Shifting Streamflow Regimes and Unusual Flood Events have an Impact on Flood Frequency Analysis for Cold Regions Watersheds

Donald H. Burn[1] **and Paul H. Whitfield**[2]
Dept. of Civil & Environmental Engineering, University of Waterloo, Waterloo, ON, Canada, N2L 3G1[1]
E-mail: donburn369@gmail.com

Center for Hydrology Coldwater Laboratory, University of Saskatchewan, Canmore, AB, Canada[2]
E-mail: paul.h.whitfield@gmail.com

ABSTRACT

Climate change is affecting flood events in complicated ways. In cold regions, the frequency of different flood drivers has shifted causing important changes in flood distributions that lead to challenges for flood frequency analysis (FFA). An assumption in FFA is that the flood series consists of independent events that are identically distributed. This assumption is unlikely to hold if there are i): changes occurring in the magnitude of flood events; ii) a mixture of flood generating processes; or iii) changes with time in the mixture of flood processes. A particular concern for FFA is when events that are large in magnitude occur outside the most common streamflow peak season. Such events can be characterized as Rogue events that are distinct from the commonly observed flood generating process in a watershed.

Significant changes in flood type fraction were found such that nival events decreased in frequency while mixed and pluvial events increased. These changes indicate a shift from nival events towards more pluvial dominated systems in other seasons. Flood frequency analysis using a combined distribution approach with the three flood types resulted in larger magnitude design flow estimates in comparison with the results from considering the data to be from a single population. The characteristics of Rogue events were explored using single dimension detection for magnitude and timing outliers, and multi-dimensional detection of magnitude/timing density outliers. The methods identify large events that are outliers in several distinct ways and are therefore considered Rogue events. The prevalence of Rogue events is explored by applying the methodology to more than 2100 hydrometric stations from Canada and the United States. The spatial and temporal distributions of these events are compared and the implications of the Rogue events for FFA are investigated.

KEYWORDS: cold regions hydrology; regime shifts; flood frequency analysis; rogue events

1 INTRODUCTION

Climate change is affecting flood events but the relationship between floods and climate change is complex. Recent research has demonstrated that the frequency of different flood drivers has shifted potentially causing changes in flood distributions (Mallakpour and Villarini 2015, Burn and Whitfield 2016). These shifts create challenges for FFA.

Changes in flood events have important ramifications for FFA, which forms the basis for the design and operation of flood protection systems and related infrastructure. An assumption in FFA is that the available flood series consists of independent events that are identically distributed (the IID assumption). The IID assumption is unlikely to hold if there are i): changes occurring in the magnitude of flood events as a function of a covariate, such as time; ii) a mixture of flood generating processes (e.g., Barth et al. 2017); or iii) changes with time in the mixture of flood processes (Burn and Whitfield 2023).

A variety of processes control the time of occurrence, duration, extent, and severity of river floods (Whitfield 2012) and the flood generating processes may change or shift over time. While no unified definition of causative mechanisms of flood events exists, Burn and Whitfield (2018) suggest viewing hydrologic regime as a nival-pluvial continuum with sites potentially moving over time from a purely nival regime towards a greater influence of pluvial events. Flood generating processes vary by region, season, and event severity. Changes in flood seasonality are most pronounced where snow accumulation and melt are the important flood generating processes (Köplin et al. 2014).

For the northern mid-latitude of North America, Burn and Whitfield (2023) showed changes in flood regime over time at 46 sites over an 80-year period. Strongly nival and pluvial watersheds show no changes, but stations with flood mixtures show an increasing influence of pluvial events. Tramblay et al. (2023) showed that changes in the flood generation mechanism, rather than trends, were occurring in the Mediterranean area based on annual peak flows from three rainfall types. In the UK, the duration and intensity of dry and wet spells increased with climate change, flash floods became less likely, but the probability of rain driven (pluvial) floods intensified (Rahmani and Fattahi 2023).

A method is presented (Burn and Whitfield, 2025) that uses circular statistics to cluster flood events and the clusters that are formed are linked to streamflow regime and climatology allowing separation of nival, mixed, and pluvial Peaks over Threshold (POT) events providing increased resolution of flood generating processes. The separation of event types allows the determination of changes in the frequency of event types over time and with temperature, precipitation, and climate indices.

Recently, emphasis has shifted to “strange” and “extreme” floods (Bertola et al. 2024). These extreme events often are not generated by the most commonly observed generating processes but rather may occur as a result of atmospheric rivers (Barth et al. 2017), hurricanes, or other large-scale processes (Whitfield and Pomeroy, 2016). The second part of this study presents a methodology to identify Rogue flood of record (FoR) events that are large in magnitude and occur outside the most common flood season or are simply different from the majority of observations. To identify Rogue R3-FoR events, the proposed method combines detection of [1] magnitude outliers, such as the flood-of-record, with [2] timing outliers, and [3] density outliers. Other “Interesting FoR” (R2-FoR) are detected when only two of the three conditions apply, specifically magnitude and timing or magnitude and density. While there is widespread familiarity with outliers in a single dimension, outliers in direction and multidimensions can be detected. These events are particularly interesting because they are outside the commonly observed generating process in a watershed, i.e., they are completely different to the ‘regular’ flood-generating mechanism, and may violate the IID assumption. An example would be a nival catchment where the annual peak is usually from snowmelt in spring/summer, but large events that occur in fall or winter are generated by an alternate process such as a hurricane or an atmospheric river. The proposed methodology is applied to more than 2100 reference and natural hydrometric stations across North America. The identified events are considered spatially and temporally to demonstrate that these are not random in space or time. This work is a step towards providing practitioners an approach that can be used to estimate appropriate design floods for cases where there are mixed flood generating processes, changes in the mixture of flood processes, or unusual flood events in the flood record.

2 METHODOLOGY

2.1 Changes in Flood Processes

All streamflow data used in the analysis were from gauging stations that are part of a Reference Hydrologic Network (RHN). Stations were selected from cold-region locations in Canada and the United States where nival or a mix of nival and pluvial flood events could be anticipated based on the location and the elevation of the gauging station. All stations were required to have an essentially complete data record for the 70-year period from 1951 to 2020. An essentially complete record was considered to entail no more than three missing years in a row, and no more than seven missing years in total where data for a year are considered to be missing if there are any days in the year with no data. A total of 202 stations were analyzed with 49 stations from Canada and 153 from the United States. POT data were used in this research, rather than annual maximum data, since POT data will generally result in more flood events than annual maximum data and have been reported to be superior to annual maximum data for flood frequency analysis (Pan et al. 2022).

The first step is to assign a flood process to each over threshold event. The approach is based on clustering of the event dates and builds on the approach of Burn and Whitfield (2023). There is a strong link between the time of the year of flood events and the flood generating process for a flood event (Blöschl et al. 2017). The date of the maximum flow for the POT events for a site are grouped into clusters using circular clustering; circular clustering is described in Whitfield (2018). The clustering was conducted using partitioning around mediods (pam), with the implementation of the pam algorithm through the R package "cluster" (Maechler et al. 2019). Five clusters were used in the subsequent analysis as a consensus preferred number of clusters based on six clustering indices. See Burn and Whitfield (2025) for further details. The POT events in five clusters were regrouped into three flood types (nival, mixed, and pluvial) using median daily streamflow and monthly climatology. Nival events are associated with the spring freshet, pluvial events arise from rainfall and mixed events result from rainfall on a snowpack, or rain on snow (ROS), that did not occur during the freshet season. Three flood types are considered herein to increase the total number of events of each flood type and decrease the uncertainty in estimates of flood quantiles. The process of identifying a flood type for the events in each cluster used polar plots (Pewsey et al. 2014) to which information regarding the flood regime for each station was added. The polar plots, and ancillary information that included median daily flows as well as monthly climatology, were used to identify flood types for events. Monthly climatological variables used were temperature, precipitation depth and snow depth for the location of the gauging station, based on the 1961 – 1990 climate normal.

Changes in the fraction of flood events of a given type were evaluated using logistic regression (Frei and Schär 2001), with a correction for overdispersion. Logistic regression is used when the independent variable in a regression relationship is binary. The variable of interest is the fraction of flood events so the independent variable for this analysis is the count of flood events of the flood type of interest (e.g., nival) and the combined count of the other two flood types (e.g., mixed and pluvial). The model estimates the probability of an event of the flood type of interest as a function of a predictor variable. For each site, logistic regression was applied considering each flood type (nival, mixed, and pluvial) as the flood type of interest. Predictor variables considered were: time (year); mean annual temperature; mean annual precipitation; and four climate indices. Further details on the implementation of logistic regression can be found in Burn and Whitfield (2025).

The final step in the methodology was an application of FFA based on a combined, or mixed, distribution approach for sites with two or three identified flood types following the approach used by Waylen and Woo (1982). The combined distribution can be defined as:

$$F_C(x) = F_N(x) \times F_M(x) \times F_P(x) \tag{1}$$

where $F_C(x)$ is the combined distribution, subscripts N, M and P refer to the distribution fit to the nival, mixed and pluvial events, respectively, and x is the over threshold event. The Generalized Pareto Distribution (GDP) was used to fit a distribution to the nival, mixed and pluvial POT data using L-Moments. If a site does not have events of a given flood type, the corresponding term in Equation (1) is not included. The results from the combined distribution approach were compared to results from fitting a distribution to the entire dataset considered as a single population. A minimum data set size of 20 over threshold events for a flood type was selected as it was thought that 20 represents a reasonable balance between the number of sites for which analysis can be conducted and the uncertainty associated with the design flood estimates.

2.2 Rogue Flood Events

Rogue flood events were identified from daily streamflow data for sites with more than 50 years of data resulting in 2123 hydrometric stations in Canada and the US. For each site, the annual maximum daily discharge and the date on which it occurred were extracted for the daily flow record. Any year with only a partial record was removed.

Methods for outlier detection were used to identify different aspects of FoR extreme events. There are many different methods to determine outliers. Recently, data mining has produced several tests that can detect outliers with other attributes than available with a single variable. The exploration of outliers sought to find methods that were available and practical to implement. To describe these, it is useful to separate into three categories: magnitude, timing, and multidimensional density (timing and magnitude). Three methods were retained to identify N-FoR, R3-FoR or R2-FoR for the analysis presented here. These methods were focused on identifying three types of FoR outliers: [1] magnitude, [2] timing, and [3] combined magnitude and timing. The combination of these three provides the information to determine if a flood-of-record is 'normal' (N-FoR) implying similar to most floods at the site, but the largest, an R2-FoR, which is similar to 'normal' but not part of the distribution of 'normal' events, or, an R3-FoR, which is a flood event that is not 'normal' but larger, with different timing, and not part of the distribution of normal events. Further details on the methods used to identify outliers can be found in Whitfield and Burn (2026).

3 RESULTS

3.1 Changes in Flood Processes

The change in the fraction of events of a given flood type with time (year) are summarized in Table 1. The results indicate that, for some categories, there are substantively more changes than would be expected to occur by chance. Noteworthy is the decrease in the fraction of nival events with time. There is also an increase in the fraction of pluvial events. The decrease in the fraction of nival events and increase in the fraction of pluvial events suggests that a shift from nival towards pluvial events is a strong climate change signal for this data set. Mixed events are experiencing fewer changes over time. Fewer changes in the fraction of mixed events likely reflects the occurrence of more mixed events based on a shift from nival events balanced by fewer mixed events as mixed events shift to a pluvial flood response. Similar results were obtained for annual temperature or annual precipitation as the explanatory variable.

Table 1. Percent stations with changes in flood processes with time using logistic regression.

Significance Level	Nival	Mixed	Pluvial
Increase 5%	1.8%	5.0%	14.9%
Increase 10%	3.0%	8.3%	20.0%
Decrease 10%	21.5%	7.5%	1.7%
Decrease 5%	16.1%	4.2%	0.6%

The results from the combined frequency analysis approach are displayed in the form of boxplots of design flow ratios that are calculated as the design flow from the combined frequency analysis approach divided by the design flow from considering the entire data set as a single population. Figure 1 presents results for sites with three flood types. Sites were only included in the analysis if there was a minimum number of events for a flood type. For thresholds of 15, 20 and 25, the numbers of sites for 3 flood types were 74, 59 and 30, as shown on Figure 1.

The boxplots in Figure 1 are arranged by increasing return period values on the horizontal axis with three boxplots for each return period showing results, from left to right, for 15, 20 and 25 as the threshold for the minimum number of events for a flood type. The data points are superimposed on the boxplots in

3 columns with the left column (black upward triangles) for sites where the combined estimate is above the 95% confidence limit for the entire dataset estimate while the middle (grey circles) and right (black downward triangles) columns indicate sites that are, respectively, within and below the 95% confidence limits.

Values greater than one on the graphs indicate that the combined estimate (EQ 1) is larger than the single dataset estimate. It can be observed that the median design flow ratio is in the range of around 1.2 to 1.3 for all results presented indicating that a larger combined estimate is the norm. Increasing the minimum record length results in a narrower spread of the results. While the median design flow ratio is often slightly lower for the 15-event minimum record length, there is no systematic pattern in median design flow ratio as a function of the minimum record length in that the values for 20 and 25 record length minimum are generally quite similar. There is a preponderance of combined estimates of design flow that are above the 95% confidence limit of the single dataset estimate (black upward triangles) with this being more prevalent for shorter return periods. There are no combined estimates that are below the 95% confidence limit for the single dataset estimate.

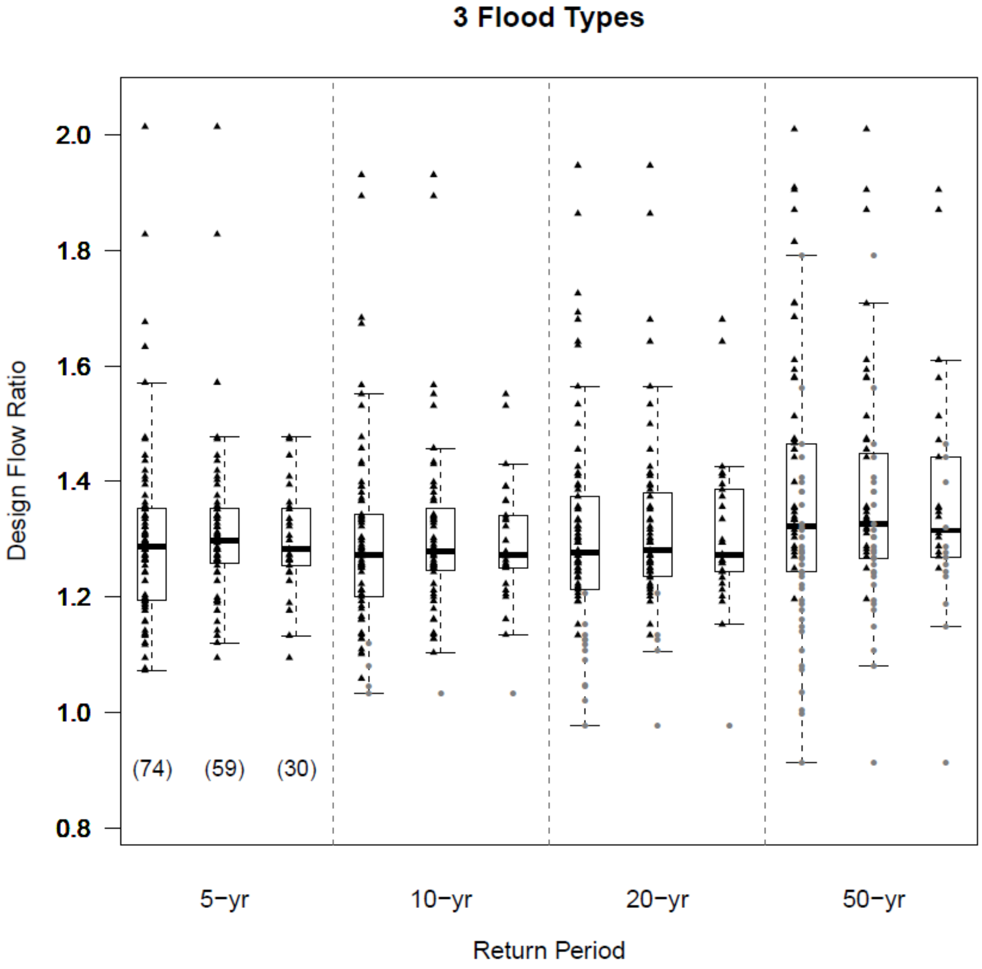

Figure 1: Boxplots of design flow ratio for sites with three flood types. The solid line inside the box indicates the median value and the box encompasses the 25th to the 75th percentile. The whiskers extend 1.5 times the interquartile range from the box. Points plotted as upwards black triangles, grey circles and downward black triangles indicate observed values that are above, within, and below the 95% confidence limits, respectively.

3.2 Rogue Flood Events

R3-FoR events, which are larger, with different timing, and not part of the distribution of normal events, are relatively rare, only about 10 % of the total number of cases (205/2123). R2-FoR events are more common as the selection is less stringent, constituting about 39 % (823) of the cases and N-FoR events represent 51 % (1095/2123).

The distribution of the flood-of-record events over time is shown in Figure 2. There are fewer FoR events in the early part of this plot because there were fewer stations in operation before 1960. There is no obvious trend in any of the N-FoR, R2-FoR, nor R3-FoR events, but a few years stick out as having more R3-FoR events such as 1937, 1954, and 2010 (Figure 2).

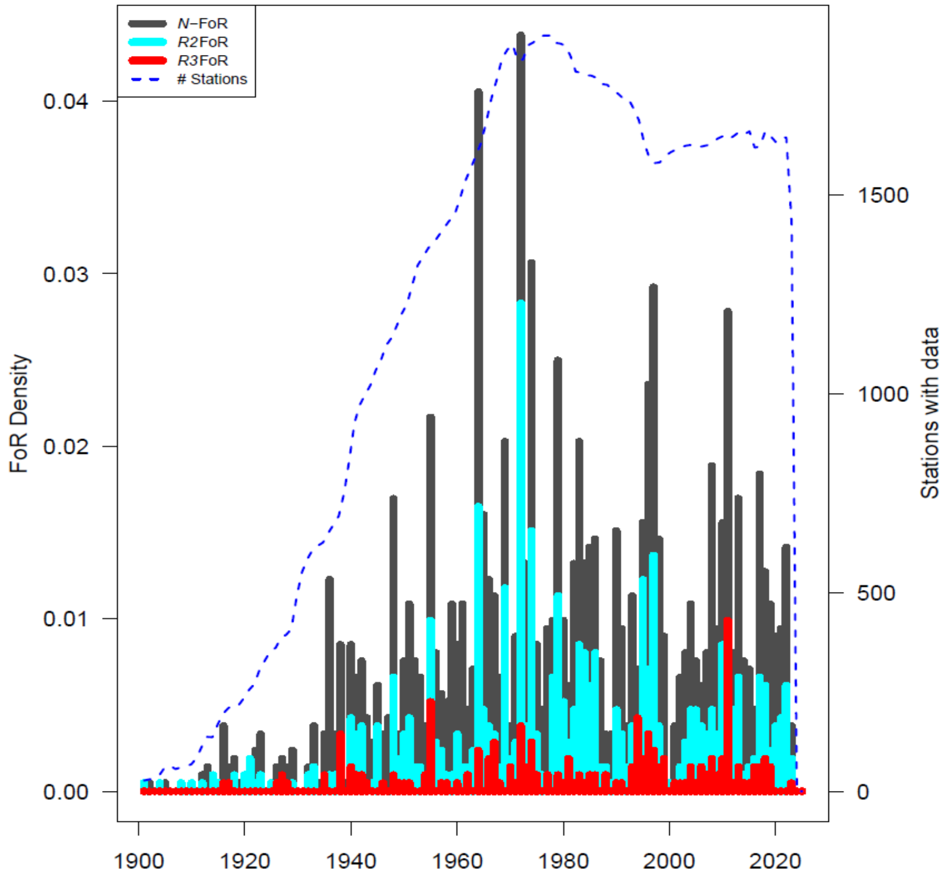

Figure 2: Histogram of flood-of-record (N-FoR), R2-FoR, R3-FoR events by year for the 2123 study stations. Also shown is the number of stations with data for each calendar year (dashed line)

4 DISCUSSION

The observed changes in the fraction of flood types have implications for FFA particularly since a comparison of design flow estimates in this work revealed that using the combined distribution approach generally resulted in higher estimates than those obtained from considering the entire data set as one population. The median increases in design flow estimates, for different return periods, were in the range of 20 to 30% (Figure 1). Large flood events are often not simply the largest of a common type of flood event in a watershed (Bertola et al. 2024). Large flood events often originate from different processes than most of the observed floods for the basin having separate and distinct flood generation mechanisms such as hurricanes, atmospheric rivers and mesoscale events (Whitfield and Pomeroy 2016). Many sites have a limited number of floods from the subpopulation that generated the flood of record, which results in large uncertainty associated with estimates of design flows calculated using the combined distribution approach. A small number of events of the flood type responsible for the flood of record, as well as other large flood events, creates challenges for water professionals who must estimate design flow values for the design of critical infrastructure. Recent studies have considered extreme floods but have not addressed the issue of timing. We approach the study of large floods in this study by explicitly addressing the nature of the flood-of-record with respect to magnitude and timing. We adopt methods that identify outliers in large floods based on magnitude, timing, and magnitude with timing using three separate outlier detection methods.

Identifying Floods-of-Record that are not the result of the dominate flood generating process in a basin has several consequences. Most importantly, such events mean that the assumptions of FFA may be violated, but, unlike low outliers, these events cannot be censored.

5 CONCLUSION

This research emphasizes the importance of considering multiple flood types in flood frequency analysis. Flood frequency analysis using a combined distribution approach with three flood types resulted in larger magnitude design flow estimates with a median increase of 20 – 30 % in comparison with the results from considering the data set to be from a single population. Flood frequency analysis also needs to consider changes in the mix of flood types by adopting a broader definition of nonstationarity that considers more than just flood magnitudes. Challenges exist for designing flood protection works, and other critical infrastructure, due to short record lengths for less frequently occurring flood types, which may include the flood of record for a site. While there are no trends over time in the occurrence of Rogue events, they are expected to increase in a warming world where runoff events are becoming more common throughout the year. This too has important implications for flood frequency analysis.

6 ACKNOWLEDGEMENTS

This research was partially supported by funding provided by the Natural Sciences and Engineering Research Council of Canada (NSERC) and Global Water Futures.

REFERENCES

Barth, N.A., Villarini, G., Nayak, M.A. and White, K. (2017). Mixed populations and annual flood frequency estimates in the western United States: The role of atmospheric rivers. Water Resources Research, 53, 257-269.

Bertola, M., Castellarin, A., Viglione, A., Valtancoli, E. and Blöschl, G. (2024). Frequency and spatial variability of European record floods. Water Resources Research, 60(10), e2023WR036767.

Blöschl, G., Hall, J., Parajka, J., Perdigão, R.A.P., Merz, B., Arheimer, B., Aronica, G.T. et al. (2017). Changing climate shifts timing of European floods. Science, 357 (6351), 588-590.

Burn, D.H. and Whitfield, P.H. (2016). Changes in floods and flood regimes in Canada. Canadian Water Resources Journal, 41, 139-150.

Burn, D.H. and Whitfield, P.H. (2018). Changes in flood events inferred from centennial length streamflow data records. Advances in Water Resources, 121, 333-349.

Burn, D.H. and Whitfield, P.H. (2023). Climate related changes to flood regimes show an increasing rainfall influence. Journal of Hydrology, 617, 129075.

Burn, D.H. and Whitfield, P.H. (2025). Shifting cold regions streamflow regimes in North America affect flood frequency analysis. Hydrological Sciences Journal, 70, 51-70.

Frei, C., and Schär. C. (2001). Detection probability of trends in rare events: Theory and application to heavy precipitation in the Alpine region. Journal of Climate, 14, 1568-1584.

Köplin, N., Schädler, B., Viviroli, D. and Weingartner, R. (2014). Seasonality and magnitude of floods in Switzerland under future climate change. Hydrological Processes, 28, 2567-2578.

Maechler, M., Rousseeuw, P.J., Struyf, A., Hubert, M. and Hornik, K.M. (2019). cluster: Cluster Analysis Basics and Extensions. R package version 2 (0).

Mallakpour, I. and Villarini, G. (2015). The changing nature of flooding across the central United States. Nature Climate Change, 5, 250–254.

Rahmani, F. and Fattahi, M.H. (2023). Investigation of alterations in droughts and floods patterns induced by climate change. Acta Geophysica, 1-14.

Tramblay, Y., Arnaud, P., Artigue, G., Lang, M., Paquet, E., Neppel, L. and Sauquet, E. (2023). Changes in Mediterranean flood processes and seasonality. Hydrology and Earth System Sciences, 27, 2973–2987.

Whitfield, P.H. (2012). Floods in future climates: A review. Journal of Flood Risk Management, 5, 336-365.

Whitfield, P.H. (2018). Clustering of seasonal events: A simulation study using circular methods. Communications in Statistics - Simulation and Computation, 47, 3008-3030.

Whitfield, P.H. and Pomeroy, J.W. (2016). Changes to flood peaks of a mountain river: implications for analysis of the 2013 flood in the upper Bow River, Canada. Hydrological Processes, 30, 4657-4673.

Whitfield, P.H. and Burn, D.H. (2026). Rogue and Extreme Floods in North America, submitted to Journal of Hydrology.