed skewed distribution for flood duration, the durations for high-altitude dams are more uniformly distributed. The timing of the low altitude dams indicates a dominance of March–April–May spring floods. The timing of the floods for high-altitude dams, however, is sharply concentrated around the summer (June–July) indicating snowmelt-driven floods. Persistence of seasonal snowmelt (from high altitude accumulated snowpack) during late spring and early summer contributes to the inflows of high-altitude dams leading to long duration floods in summer. Since most of the long duration floods in the high-altitude dams are predominantly snowmelt-driven, from here on, we only focus on understanding the spatiotemporal properties and the driving climate and atmospheric processes for floods in the low-altitude dams. We consider floods of varying durations, i.e. from a single day event through long duration floods over 30 days.

3.2. Hierarchical Clustering Analysis (HCA) to identify temporal clustering in flood duration for low-altitude dams

We employ the hierarchical clustering analysis (HCA) on the time series of flood duration (**D**) to group the features based on the most similarity (least dissimilarity) [Rokach and Maimon, 2005; Hartigan, 1975]. HCA involves the following overall steps:

- Find the similarity or dissimilarity between every pair of objects in the data set;
- Group the objects into a binary called hierarchical cluster tree;
- Determine where to cut the hierarchical tree into clusters.

HCA constructs a hierarchy of sets of groups formed by merging one pair from the collection of previously defined groups. The HCA begins by placing each value (of the Flood duration time series in this case) into separate clusters. Next, the distance between the entire possible combinations of two rows is computed by considering a reasonable distance mode. For instance, the centroid method (UPGMC: Unweighted Pair-Group Method using Centroid) will weight each component equally in the candidate cluster regardless of its structural subdivision [Rokach and Maimon, 2005]. Then, the two most similar clusters are grouped together and would form a new larger cluster. The number of clusters is thereby reduced by one in each calculation step in which the distance between the new cluster and all those remaining clusters is recalculated in subsequent steps using that predetermined distance method [Lundberg, 2005]. Eventually, all rows (here the flood duration) are grouped into one large cluster. This procedure is implemented for all sets of values separately to cluster similar values within each attribute. Finally, the hierarchical clustering output will be a range of indexed values in specific clusters (such as Cluster 1, Cluster 2) associated to a dendrogram that is a branching diagram indicating the relationships of similarity among a group of values.

Figure 5(a) presents the output of HCA on the yearly flood duration values for the five low-altitude dams from 1966 to 2014. One objective of this analysis is to choose the level of aggregation in the dendrogram at which to stop further merging. A standard practice is to find that level of clustering that maximizes similarity within clusters and minimizes similarity between clusters. *Jolliffe et al.* [1986] and *Fovell and Fovell* [1993] provide discussion on various objective stopping criteria for HCA.

Fig 5. (a) Dendrogram of hierarchical clustering applied to flood duration of low-altitude dams; (b) Illustration of Cluster 1 (short flood duration; C1: light) and Cluster 2 (long flood duration; C2: dark) for annual duration time series of low-altitude dams from 1966 to 2014; (c) Boxplot of aforementioned C1 and C2 clusters (groups); there is a clear discrimination between C1 and C2 in terms of flood duration.

For further details on the HCA, readers are referred to *Wilks* [1999]. We choose two clusters based on the summary statistics of each group. The best number of clusters for a given problem is not obvious and requires a subjective choice that depends on the goals of the analysis. Figure 5(b) presents the temporal manifestation of these clusters for the flood durations of five low-altitude dams. The grouping is prominent during 1993–1998 and again during 2009–2012. We can also identify these groups during the 1970s and the mid-1980s. Finally, Figure 5(c) presents the boxplots of the durations for each cluster and the corresponding timing of the flood. There is a clear separation in the duration of two clusters with Cluster 1 having durations less than 30 days and Cluster 2 having durations greater than 30 days. However, the distribution of the timing of floods indicates that there is no difference in the timing of occurrence of these short (< 30 days) and long (> 30 days) duration floods. This result provides a first order understanding that while the time of occurrence of the floods is the same for both the groups, depending on the particular climate mechanisms and the variability in the atmospheric processes, the duration of the floods vary as a result of the frequency of manifested rainfall.

3.3. Assessing the periodicity of long duration floods using wavelet transform

The wavelet transform can be applied to a time series to obtain an orthogonal decomposition of the original signal in the time and frequency domain [*Daubechies*, 1990; *Foufoula-Georgiou and Kumar*, 1995]. It enables the identification of dominant frequency components as well as their temporal variation. The continuous wavelet transform of a time series $x(t)$ is defined by *Chui* [1992] as:

$$W(s, t') = |s|^{-\frac{1}{2}} \int_{-\infty}^{\infty} x(t) \varphi^* \left(\frac{t-t'}{s} \right) dt \quad (7)$$

In equation (7), $W(s, t')$ is defined as the wavelet spectrum, $\varphi(t)$ is a wavelet function, $(*)$ is the complex conjugate, t' is the localized time index, $s \neq 0$ is the scale parameter. We can localize the wavelet function at $t = t'$ in order to compute the coefficients $W(s, t')$ and explore the behavior of $x(t)$ at $t = t'$. We apply the wavelet transform on the time series of the spatially averaged (of five

low-altitude dams) flood duration. The Morlet wavelet defined as $\varphi(t) = \pi^{-\frac{1}{4}} e^{i\omega_0 t} e^{-\frac{t^2}{2}}$ [*Farge*, 1992], where ω_0 is a frequency employed here. Figure 6 shows the wavelet decomposition of this time series based on *Torrence and Compo* [1998] wavelet tool. Figure 6(a) and (b) presents the raw time series and the time variation of wavelet power versus the scale, respectively. Moreover, Figure 6(c) represents the global wavelet power, i.e. the time integrated variance of energy coefficients at every scale. A red-noise significance level for the global wavelet power is also shown in Figure 6(c). An autoregressive (AR(1)) model is fit to $x(t)$, and then its Fourier spectrum and associated one-sided 95% confidence limits are computed as a function of frequency. Global wavelet power spectrums that are higher than the red noise significance level are deemed to be statistically significant. For the flood duration time series, we note that the 8 to the 16-year band has a global wavelet power level higher than the red noise significance level, thus providing evidence for decadal to bi-decadal variability in long duration floods.

Fig 6. The wavelet spectrum of averaged flood duration values for low-altitude dams with the average global wavelet spectrum; (a) Annual averaged flood duration variation of all low-altitude dams; (b) Wavelet power spectrum of flood duration time-series in panel a; (c) Global averaged wavelet power.

Such decadal scale variability often relates to large-scale climate phenomena such as the Pacific Decadal Oscillation (PDO), the Atlantic Multidecadal Oscillation (AMO) and the North Atlantic Oscillation (NAO) through their influence on atmospheric pressure disturbances and local wind patterns [Maul and Hanson, 1991; Wakelin *et al.*, 2003; Bindoff *et al.*, 2007; Bingham and Hughes, 2009]. Trenberth *et al.* [2009] and Higgins *et al.* [1997] have shown that both Atlantic and Pacific Ocean conditions can influence the hydroclimatic variability of the eastern portion of the MRB [Yu and Zwiers, 2007; Yu *et al.*, 2007; Villarini *et al.*, 2011; Villarini *et al.*, 2013; Farneti *et al.*, 2013; Smith *et al.*, 2013; Villarini *et al.*, 2014; Nayak *et al.*, 2014; Wang *et al.*, 2014; Newman *et al.*, 2016].

3.4. Identifying the atmospheric circulation patterns for long duration floods

In this section, we investigate the meteorological context and the conditions that lead to long duration floods (Cluster 2) in MRB. Long duration floods, especially over large river basins are typically associated with persistent rainfall and high antecedent moisture conditions that are ultimately linked to oceanic moisture sources [Hirschboeck, 1991].

Fig 7. The pattern of atmospheric processes (plotted as anomalies) concurrent with flood events. The Precipitable Water Content (PWC) anomalies, Surface Air Temperature (SAT) anomalies, and Precipitation Rate (PR) anomalies are shown along with the top 75th percentile of wind vectors at 500-mb pressure column for flood events in Cluster 1 (short duration flood events) and Cluster 2 (long duration flood events).

Hence, we examine the large-scale weather and the atmospheric flow anomalies associated with organized moisture transport.

For each of the flood duration clusters (C1 and C2), we identify the time of occurrence of the flood as the day of the year when annual maximum occurs. There are 194 events/dates under Cluster 1 (floods with duration < 30 days) and 51 events/dates under Cluster 2 (floods with duration > 30 days). Among these events, floods occur at two or more stations at least 56% of the times. The typical distribution of the flood dates for both clusters reveals a dominance of March–April–May spring floods (Figure 5(c)). To analyze the meteorological patterns of Cluster 1 and Cluster 2, we use the daily averaged NCEP-NCAR Reanalysis V2 data as the primary source of atmospheric data.

Figure 7 presents the composites of precipitation rate, surface air temperature anomalies, and precipitable water content anomalies averaged over the 194 events of Cluster 1 and 51 events of Cluster 2. Similarly, Figure 8 presents the composites of the sea level pressure and geopotential height anomalies averaged over the events. We can see from Figure 7 that there is a strong temperature anomaly for Cluster 2 (third horizontal panel from bottom), indicating frontal boundary separation of cold air and warm air along the Missouri River that leads to upliftment, adiabatic cooling, cumuliform cloud formation and intense precipitation ahead of the front. It reveals a common cold front based mid-latitude cyclonic phenomenon with intense, short-lived precipitation events. These surface temperature gradients are the primary drivers of moisture transport from tropics to higher latitudes [Karamperidou *et al.*, 2012; Jain *et al.*, 1999]. To show the spatial extent of the precipitation patterns during the flood events, composite of the precipitation rate (second horizontal panel from bottom) and precipitable water content anomaly (fourth horizontal panel from the bottom) is also plotted in Figure 7. Strong positive anomalies exist over much of the MRB while negative anomalies exist southwest of the basin.

Figure 8 shows the composites of sea level pressure anomalies and the 500mb geopotential height anomalies for Cluster 1 and Cluster 2. We see a strong negative anomaly (low pressure) for sea level pressure over the MRB and a corresponding upper-level divergence as a result of upper atmospheric ridge and trough manifestation. A dipole pattern of a significant positive geopotential height anomaly to the east of the basin together with a weak low anomaly to the west is revealed from the composites. The top 75th percentile wind vectors are also plotted to show the convergence and divergence of the flow aloft along the favored position on Rossby wave creating low pressure at the surface. Divergence in the upper atmosphere, caused by decreasing vorticity provides a lifting mechanism for the column of air. This upper-level divergence maintains the surface low-pressure systems resulting in anomalous precipitation condition. The anomalous convergence above the MRB is associated with a persistent, anomalous circulation feature accompanied by strong upward vertical motion over the basin indicating that long duration floods in a large basin for a season are not due to a collection of random unrelated events [Hirschboeck, 1991]. Nakamura *et al.* [2013], in their work on floods in the Ohio River identified similar association to anomalous atmospheric circulations. Similarly, Lu *et al.* [2013], Wernli, [1997] and Bao *et al.* [2006] have previously investigated atmospheric circulations and associated extreme flood events.

Fig 8. Geopotential Height (GPH) anomalies with top 75th percentile wind vectors and Sea Level Pressure (SLP) anomalies at 500-mb pressure level for flood events in Cluster 1 (short duration flood events) and Cluster 2 (long duration flood events).

3.5 Dependence on Concurrent Rainfall Events (Rainfall Duration and Intensity), Antecedent Flow Conditions, and Precursor Moisture Delivery Tracks

The concurrent rainfall events (in the flood window of ± 30 days around the peak) for Cluster 1 and Cluster 2 are considered to evaluate precisely, the impacts of rainfall duration and intensity

on flood duration and flood volume. The initial flow fraction (Q_f) of each flood event is used in conjunction to determine the effectiveness of rainfall duration or rainfall intensity (or a combination of them) on flood duration and volume variability in clusters C1 (short duration floods) and C2 (long duration floods). Figure 9 illustrates the variability of flood duration (D), flood volume (V) and initial flow fraction (Q_f) with respect to the rainfall duration and intensity. The points indicated by an open triangle (circle) correspond to Cluster 1 (Cluster 2). In Figure 9(a), 9(b) and 9(c), the x-axis represents rainfall intensity, the y-axis represents the rainfall duration and the intensity of the color of the points represents the flood duration, flood volume and initial flow fraction, respectively. Two cases are presented here to emphasize the importance of antecedent flow conditions.

Case 1: Flood events with similar rainfall duration but different rainfall intensities

Notice Case 1 in Figure 9 that shows two flood events with rainfall duration of 15 days and rainfall intensities of 3.5 mm/day and 12.37 mm/day. The event with low-intensity rainfall corresponds to longer flood duration as opposed to the event with high-intensity rainfall for the same rainfall duration. The low intensity rainfall event has flood duration of 51 days (indicated by circle) and the high intensity rainfall event has flood duration of 12 days (indicated by a triangle). We can find a similar pattern in the flood volume. An initial hypothesis is that the floods with high rainfall intensity for the same number of days should manifest as larger flood volume events. However, notice in Figure 9(c) that the initial flow fraction (Q_f) is close to 0 ($Q_f = 0.053$) for the high-intensity rainfall event and greater than 1 ($Q_f = 1.36$) for the low-intensity rainfall event. When the river is in an imminent flood condition (high flow fractions), low-intensity rainfall for several days can cause a flood with larger volume and duration. On the other hand, high-intensity rainfall for several days can result in low flood volume and duration if the river is in dry conditions (low flow fractions).

Case 2: Flood events with similar rainfall intensity but different rainfall durations

Notice Case 2 in Figure 9 that shows two flood events with a rainfall intensity of 13.8 mm/day and different rainfall durations of 13 days and 25 days. The event with low rainfall duration corresponds to high flood duration and volume, while the event with high-rainfall duration corresponds to low flood volume and duration. The low rainfall duration event has flood duration of 53 days (indicated by circle) and the high rainfall duration event has flood duration of 14 days (indicated by a triangle). Similar to Case 1, the initial flow fractions for these events are determining the flood volume and duration. While the river is in imminent flood conditions, high-intensity rainfall for a few days leads to high flood volumes and duration, while the high-intensity rainfall for several days is still not sufficient to cause a flood when the river is in dry conditions. In other words, although under an initial hypothesis that large rainfall duration will influence the flood duration, the initial flow fraction is determining how much more/less rainfall duration (assuming a constant rainfall intensity) will be effective on flood duration and volume.

Figure 9(d) presents the relationship between initial flow fraction and flood duration for all the events. We observe that most of the long duration floods have an initial flow fraction close to or greater than 1 indicating that the antecedent flow conditions play a major role in explaining the long duration floods.

(a)

(b)

(c)

(d)

Fig 9. Variation of flood duration, flood volume and initial flow fraction with respect to rainfall duration and rainfall intensity for low-altitude dams (Dam#12, 6, 10, 9 and 13 in descending drainage area ordered in circle/triangle marker size); (a) Flood duration versus rainfall duration [days] and averaged rainfall intensity [mm]; (b) Flood volume [$\text{ft}^3 \cdot \text{s}^{-1}$] variation with respect to rainfall duration and averaged rainfall intensity; (c) Initial flow fraction [$\text{ft}^3 \cdot \text{s}^{-1} / \text{ft}^3 \cdot \text{s}^{-1}$] variabilities versus rainfall duration and rainfall intensity for Cluster1 (triangle) and Cluster 2 (circle); (d) Variation of flood duration with respect to initial flow fraction; it is clear that long duration flood events (Cluster 2) have larger initial flow fraction (this is not limited to specific drainage area size, as the latter statement is true for both small, moderate and large-size drainage area).

The flood events from C2 have much larger initial flow fraction compared to that in C1 (Figure 10(a)). This analysis demonstrates that long duration rainfall occurring proportionally with a significant initial flow fraction will ultimately generate extremely high flood volumes (see the darkest red, blue, green-colored markers in all panels of Figure 9). Similarly, it can be seen that high rainfall intensities are not necessarily contributing to large flood duration and flood volumes. In fact, a long duration rainfall event with a low rainfall intensity can cause long flood duration in the presence of a large initial flow fraction (i.e. $Q_f > 1$) or imminent flood condition. The latter statement is more valid specifically for those reservoirs whose drainage areas are relatively small; such as Oahe Dam (dam#9; smallest circle/triangle markers in Figure 9). This indicates the essence of initial flow's contribution on flood duration (and volumes), in particular for C2. For further verification, we investigated the rainfall frequency (counts) for the top 50 and bottom 50 flood events in both clusters. The boxplots of the rainfall events in each cluster (presented in Figure 10(b)) show a fatter tail for rainfall counts for the long duration cluster indicating that the long duration flood events in C2 are occurring due to persistent rainfall events closer to the peak which result in large cumulative flow volumes. Thus, this scenario can cause an extremely destructive flood which leads to dam water-level rise and spillway inundations.

Next, we analyzed the number of moisture tracks entering the basin and the amount of moisture released seven days prior to the starting of the flood window (i.e. 7 days before the antecedent flow condition), for both Cluster1 events and Cluster2 events, in order to understand the atmospheric conditions that lead to high antecedent flow. The analysis of associated moisture release follows the method provided in *Lu et al.* [2013] and *Lu and Lall*, [2016], using the Tropical Moisture Exports (TME) dataset [*Knippertz and Wernli*, 2010; *Knippertz et al.*, 2013]. The TME dataset tracks moisture transports that originated in the tropics, between 0° and 20°N , and propagated to higher latitudes up to 7 days, with a 6-hourly updates of a set of meteorological parameters of the moist air parcels. The variables used for this analysis are the longitudinal and latitudinal coordinates of the moist air parcels and the specific humidity at each update. Only significant moisture transports were retained in the TME dataset, the detailed selection criteria can be found in *Knippertz and Wernli*, 2010; *Knippertz et al.*, 2013 and *Lu et al.* [2013]. Previous studies [*Lu et al.*, 2013; *Lu and Lall*, 2016] have shown the utility of this dataset for moisture trajectory tracking and analysis of moisture release, and successfully linked the moisture release directly to extreme precipitation and associated atmospheric circulation patterns. The diagnosis of associated moisture release and moist air tracks entering this study area [$115^\circ\text{W} - 90^\circ\text{W}$, $35^\circ\text{N} - 55^\circ\text{N}$] followed similar approach as *Lu et al.* [2013] with improvement on the organization of tracks by entering dates rather than their birth dates used in *Lu et al.* [2013]. Thus the analysis of moisture release is directly linked to the rainfall that triggered floods in the region. Comparisons are done on the two clusters, i.e., C1 and C2 and presented in Table 3.

Table 3. Comparisons of moisture air tracks, moisture air releases and average rainfall intensity during 7 days before Q_f (i.e. from T-38 to T-31)

7 days before Q_f		
Cluster	C1	C2
Average No. of Moisture Air Track	130	146
Moisture Release [g/kg]	21	46
Average Rainfall Intensities [mm]	6.6	11.6

Average Q_f [$\text{ft}^3 \cdot \text{s}^{-1} / \text{ft}^3 \cdot \text{s}^{-1}$]	0.347	1.513
---	-------	-------

The results show that up to 7 days before all Q_f dates, the average moisture release over all 7 days prior to Q_f of C1 (21 g/kg) is significantly lower than that of C2 (46 g/kg), which suggests a more saturated condition of C2, i.e. higher Q_f prior to flood-triggering rainfall events. The average number of tracks entering the area up to 7 days before Q_f does not differ much between C1 (130) and C2 (146). Average rainfall intensity during these 7 days is 6.6 mm for C1 and 11.6 mm for C2.

Fig 10. (a) The boxplot of initial flow fraction for the flood events in Cluster 2 (C2) and Cluster 1 (C1); (b) Duration of 50 rainfall events as top longest (from C2) and bottom shortest (from C1) in the nearest wet spell.

In order to quantify the variance in flood duration and volume that can be explained by the initial flow fraction, rainfall duration, and rainfall intensity, we fit simple empirical log-linear models for Cluster 1 and Cluster 2 as follows:

$$F_D = \alpha_D \cdot (Q_f)^{\beta_1} \cdot (R_D)^{\beta_2} \cdot (R_I)^{\beta_3} \rightarrow \log(F_D) = \log(\alpha_D) + \beta_1 \log(Q_f) + \beta_2 \log(R_D) + \beta_3 \log(R_I) \quad (8)$$

$$F_V = \alpha_V \cdot (Q_f)^{\beta_4} \cdot (R_D)^{\beta_5} \cdot (R_I)^{\beta_6} \rightarrow \log(F_V) = \log(\alpha_V) + \beta_4 \log(Q_f) + \beta_5 \log(R_D) + \beta_6 \log(R_I) \quad (9)$$

where F_D and F_V denote the flood duration and volume, Q_f , R_D and R_I are referring to initial flow fraction, rainfall duration, and rainfall intensity, respectively. The models are fit to non-zero values (i.e. F_D , F_V , Q_f , R_D , $R_I \neq 0$) and the explanatory variables are ranked based on their relative importance [Feldman, 2005; Chevan and Sutherland, 1991; Groemping, 2006] to understand which hydroclimatic driver (Q_f , R_D , and R_I) would contribute significantly to the flood duration and flood volume's variability and the percent variance they can explain. We used the package "relaimpo" developed in R by Groemping [2006] who also discussed different metrics to assess the relative importance of the explanatory variables in the model. The averaging over ordering of explanatory variables and the proportional marginal decomposition [Feldman, 2005] methods were chosen here. The "relaimpo" package presents the importance relating to the amount of explained variance along with the bootstrap confidence intervals on the metrics. The metrics present the proportionate contribution each predictor makes to R^2 , considering both its direct effect (i.e., its correlation with the predictand) and its effect when combined with the other variables in the regression equation [Johnson and Lebreton 2004; Achen, 1982].

Table 4 presents these results. We can explain around 32% of variance in the long duration floods (Cluster 2) using Q_f , which contributes 93.33% of the R^2 . We find that the R_D and R_I do not have an influence on the flood durations. In other words, initial flow fraction is the dominant explanatory variable for long duration floods. For the flood volume, we can explain around 30% of the variance with 60.39% of that R^2 contributed from Q_f and 36.92% contributed from R_I . This indicates that R_I (rainfall intensity) will determine how much flood volume may be generated beyond existing initial flow. Furthermore, we also find that using rainfall duration, intensity and initial flow fraction, we can only explain around 2% of the variance in flood duration and 15% of variance in flood volume in Cluster 1 (short duration) indicating their lack of predictive capacity. Through this analysis, we make an attempt to identify interesting aspects of local climatological features and their organization during floods that provoke thinking towards systematic predictive flood model building.

Table 4. The contribution of Q_f , R_D and R_I to flood duration and volume variability in Cluster 2 (C2) and Cluster 1 (C1) based on relative importance analysis on log-linear regression.

Model		Relative Importance of Explanatory Variables in the log-linear Regression Model			
Cluster No.		C2		C1	
Driver	Component	Flood Duration	Flood Volume	Flood Duration	Flood Volume
	Q_f	93.33%	60.39%	-	83.75%
	R_D	-	2.67%	-	7.49%
	R_I	6.35%	36.92%	-	8.75%
	Percent	32%	30%	2%	15%

F_D : Flood duration
 Q_f : Initial flow fraction ($= Q_{T,30}/Q^{*90\%}$); $Q_{T,30}$: Discharge in 30 days before flood timing (T), $Q^{*90\%}$: 90th percentile of streamflow time-series
 R_f : Rainfall intensity
 R_D : Rainfall duration

4. Implication for Flood Frequency Analysis and Flood Control

Dam control becomes increasingly important during recurrent rain and snow events. Meteorologists have considered intensity-duration-frequency (IDF) curves for extreme rainfall in a region for different durations. Typically storm durations from 1 hour to 72 hours are considered, and the rainfall totals associated with each duration for a specified return period are estimated. Operational managers often use the IDF curves together with assumptions as to antecedent soil moisture conditions (AMC) from prior rainfall in a watershed to assess or update flood risk from an event. During these larger floods, the soil is saturated and does not have the capacity to absorb additional rainfall. Under these conditions, essentially all of the rain that falls on saturated soil, run off and becomes streamflow. One could look at the current methods used in operating dams and utilize ideas that integrate the peak flow, volume, duration, and timing of such events. This multifaceted approach could yield a more comprehensive analysis for the release and control of water within a dam. From an infrastructure design perspective, the analysis of the flood volume and duration and how the statistics of such events may change with time is of critical importance to engineers in designing safe and robust infrastructure components.

Further, traditional tools for flood risk assessment are fundamentally indexed to an instant peak flow occurrence as a particular recording measurement on a river and then inferred through dynamic modeling to the potential region that is likely to be flooded. However, as discussed earlier, the determination of property losses has to factor in, the duration, velocity, and depth of floods. The current idea of flood return period centered on the instantaneous peak flow of a stream requires review in particular, for long duration floods. As we have seen above, there may be rainfall events that occur in which a low peak is experienced, however, this low peak has an extended duration. The results will be such that the volume of the event may exceed that of a much higher peak, shorter duration event. The introduction of controlling spillway releases through a peak, volume and duration examination versus just the traditional peak flow may have more merit in addressing the control issues of dam operation.

To investigate this phenomenon, we selected two flood control dams, Oahe (Dam # 9) and Smithville (Dam # 13) dams from the low-altitude dams' category and presented their joint distribution of the flood peak, volume, and duration in Figure 11. In Figure 11(a) (horizontal panel), the color of the scatter plot indicates the flood duration for the event. The 10-year return

period flood peak is also shown for both the dams. In both the dams, there is a general homogeneous one-to-one correspondence for flood peak and volume. However, it is not necessarily true that each large flood peak value will correspond to a large flood volume and vice versa. In fact, it is not surprising to see a relatively average flood peak value corresponds to largest flood volume (e.g. see a flood event with P and V values approximately equal to 14000 and 230000 $\text{ft}^3 \cdot \text{s}^{-1}$, respectively in Oahe Dam) due to the long duration of the flood.

Flood Control Dams

Fig 11. (a) Flood volume and peak variations in the presence of flood duration (color-bar) for Oahe Dam (Dam # 9) and Smithville Dam (Dam # 13) (flood control dams from Cluster 2 –long duration flood events– amongst low-altitude dams) and the corresponding Pearson correlation; (b) The recurrence intervals of 10-year flood P and log-Pearson Type III; the vertical orange colored dashed-line indicates the corresponding peak value for the recurrence intervals of 10-year flood; top 5 largest flood volume values have been labeled in a descending order from 1 to 5.

We can see that there are such events in which the volume and duration of a flood event are high with low peaks. For instance, for the Oahe dam, the event with highest flood peak ($P=19800 \text{ ft}^3.\text{s}^{-1}$) correspond to a low flood volume and duration ($V= 76726 \text{ ft}^3.\text{s}^{-1}$ and $D= 23$ days). An event like this could indicate a possible flash flood which tapers off after a short period. Compare this to the event, in the same dam, with the peak at approximately $13300 \text{ ft}^3.\text{s}^{-1}$, a volume of almost $225034 \text{ ft}^3.\text{s}^{-1}$, and duration of 60 days. This event may be more hazardous than the event mentioned above because of the high accumulation of rainfall over an extended period. Similarly, for both the dams, there are many events which have a very high peak but low duration. These events may need to be managed differently than the events occurring at a peak of approximately 12000 to $14000 \text{ ft}^3.\text{s}^{-1}$ which have high volumes and durations. More importantly, designing the flood control dams according to the flood duration is crucial for smaller dams.

The 10-year flood peak recurrence interval (which is analogous to the annual-exceedance probability of 10 %) is highlighted and presented in Figure 10(b) for both Oahe Dam and Smithville Dam. We computed this based on Bulletin 17B (B17B) of the Interagency Advisory Committee on Water Data [Flynn *et al.*, 2006]. In Table 5, we compare the flood P , V and D values as well as the corresponding annual exceedance probabilities for top 5 flood volume events in Oahe (dam # 9) and Smithville (dam # 13) dams.

Table 5. The statistics of top 5 largest flood volumes for Oahe and Smithville flood control dams; 50, 20, 10, 4 and 2 % annual exceedance probabilities indicate 2, 5, 10, 25 and 50 years flood P recurrence interval, respectively.

Oahe Dam (Dam. No. 9)						Smithville Dam (Dam No. 13)				
Event No.	1	2	3	4	5	1	2	3	4	5
V [$\text{ft}^3.\text{s}^{-1}$]	225034	153740	129563	120292	118706	62620	56815	43596	39589	37110
P [$\text{ft}^3.\text{s}^{-1}$]	13300	11800	11900	18000	16300	2390	21100	2370	2260	2160
D [days]	60	33	56	27	43	51	15	42	28	30
Annual Exceedance Probability [%]	10	20	16	4	6	38	2	40	42	44

Legend					
	Number	Number	Number	Number	Number
	Max.	-----	-----	-----	Min.

The 10-year recurrence interval event for the Oahe Dam as defined by the flood peak corresponds to the highest flood volume. Similarly, for the Smithville Dam, Event No. 1 (see Figure 11(b)) has the largest flood volume with $62620 \text{ ft}^3.\text{s}^{-1}$ and second largest flood Peak value ($2390 \text{ ft}^3.\text{s}^{-1}$) which is featured as 51 days flood event (see Table 5). The annual exceedance probability for Event No. 1 is 38% that indicates a 2–5 years flood P recurrence interval. The 10% probability (10-year flood P recurrence interval) event has less flood volume compared to this event. This information on the recurrence interval for the joint distribution of the flood peak, volume and duration can improve the flood control management strategies.

5. Summary and Conclusions

A spatiotemporal climate-informed framework for providing insights on flood features for the large reservoirs (dams) is presented here to improve on the previous flood risk assessments and floodplain management strategies which are mostly based on analyzing the instantaneous peak flow events. It is demonstrated here that while the peak flow rate plays a critical role in flood risk assessment and reservoirs operations, the duration of the flow and cumulative peak flow volumes have to be taken into consideration for dynamic flood risk management and precaution warning system's design. The connections of the large-scale atmospheric processes to flood duration as well as the frequency of the rainfall events concurrent to these events were shown and discussed here for MRB in details. The long duration of the flood peak can induce a vast amount of exceedance cumulative flow volume that could cause a massive flood event. Synoptic scale atmospheric processes that induce these anomalous flood events were identified and discussed. It is also discussed that the initial flow fraction can dictate the long duration flood events, while rainfall duration can contribute to flood volume in addition to the effect of initial flow fraction for the long duration flood events. However, the latter understanding cannot be generalized for short duration flood. The frequency of rainfall events was significantly greater for those floods with longer durations. Similarly, the number of storm tracks and the amount of moisture-released prior to the flood triggering events are significantly more for long duration floods. Subsequently, an enormous amount of flooding water will be accumulated due to recurrent rainfall events. This should be of crucial considerations to safely operate the flood control dams during the rainfall season and for updating the floodplain management strategies. In other words, it is critical to design and operate the flood control dams (e.g. dam maximum capacity, spillway capacity, and releases during floods) based on the flood duration in addition to the flood flow peak values. A moderate flood peak event with long duration can have significant flood volumes that can ultimately endanger the hydraulic structures.

The results of this study present a statistical hydroclimatologic framework to assess the spatiotemporal properties of flood attributes, especially flood durations and their driving large-scale atmospheric processes and teleconnections, however it is still important to investigate other factors such as the basin geomorphological features or the contribution of snowmelt to peak flow rate in addition to rainfall events. Furthermore, there is a timing delay ahead of peak flow corresponded to snowmelt and rainfall mechanisms that make peak flow durations more complicated than before. It is also interesting to focus on different patterns of floods (e.g. large flood peak with small flood duration, small flood peak with long flood duration, large flood peak with long flood duration, etc.) and their associated large-scale atmospheric processes and rainfall statistics. The connection of variables as mentioned above with atmospheric characteristics and climate drivers should be assessed further in both regional and global scenarios. These are the focus of our current research.

Acknowledgements

This research is supported by:

- National Science Foundation, Paleo Perspective on Climate Change (P2C2) program – award number: 1401698

- National Science Foundation, Water Sustainability and Climate (WSC) program – award number: 1360446
- Ralph E Powe Junior Faculty Award year 2015-2016
- PSC-CUNY award 68670-00 46

We are greatly thankful to the editors and reviewers and for their constructive comments and useful suggestions to improve the manuscript.

The statements contained within the manuscript/research article are not the opinions of the funding agency or the U.S. government but reflect the authors' opinions.

Data used in this research are available from (a) The MRB water management division operated by the Northwest division of the United States Army Corp. of Engineers (USACE), (b) National Inventory of Dams (NID) (<http://nid.usace.army.mil/>), (c) National Water Information System (NWIS) of U.S. Geological Survey (USGS) (<http://waterdata.usgs.gov/nwis/>), (d) Global Historical Climatology Network (GHCN) processed in National Climatic Data Center (NCDC) of National Oceanic and Atmospheric Administration (NOAA) (<https://www.ncdc.noaa.gov/>), and (e) Data on the atmospheric circulation variables are obtained from NOAA's Climate Diagnostics Center (<http://www.cdc.noaa.gov/>). The model output data (clusters of flood durations) are available from the authors upon request.

References

- Achen, C. H. (1982), *Interpreting and Using Regression*. SAGE Publications, Series of Quantitative Applications in the Social Sciences (Book 29), Thousand Oaks, CA, USA.
- Archfield, S. A., R. M. Hirsch, A. Viglione, and G. Blöschl (2016), Fragmented patterns of flood change across the United States, *Geophys. Res. Lett.*, 43, doi:10.1002/2016GL070590.
- Bao, J. W., S. A. Michelson, P. J. Neiman, F. M. Ralph, and J. M. Wilczak (2006), Interpretation of enhanced integrated water vapor bands associated with extratropical cyclones: Their formation and connection to tropical moisture, *Mon. Weather Rev.*, 134, 1063–1080, doi:10.1175/MWR3123.1.
- Bindoff, N. L., J. Willebrand, V. Artale, A. Cazenave, J. Gregory, S. Gulev, K. Hanawa, C. Le Quéré, S. Levitus, Y. Nojiri, C. K. Shum, L. D. Talley and A. Unnikrishnan (2007), *Observations: oceanic climate change and sea level*, In: *Climate Change 2007: The Physical Science Basis, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
- Bingham, R. J., and C. W. Hughes (2009), Signature of the Atlantic meridional overturning circulation in sea level along the east coast of North America, *Geophys. Res. Lett.*, 36, L02603, doi:10.1029/2008GL036215.
- Bowman, A.W., and A. Azzalini (1997), *Applied Smoothing Techniques for Data Analysis*. 280 pages, ISBN-13: 978-0198523963, Oxford University Press Inc., New York, USA.

- Brown, D. P., and A. C. Comrie (2004), A winter precipitation ‘dipole’ in the western United States associated with multidecadal ENSO variability, *Geophys. Res. Lett.*, 31(9), L09203, doi:10.1029/2003GL018726.
- Chevan, A. and M. Sutherland (1991), Hierarchical Partitioning, *Am. Stat.*, 45, 90–96, doi: 10.2307/2684366.
- Chui, C. K. (1992), *An introduction to wavelets*, Academic Press, San Diego, CA, USA.
- Dalrymple, T. (1960), Flood frequency analyses, U.S. Geol. Surv. Water Supply Pap., 1543-A, 80 pp. <https://pubs.er.usgs.gov/publication/wsp1543A>.
- Daubechies, I. (1990), The wavelet transform time-frequency localization and signal analysis, *IEEE Trans. Inform. Theory*, 36, 961–1004, doi:10.1109/18.57199.
- de Moel, H., B. Jongman, H. Kreibich, B. Merz, E. Penning-Rowsell, and P. J. Ward (2015), Flood risk assessments at different spatial scales, *Mitig. Adapt. Strategies Glob. Chang.*, 20(6), 865–890, doi:10.1007/s11027-015-9654-z .
- Dettinger, M., D. R. Cayan, H. F. Diaz, D. M. Meko (1998), North–south precipitation patterns in western North America on interannual to decadal timescales, *J. Climate*, 11, 3095–3111, [http://dx.doi.org/10.1175/1520-0442\(1998\)011%3C3095:NSPPIW%3E2.0.CO;2](http://dx.doi.org/10.1175/1520-0442(1998)011%3C3095:NSPPIW%3E2.0.CO;2).
- Farge, M. (1992), Wavelet transforms and their applications to turbulence. *Annu. Rev. Fluid Mech.*, 24, 395–457, doi: 10.1146/annurev.fl.24.010192.002143.
- Farneti, R., F. Molteni, and F. Kucharski (2013), Pacific interdecadal variability driven by tropical–extratropical interactions, *Clim. Dyn.*, 42(11), 1–19, doi:10.1007/s00382-013-1906-6.
- Feldman, B. (2005), Relative Importance and Value. Manuscript (Version 1.1, March 19 2005), downloadable at [<http://www.prismanalytics.com/docs/RelativeImportance050319.pdf>]
- Flynn, K. M., W. H. Kirby, and P. R. Hummel (2006), User’s Manual for Program PeakFQ Annual Flood-Frequency Analysis Using Bulletin 17B Guidelines, U.S. Geological Survey, Techniques and Methods Book 4, Chapter B4, 1-42; <http://water.usgs.gov/software/PeakFQ>. [accessed on March 10, 2016].
- Foufoula-Georgiou, E., and P. Kumar (1995), *Wavelets in Geophysics Volume 4 (Wavelet Analysis and Its Applications)*, Academic Press, California, USA, 1–373.
- Fovell, R., and M. Y. Fovell (1993), Climate zones of the conterminous United States defined using cluster analysis, *J. Climate*, 6, 2103–2135, [http://dx.doi.org/10.1175/1520-0442\(1993\)006%3C2103:CZOTCU%3E2.0.CO;2](http://dx.doi.org/10.1175/1520-0442(1993)006%3C2103:CZOTCU%3E2.0.CO;2).
- Gromping, U. (2006), Relative importance for linear regression in R: the package relaimpo, *J. Stat. Softw.*, 17, 1–27, doi:10.18637/jss.v017.i01.
- Galat D. L., C. S. Berry, W. M. Gardner, J. C. Hendrickson, G. E. Mestle, G. J. Power, C. Stone, and M. R. Winston (2005), Spatiotemporal patterns and changes in Missouri River fishes, in Rinne, J. N., R. M. Hughes, B. Calamusso (Eds.), *Historical Changes in Large River Fish Assemblages of the Americas*, American Fisheries Society, Bethesda, MD, USA, 249–291.

- Haraguchi, M., and U. Lall (2015), Flood risks and impacts: A case study of Thailand's floods in 2011 and research questions for supply chain decision making, *Int. J. Disaster Risk Reduct.*, 14(3), 256–272, doi:10.1016/j.ijdrr.2014.09.005.
- Hartigan, J. A. (1975), *Clustering algorithms*, John Wiley, ISBN: 047135645X, John Wiley & Sons, Inc. New York, NY, USA.
- Higgins, R. W., Y. Yao, and X. L. Wang (1997), Influence of the North American monsoon system on the U.S. summer precipitation regime, *J. Climate*, 10, 2600–2622, [http://dx.doi.org/10.1175/1520-0442\(1997\)010%3C2600:IOTNAM%3E2.0.CO;2](http://dx.doi.org/10.1175/1520-0442(1997)010%3C2600:IOTNAM%3E2.0.CO;2).
- Hirschboeck, K. K. (1991), Climate and floods, National Water Summary 1988–89: Hydrologic events and floods and droughts, USGS Water Supply Paper 2375, R. W. Paulson et al., Eds., U.S. Dept. of the Interior, 67–88, [Available online at <http://pubs.er.usgs.gov/publication/wsp2375>]. [accessed on March 15, 2016].
- Hoerling, M., J. Eischeid, and R. Webb (2013), Understanding and Explaining Climate Extremes in the Missouri River Basin Associated with the 2011 Flooding, Climate Assessment Report, 1–34, <http://www.esrl.noaa.gov/psd/csi/factsheets/pdf/noaa-mrb-climate-assessment-report.pdf>. [accessed on March 15, 2016].
- IPCC: Intergovernmental Panel on Climate Change (2007), *Climate change 2007: The physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor, and H.L. Miller (eds), Cambridge University Press, Cambridge, UK, 1–996.
- Irvine, K. N., P. R. Waylen (1986), Partial series analysis of high flows in Canadian rivers, *Can. Water Resour. J.*, 11(20) 83–91, doi:10.4296/cwrj1102083.
- Jain, S., U. Lall, and M. E. Mann (1999), Seasonality and interannual variations of northern hemisphere temperature: equator-to-pole gradient and ocean–land contrast, *J. Clim.*, 12, 1086–1100, doi:10.1175/1520-0442(1999)012<1086:SAIVON>2.0.CO;2.
- Johnson, J. W., J. M. Lebreton (2004), History and Use of Relative Importance Indices in Organizational Research, *Organ. Res. Meth.*, 7, 238–257, doi: 10.1177/1094428104266510.
- Jolliffe, I. (1986), *Principal Component Analysis*, Springer Verlag, New York, NY, USA.
- Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. C. Mo, C. Ropelewski, J. Wang, A. Leetmaa, R. Reynolds, R. Jenne, and D. Joseph (1996), The NCEP/NCAR 40-Year Reanalysis Project, *Bull. Amer. Meteor. Soc.*, 77, 437–471, [http://dx.doi.org/10.1175/1520-0477\(1996\)077%3C0437:TNYRP%3E2.0.CO;2](http://dx.doi.org/10.1175/1520-0477(1996)077%3C0437:TNYRP%3E2.0.CO;2).
- Karamperidou, C., F. Cioffi, and U. Lall (2012), Surface temperature gradients as diagnostic indicators of midlatitude circulation dynamics, *J. Clim.*, 25, 4154–4171, doi:10.1175/JCLI-D-11-00067.1.
- Kistler, R., E. Kalnay, W. Collins, S. Saha, G. White, J. Woollen, M. Chelliah, W. Ebisuzaki, M. Kanamitsu, V. Kousky, H. van den Dool, R. Jenne, and M. Fiorino (2001), The NCEP–NCAR 50-Year Reanalysis: Monthly Means CD-ROM and Documentation, *Bull. Amer.*

- Meteor. Soc., 82, 247–268, [http://dx.doi.org/10.1175/1520-0477\(2001\)082%3C0247:TNNYRM%3E2.3.CO;2](http://dx.doi.org/10.1175/1520-0477(2001)082%3C0247:TNNYRM%3E2.3.CO;2).
- Knippertz, P., and H. Wernli (2010), A Lagrangian climatology of tropical moisture exports to the northern hemispheric extratropics, *J. Clim.*, 23, 987–1003, doi:10.1175/2009JCLI3333.1.
- Knippertz, P., H. Wernli, and G. Gläser (2013), A global climatology of tropical moisture exports, *J. Clim.*, 26, 3031–3045, doi:10.1175/JCLI-D-12-00401.1.
- Lang, M., T. B. M. J. Ouarda, and B. Bobée (1999), Towards operational guidelines for over-threshold modeling, *J. hydrol.*, 225, 103–117, doi:10.1016/S0022-1694(99)00167-5.
- Lima, C. H. R., U. Lall (2010), Spatial scaling in a changing climate: A hierarchical bayesian model for non-stationary multi-site annual maximum and monthly streamflow, *J. Hydrol.*, 383, 307–318, doi:10.1016/j.jhydrol.2009.12.045.
- Lundberg, J., Classifying Dialects Using Cluster Analysis. Master's Thesis in Computational Linguistics, Department of Linguistics, University of Gothenburg, Gothenburg, Sweden, 1–99, 2005.
- Lu, M., U. Lall, A. Schwartz, and H. Kwon (2013), Precipitation predictability associated with tropical moisture exports and circulation patterns for a major flood in France in 1995, *Water Resour. Res.*, 49(10), 6381–6392, doi:10.1002/wrcr.20512.
- Lu, M., and U. Lall (2016), Tropical Moisture Exports, Extreme Precipitation and Floods in Northeast US, *Hydrol. Earth Syst. Sci.*, 0, 1–40, doi:10.5194/hess-2016-403.
- Maul, G. A. and K. Hanson (1991), Interannual coherence between North Atlantic atmospheric surface pressure and composite southern U.S.A. sea level, *Geophys. Res. Lett.*, 18, 653–656, doi: 10.1029/91GL00141.
- Menne, M. J., I. Durre, R. S. Vose, B. E. Gleason, and T. G. Houston (2012), An overview of the Global Historical Climatology Network-Daily Database, *J. Atmos. Oceanic. Technol.*, 29, 897–910, doi:10.1175/JTECH-D-11-00103.1.
- Nakamura, J., U. Lall, Y. Kushnir, A. W. Robertson, and R. Seager (2013), Dynamical structure of extreme floods in the U.S. Midwest and the United Kingdom, *J. Hydrometeor.*, 14, 485–504, doi:10.1175/JHM-D-12-059.1.
- National Inventory of Dams (NID) (2005), U.S. Army Corps of Engineers in cooperation with FEMA's National Dam Safety Program, <http://nid.usace.army.mil/>. [accessed on July 30, 2015].
- Nayak, M. A., G. Villarini, and D. A. Lavers (2014), On the skill of numerical weather prediction models to forecast atmospheric rivers over the central United States, *Geophys. Res. Lett.*, 41, 4354–4362, doi:10.1002/2014GL060299.
- Newman, M., M. A. Alexander, T. R. Ault, K. M. Cobb, C. D. Emanuele Di Lorenzo, N. J. Mantua, A. J. Miller, S. Minobe, Hi. Nakamura, N. Schneider, D. J. Vimont, A. S. Phillips, J. D. Scott, and C. A. Smith (2016), The Pacific Decadal Oscillation, Revisited, *J. Climate*, 29(12), 4399–4427, doi:10.1175/JCLI-D-15-0508.1.

- NOAA National Climatic Data Center (NOAA-NCDC) (2015), <http://www.ncdc.noaa.gov/> [accessed on July 30, 2015].
- NWS-NOAA. Office of Climate, Water, and Weather Services 2011, The Missouri/Souris River Floods of May–August 2011, <http://www.nws.noaa.gov/os/assessments/>. [accessed on July 30, 2015].
- Promchote, P., S. Y. S. Wang, and P. G. Johnson (2015), The 2011 great flood in Thailand: Climate diagnostics and Implications from climate change, *J. Clim.*, 29, 367–379, doi:10.1175/JCLI-D-15-0310.1.
- Robertson, A. W., Y. Kushnir, U. Lall, and J. Nakamura, On the Connection Between Low-frequency Modulation of Large-scale Weather Regimes and Springtime Extreme Flooding over the Midwest of the United States, *Science and Technology Infusion Climate Bulletin*, Fort Worth, TX, USA, 150–152, 2011.
- Rokach, L., and O. Maimon (2005), Clustering methods. In: Rokach L, Maimon O, editors. *Data mining and knowledge discovery handbook*, Springer, New York, USA; 321–352, doi:10.1007/b107408.
- Smith, J. A., M. L. Baeck, G. Villarini, D. B. Wright, and W. Krajewski (2013), Extreme flood response: The June 2008 flooding in Iowa, *J. Hydrometeorol.*, 14(6), 1810–1825, doi:10.1175/JHM-D-12-0191.1.
- The World Bank (2011), The World Bank Supports Thailand’s Post- Floods Recovery Effort. Available from: <http://www.worldbank.or.th> [accessed on 06 April 2016].
- Thomas, D. M. and M. A. Benson (1970) Generalization of streamflow characteristics from drainage-basin characteristics. Tech. Rep., US Geol. Surv. Water Supply Paper 1975, 55p.
- Torrence, C., and G. P. Compo (1998), A practical guide to wavelet analysis, *Bull. Amer. Meteor. Soc.*, 79, 61–78, [http://dx.doi.org/10.1175/1520-0477\(1998\)079<0061:APGTWA>2.0.CO;2](http://dx.doi.org/10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2)
- Trenberth, K. E., J. Fasullo, J. Kiehl (2009), Earth’s global energy budget, *Bull. Amer. Meteor. Soc.*, 90(3), 311–323, doi:10.1175/2008BAMS2634.1.
- Villarini, G., J. A. Smith, M. L. Baeck, and W. F. Krajewski (2011), Examining flood frequency distributions in the Midwest U.S., *J. Amer. Water Resour. Assoc.*, 43, 447–463, doi:10.1111/j.1752-1688.2011.00540.x.
- Villarini, G., J. A. Smith, R. Vitolo, and D. Stephenson (2013), On the temporal clustering of US floods and its relationship to climate teleconnection patterns, *Int. J. Climatol.*, 33, 629–640, doi:10.1002/joc.3458.
- Villarini, G., R. Goska, J. A. Smith, G. A. Vecchi (2014), North Atlantic tropical cyclones and U.S. flooding, *Bull. Am. Meteorol. Soc.*, 95(9), 1381–1388, doi:10.1175/BAMS-D-13-00060.1.
- Wakelin, S. L., P. L. Woodworth, R. A. Flather, and J. A. Williams (2003), Sea-level dependence on the NAO over the NW European Continental Shelf, *Geophys. Res. Lett.*, 30, 1–4, doi: 10.1029/2003GL017041.

- Wang, S., J. Huang, Y. He, and Y. Guan (2014), Combined effects of the Pacific decadal oscillation and El Niño–Southern oscillation on global land dry-wet changes, *Sci. Rep.*, 4(6651), 1–8, doi:10.1038/srep06651.
- Ward, P. J., M. Kummu, and U. Lall (2016), Flood frequencies and durations and their response to El Niño Southern Oscillation: Global analysis, *J. Hydrol.*, 539, 358–378, doi:10.1016/j.jhydrol.2016.05.045.
- Waylen, P., M. K. Woo (1983), Stochastic analysis of high flows generated by mixed processes. *Can J. Civil. Eng.*, 10, 639–648, doi:10.1139/l83-092.
- Wernli, H. (1997), A lagrangian-based analysis of extratropical cyclones. II: A detailed case-study, *Q. J. R. Meteorol. Soc.*, 123, 1677–1706, doi:10.1002/qj.49712354211.
- Wilks, D. S. (1999), Interannual variability and extreme-value characteristics of several stochastic daily precipitation models, *Agric. For. Meteorol.*, 93, 153–169, doi:10.1016/S0168-1923(98)00125-7.
- Woodhouse, C. A., D. M. Meko, G. M. MacDonald, D. W. Stahle, and E. R. Cook (2010), A 1200-year perspective of 21st century drought in southwestern North America, *Proc. Natl. Acad. Sci. USA*, 107(50), 21283–21288, doi:10.1073/pnas.0911197107.
- Yu, B., A. Shabbar, and F. W. Zwiers (2007), The enhanced PNA-like climate response to Pacific interannual and decadal variability, *J. Climate*, 20, 5285–5300, doi:10.1175/2007JCLI1480.1.
- Yu, B., and F. W. Zwiers (2007), The impact of combined ENSO and PDO on the PNA climate: a 1,000-year climate modeling study, *Clim. Dyn.*, 29, 837–851, doi:10.1007/s00382-007-0267-4.
- Ziegler, A. D., H. S. Lim, C. Tantasarin, N. R. Jachowski, and R. Wasson (2012), Floods, false hope, and the future, *Hydrol. Process.*, 26(11), 1748–1750, doi:10.1002/hyp.9260.

Graphical abstract