Hydroclimate drivers and atmospheric teleconnections of long duration floods: An application to large reservoirs in the Missouri River Basin

Nasser Najibi, Naresh Devineni, Mengqian Lu

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 appear to occur random across the MRB 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 precipitation, 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., 2013). 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 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 (Slater and Villarini, 2016), 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. Hoeling 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 mitigation strategies (de Moel et al., 2015).

In this study, we attempt to develop an exploratory data analysis based inference framework 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 National Centers for Environmental Prediction /National Center for Atmospheric Research (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 also present the case for developing the flood frequency analysis in a multivariate context. 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).

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 MRB as an essential part of Mississippi River System encompasses around one-sixth of the U.S. (1,371,000 km²) including the states 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 4180 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. Fig. 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.

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 (e.g. Villarini, 2016), 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 Corps of Engineers (USACE) hosts a suite of information regarding the major reservoirs in the basin. We carefully identified 13 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 flows and annual mean flow for various recurrence intervals show scaling relationships with drainage area of the catchment (Thomas and Benson, 1970; Lima and Lall, 2010). 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 gages on the major tributaries corresponding to inflow into each reservoir are also provided in Table 1.

The purpose and height of dams vary across the basin (Fig. 1b, 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.

2.2. GHCN rainfall and NCEP/NCAR reanalysis data

The rainfall data are obtained from the Global Historical Climatology Network (GHCN) (Menne et al., 2012; NOAA-NDCI, 2015) processed in National Climatic Data Center (NCDI) 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° by 2.5° 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

Fig. 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 Fig. 2(b).

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:

\[
P_{i,j} = \max\left(Q^*_{i,j}\right)
\]

\(i = \text{days in the water year (October 1 – September 30)}\)

\(\tau = 1968 \div 2014;\)

\(j = 1 : 13 \text{ streamflow stations}\)

\(T_{i,j} = i \text{ when } R_{i,j} = Q^*_{i,j}\)

\(D_{i,j} = \sum_{i=1}^{T_{i,j}} \delta_{i,j} \text{, } k = \pm 30 \text{ days}\)

\(V_{i,j} = \frac{T_{i,j}}{i} \sum_{i=1}^{T_{i,j}} \delta_{i,j} \cdot Q^*_{i,j}\)

where:

\(\delta_{i,j} = \begin{cases} 
1 \text{ if } Q^*_{i,j} > Q^*_j \\
0 \text{ if } Q^*_{i,j} \leq Q^*_j 
\end{cases}\)

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 and 2 years for the selected stations. Dalrymple (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 and 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.

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

<table>
<thead>
<tr>
<th>No.</th>
<th>Dam</th>
<th>ID</th>
<th>Name</th>
<th>Latitude (deg.)</th>
<th>Longitude (deg.)</th>
<th>USGS streamflow station</th>
<th>GHCN rainfall station</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Hebgen</td>
<td>USC00248324</td>
<td>Madison</td>
<td>44.865</td>
<td>−111.347</td>
<td>06038500</td>
<td>44.867</td>
</tr>
<tr>
<td>2</td>
<td>Toston</td>
<td>USC00248324</td>
<td>Teton River</td>
<td>46.120</td>
<td>−111.408</td>
<td>06054500</td>
<td>46.135</td>
</tr>
<tr>
<td>3</td>
<td>Holter</td>
<td>USC00248324</td>
<td>Holter Lake</td>
<td>46.592</td>
<td>−112.086</td>
<td>06086500</td>
<td>46.984</td>
</tr>
<tr>
<td>4</td>
<td>Morony</td>
<td>USC00248324</td>
<td>Morony Lake</td>
<td>47.582</td>
<td>−111.457</td>
<td>06078200</td>
<td>47.435</td>
</tr>
<tr>
<td>5</td>
<td>Tiber</td>
<td>USC00248324</td>
<td>Tiber</td>
<td>48.322</td>
<td>−111.098</td>
<td>06101500</td>
<td>48.310</td>
</tr>
<tr>
<td>6</td>
<td>Fort Peck</td>
<td>USC00248324</td>
<td>Fort Peck</td>
<td>48.003</td>
<td>−106.416</td>
<td>06132000</td>
<td>48.036</td>
</tr>
<tr>
<td>7</td>
<td>Boysen</td>
<td>USC00248324</td>
<td>Boysen</td>
<td>43.417</td>
<td>−108.178</td>
<td>06259000</td>
<td>43.417</td>
</tr>
<tr>
<td>8</td>
<td>Yellowtail</td>
<td>USC00248324</td>
<td>Yellowtail</td>
<td>45.307</td>
<td>−107.958</td>
<td>06287000</td>
<td>45.317</td>
</tr>
<tr>
<td>9</td>
<td>Oahe</td>
<td>USC00248324</td>
<td>Oahe</td>
<td>44.452</td>
<td>−100.399</td>
<td>06441500</td>
<td>44.318</td>
</tr>
<tr>
<td>10</td>
<td>Cottonwood Creek</td>
<td>USC00248324</td>
<td>Cottonwood Creek</td>
<td>46.298</td>
<td>−98.268</td>
<td>06470500</td>
<td>46.355</td>
</tr>
<tr>
<td>11</td>
<td>Glendo</td>
<td>USC00248324</td>
<td>Glendo</td>
<td>42.483</td>
<td>−104.950</td>
<td>06652800</td>
<td>42.457</td>
</tr>
<tr>
<td>12</td>
<td>Lake Babcock-North Columbus</td>
<td>USC00248324</td>
<td>Lake Babcock-North Columbus</td>
<td>41.467</td>
<td>−97.367</td>
<td>06774000</td>
<td>41.368</td>
</tr>
<tr>
<td>13</td>
<td>Smithville</td>
<td>USC00248324</td>
<td>Smithville</td>
<td>39.399</td>
<td>−94.555</td>
<td>08821150</td>
<td>39.388</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>No.</th>
<th>Dam name</th>
<th>State</th>
<th>Reservoir</th>
<th>Height (ft)</th>
<th>Storage capacity (acre ft)</th>
<th>Maximum capacity (acre ft)</th>
<th>Normal capacity (acre ft)</th>
<th>Surface capacity (acre ft)</th>
<th>Drainage area (mi²)</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Hebgen</td>
<td>MT</td>
<td>Madison River</td>
<td>88</td>
<td>325,000</td>
<td>525,000</td>
<td>273,000</td>
<td>384,800</td>
<td>904</td>
<td>1915</td>
</tr>
<tr>
<td>2</td>
<td>Toston</td>
<td>MT</td>
<td>Teton Reservoir</td>
<td>56</td>
<td>3000</td>
<td>32,362</td>
<td>3000</td>
<td>23,600</td>
<td>14,641</td>
<td>1940</td>
</tr>
<tr>
<td>3</td>
<td>Holter</td>
<td>MT</td>
<td>Holter Lake</td>
<td>124</td>
<td>243,000</td>
<td>265,000</td>
<td>245,000</td>
<td>265,000</td>
<td>16,924</td>
<td>1918</td>
</tr>
<tr>
<td>4</td>
<td>Morony</td>
<td>MT</td>
<td>Missouri River</td>
<td>59</td>
<td>3000</td>
<td>13,000</td>
<td>7800</td>
<td>13,000</td>
<td>20,605</td>
<td>1930</td>
</tr>
<tr>
<td>5</td>
<td>Tiber</td>
<td>MT</td>
<td>Tiber Reservoir</td>
<td>211</td>
<td>6081</td>
<td>1,555,898</td>
<td>967,320</td>
<td>1,337,000</td>
<td>4944</td>
<td>1956</td>
</tr>
<tr>
<td>6</td>
<td>Fort Peck</td>
<td>MT</td>
<td>Fort Peck lake</td>
<td>250</td>
<td>18,463,000</td>
<td>18,910,000</td>
<td>15,200,000</td>
<td>18,910,000</td>
<td>56,487</td>
<td>1937</td>
</tr>
<tr>
<td>7</td>
<td>Boysen</td>
<td>WY</td>
<td>Boysen Reservoir</td>
<td>220</td>
<td>892,226</td>
<td>1,473,000</td>
<td>802,000</td>
<td>819,800</td>
<td>7701</td>
<td>1952</td>
</tr>
<tr>
<td>8</td>
<td>Yellowtail</td>
<td>MT</td>
<td>Bighorn Lake</td>
<td>525</td>
<td>958</td>
<td>1,375,000</td>
<td>873,000</td>
<td>1,375,000</td>
<td>19,672</td>
<td>1966</td>
</tr>
<tr>
<td>9</td>
<td>Oahe</td>
<td>SD</td>
<td>Lake Oahe</td>
<td>245</td>
<td>23,500,000</td>
<td>23,300,000</td>
<td>18,900,000</td>
<td>23,300,000</td>
<td>3147</td>
<td>1966</td>
</tr>
<tr>
<td>10</td>
<td>Cottonwood Creek</td>
<td>ND</td>
<td>Cottonwood Creek</td>
<td>54</td>
<td>11,400</td>
<td>11,400</td>
<td>7540</td>
<td>4160</td>
<td>4390</td>
<td>1922</td>
</tr>
<tr>
<td>11</td>
<td>Glendo</td>
<td>WY</td>
<td>Glendo</td>
<td>170</td>
<td>1,170,505</td>
<td>1,124,000</td>
<td>795,200</td>
<td>798,400</td>
<td>15,548</td>
<td>1958</td>
</tr>
<tr>
<td>12</td>
<td>Lake Babcock-North Columbus</td>
<td>NE</td>
<td>Loup Canal</td>
<td>32</td>
<td>20,000</td>
<td>20,000</td>
<td>16,000</td>
<td>5270</td>
<td>59,300</td>
<td>1937</td>
</tr>
<tr>
<td>13</td>
<td>Smithville</td>
<td>MO</td>
<td>Little Platte River</td>
<td>101</td>
<td>246,500</td>
<td>246,500</td>
<td>144,600</td>
<td>144,600</td>
<td>213</td>
<td>1965</td>
</tr>
</tbody>
</table>
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.

Fig. 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 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.

\[ Q_{f,i} = \frac{Q_{f,k}}{Q_{f,j}} \]  

\( Q_{f,i} \) is the initial flow fraction representing the antecedent flow conditions (f stands for fraction), \( Q_{f,k} \) 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). \( Q_{f,j} \) is the threshold \( Q_{90} \) as defined in Eq. (5). \( Q_{f,j} \) greater than 1 indicates that the initial flow at the beginning of the flood is greater than the threshold – an incipient flood condition. \( Q_{f,j} \) 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.

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). Fig. 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 (Fig. 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 Fig. 4(b)).

Fig. 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 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 predominately 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 (UPMC: 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.

Fig. 5(a) presents the output of HCA on the yearly flood duration values for 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.

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. Fig. 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, Fig. 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.

![Fig. 5](image-url)

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) (the boxplot for each cluster -of flood duration and timing- shows the 25th, 50th and 75th percentiles in the middle with whiskers extending to 1st and 99th percentile of the data); there is a clear distinction between C1 and C2 in terms of flood duration.

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:

\[
W(s, t') = |s|^{\frac{1}{2}} \int_{-\infty}^{\infty} x(t) \psi^{*}\left(\frac{t - t'}{s}\right) dt
\]

In Eq. (7), \( W(s, t') \) is defined as the wavelet spectrum, \( \psi(t) \) is a wavelet function, \( \psi^* \) 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 5 low-altitude dams) flood duration. The Morlet wavelet is defined as \( \psi(t) = \pi^{-\frac{1}{4}} e^{\omega_0^2 t^2} \) (Farge, 1992), where \( \omega_0 \) is a frequency employed here. Fig. 6 shows the wavelet decomposition of this time series based on Torrence and Compo (1998) wavelet tool. Fig. 6(a) and (b) present the raw time series and the time variation of wavelet power versus the scale, respectively. Moreover, Fig. 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 Fig. 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
associated with the floods. Recently, Ding and Wang, 2005; Bindoff et al., 2007; Bingham and Hughes, 2009). NAO is an atmospheric mode derived from sea level pressure that is often related to AMO in exerting decadal scale variability. 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). Recently, Mallakpour and Villarini (2016), have demonstrated that large-scale climate indices can explain the inter-annual variability in the frequency of flood events.

We investigated the joint relationships of lagged Pacific/North American (PNA) pattern and NAO with flood durations and volumes. At a 15-day lag time, we find a clear non-linear dependence (figure not shown). Moreover, there is a clear separation in the distribution of these climate/atmospheric variables between short duration and long duration floods. In fact, during the long duration floods (C2, D > 30 days), NAO and PNA are anomalous and anti-correlated, whereas the NAO and PNA anomalies are mostly under neutral conditions for the short duration floods (C1, D < 30 days).

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 always related to the slowly-moving weather systems and persistent moisture supply which ultimately will be linked to the oceanic moisture sources (Hirschboeck, 1991).

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 (Fig. 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.

Fig. 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, Fig. 8 presents the composites of the sea level pressure and geopotential height anomalies averaged over the events. We can see from Fig. 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 Fig. 7. Strong positive anomalies exist over much of the MRB while negative anomalies exist southwest of the basin.

Fig. 8 shows the composites of sea level pressure anomalies and the 500 mb 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. In summary, from Figs. 7 and 8, one can relate the long duration floods to anomalies in the geopotential height fields (e.g. PNA), which is in turn modulated by the large scale oceanic tele-connections.
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 \( Q_i \) 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). Fig. 9 illustrates the variability of flood duration (D), flood volume (V) and initial flow fraction \( Q_i \) with respect to the rainfall duration and intensity. The points indicated by a triangle (circle) correspond to Cluster 1 (Cluster 2). In Fig. 9(a), (b) and (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 Fig. 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 Fig. 9(c) that the initial flow fraction \( Q_i \) is close to 0 (\( Q_i = 0.053 \)) for the high-intensity rainfall event and greater than 1 (\( Q_i = 1.36 \)) 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 Fig. 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.

Fig. 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.

The flood events from C2 have much larger initial flow fraction compared to that in C1 (Fig. 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 Fig. 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. Qf > 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 Fig. 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 Fig. 10(b)) show a wider-distribution 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 Cluster 1 events and Cluster 2 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.
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.

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:

$$F_D = \exp^{\beta_0} \cdot (Q_f)^{\beta_1} \cdot (R_D)^{\beta_2} \cdot (R_I)^{\beta_3} \cdot \epsilon_D \rightarrow$$

$$\ln(F_D) = \alpha_D + \beta_1 \ln(Q_f) + \beta_2 \ln(R_D) + \beta_3 \ln(R_I) + \epsilon_D$$

$$F_V = \exp^{\beta_V} \cdot (Q_f)^{\beta_4} \cdot (R_D)^{\beta_5} \cdot (R_I)^{\beta_6} \cdot \epsilon_V \rightarrow$$

$$\ln(F_V) = \alpha_V + \beta_4 \ln(Q_f) + \beta_5 \ln(R_D) + \beta_6 \ln(R_I) + \epsilon_V$$

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; Groenping, 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 Groenping (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.

<table>
<thead>
<tr>
<th>Driver/Component</th>
<th>C2</th>
<th>C1</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Q_f$</td>
<td>93.33%</td>
<td>60.39%</td>
</tr>
<tr>
<td>$R_D$</td>
<td>2.67%</td>
<td>7.49%</td>
</tr>
<tr>
<td>$R_I$</td>
<td>6.35%</td>
<td>36.92%</td>
</tr>
<tr>
<td>Percent</td>
<td>32%</td>
<td>30%</td>
</tr>
</tbody>
</table>

$F_D$: Flood duration.
$Q_f$: Initial flow fraction ($= Q_{\text{initial}}/Q_{90\%}$); $Q_{90\%}$: Discharge in 30 days before flood timing (T); $Q_{\text{initial}}$: 90th percentile of streamflow time-series. $R_I$: Rainfall intensity. $R_D$: Rainfall duration.
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; the symbol size varies with respect to the flood duration values; and top 5 largest flood volume values have been labeled in a descending order from 1 to 5.

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 h to 72 h are considered, and the rainfall totals associated with each duration for a specified return period are estimated. Operational managers...
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.

<table>
<thead>
<tr>
<th>Event No.</th>
<th>Oahe Dam (Dam No. 9)</th>
<th>Smithville Dam (Dam No. 13)</th>
</tr>
</thead>
<tbody>
<tr>
<td>V [ft³/s¹]</td>
<td>250536 153740 129563 120292 118706</td>
<td>62620 56815 43596 39589 37110</td>
</tr>
<tr>
<td>P [ft³/s¹]</td>
<td>13300 11800 11900 18000 16300</td>
<td>2390 21100 2370 2260 2160</td>
</tr>
<tr>
<td>D [days]</td>
<td>60 33 56 27 43</td>
<td>51 15 42 28 30</td>
</tr>
</tbody>
</table>

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 = 19,800 ft³ s⁻¹) correspond to a low flood volume and duration (V = 76,726 ft³ 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 13,300 ft³ s⁻¹, a volume of almost 225,034 ft³ 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 12,000–14,000 ft³ 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 Fig. 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.

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 Fig. 11(b)) has the largest flood volume with 62,620 ft³ s⁻¹ and second largest flood Peak value (2390 ft³ 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 reservoir 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.

Acknowledgments

This research is supported by: National Science Foundation, Paleoclimate 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 for the second author. Year 2015–2016 PSC-CUNY award 68670-00 46

We are grateful to the editors and reviewers and for their constructive comments and useful suggestions that eventually led to an improved 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://mid.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 (CHCN) 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 also available from the authors upon request.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.adwatres.2016.12.004.

References


