Flow Frequency Mixed-Population Analysis: Examples, Tools, and Challenges

Jennifer P. Christensen, Hydrologic and Climate Resilience Engineer, U.S. Army Corps of Engineers, Omaha, NE, Jennifer.P.Christensen@usace.army.mil

George W. Hayes III, Geographer, U.S. Army Corps of Engineers, Omaha, NE, George.W.Hayes@usace.army.mil

Abstract

Flow frequencies are an essential part of floodplain mapping and flood mitigation projects. Fitting flow frequencies to observed annual peak flows is both an art and a science involving methods in Bulletin 17C, available information, and engineering judgment. In most studies, all the observed annual peak flows collected at a site can be used to fit a single Log-Pearson Type III distribution. However, in the case of gage sites where peak flows are driven by a combination of snowmelt, rainfall, ice, and/or rain-on-snow events the peak flow frequency might not provide a realistic fit to the data. A mixed-population flow frequency can be used in these cases to provide a better fit to the observed events. This paper provides examples from two studies using mixed-population analysis. In the one of these studies, mixed-population analysis produced a better fit to the observed annual peak streamflow. In the case of the other study, using mixed-population analysis resulted in nearly identical results to the single-population analysis. This paper discusses why mixed-population analysis was considered, highlights some of the methodology and tools, and describes challenges encountered. Gages considered are from the Elkhorn River Basin in Nebraska and the Yellowstone River in Montana. These sites included historical data but no paleoflood data.

Introduction

Flooding can be produced by many types of event mechanisms depending on the location of the watershed. Example event mechanisms include rainfall, snowmelt, rain-on-snow, ice-affected flows, and hurricane flooding. In many cases, a single-population peak flow frequency (PFF) will realistically represent the probability of flooding at a site. Single population PFFs treated the record of peak flows at a site as one homogenous record and are the state of practice. However, in the case of some watersheds with very different seasonal hydrometeorological events or where extremes are driven by hurricanes, the use of mixed-population analysis can produce a more realistic fit to extreme events.

Challenges presented in this paper related to mixed-population analysis include (1) when to use mixed-population analysis, (2) how to separate annual peak flow events efficiently by mechanism, (3) how to identify ice-affected events, (4) how to extend the record or fill in the data gaps left by assigning the annual peak flow to a mechanism, (5) how to find information for perception thresholds, and (4) how to develop seasonal regional skews for the separate population curves. Examples for this paper are drawn from studies in the Elkhorn River Basin in Nebraska (USACE, 2022a) and in Glendive, Montana on the Yellowstone River (USACE, 2022b). These are referred to as the Elkhorn and the Glendive studies in this paper.

The Elkhorn River Basin is approximately 7,000 square miles. The Elkhorn River extends from the headwaters in the eastern Sandhills and enters the Platte River as a left bank tributary just southwest of Gretna, Nebraska. The Elkhorn Basin has two main mechanisms: snowmelt and rainfall. Snowmelt events typically occur early in the calendar year while rainfall events occur mainly later in the year. In addition, ice jams and rain-on-snow events also are observed in this study area.

Glendive is located on the mainstem of the Yellowstone River in east-central Montana. The city has a long history of ice-jam flooding due to the Yellowstone River's direction of flow from south to north. This results in thawing of snow and ice in the river headwaters before the ice in the lower river thaws. This creates a greater likelihood of ice-jams in the lower river because water moving downstream lifts channel ice in the colder reaches to the north, fracturing it and conveying it downstream. Ice-jams occur when this fractured ice becomes caught in channel bends, on islands, and on structures like bridges in the channel. Ice-affected flows have long been known to impact flooding in Glendive, at least three events have come close to overtopping the levee, and the recent March 2014 event resulted in evacuations and power outages.

Figures 1 and 2 show the location of the Elkhorn River Basin and the city of Glendive, respectively.

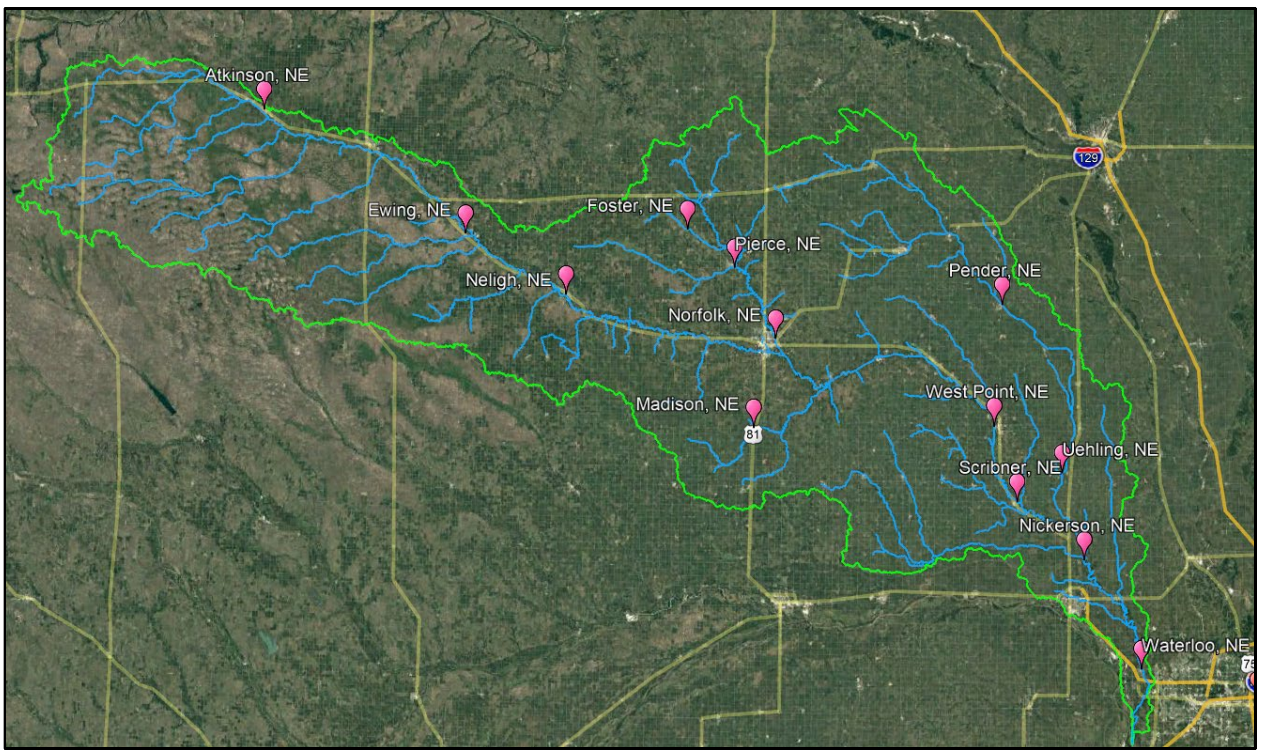

Figure 1. Elkhorn River, Nebraska.

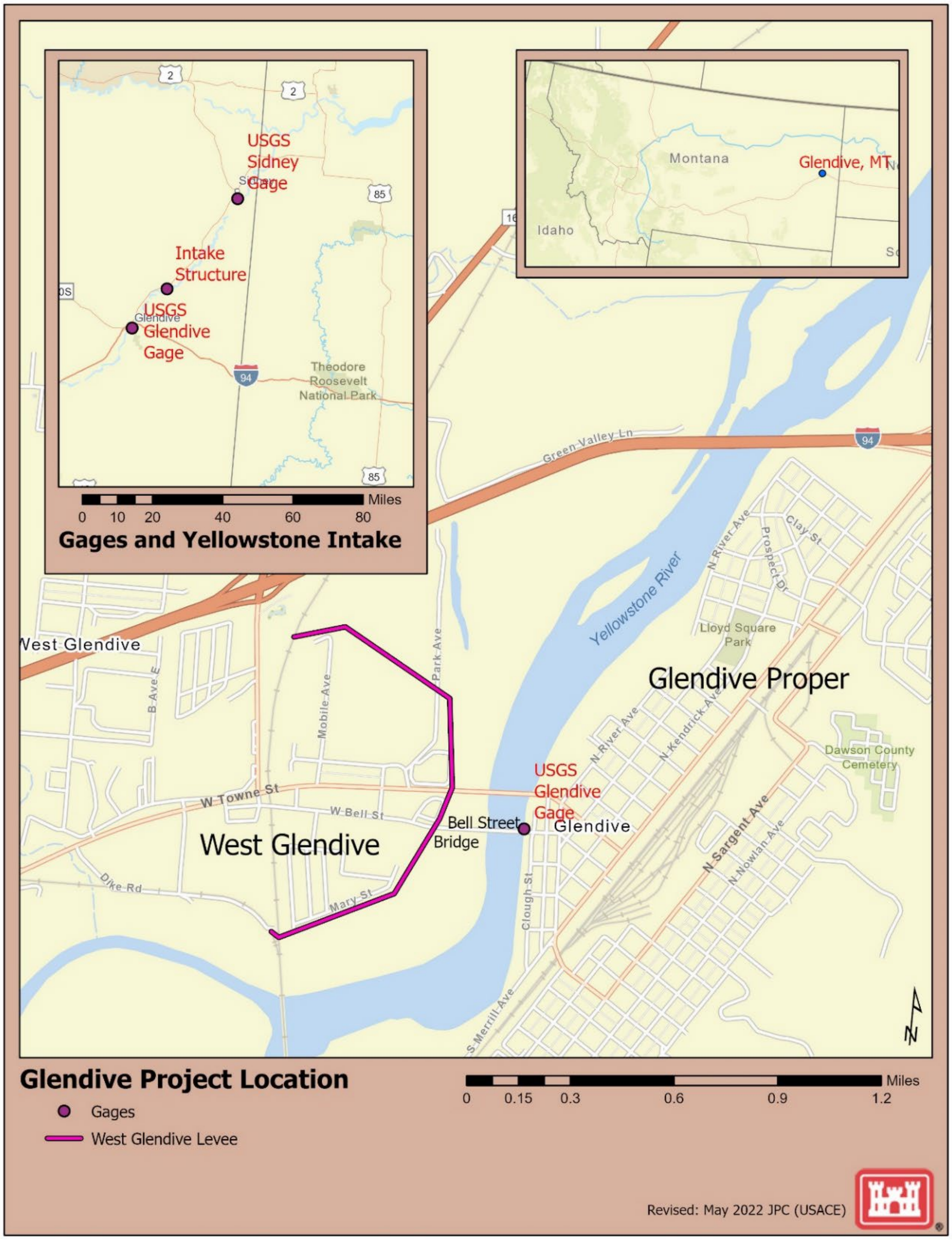

Figure 2. Glendive, Montana.

When to Use Mixed-Population Analysis

A significant challenge in mixed-population analysis is determining when to allocate the additional funding and time needed to do the analysis. In the case of the Elkhorn study a mixed-population analysis was not an apparent need until the extraordinary March 2019 event occurred. In the case of the Glendive study, a mixed-population PFF was considered early on because ice-jam flooding was known to be a significant issue.

The March 2019 event over the Elkhorn River Basin was an "extraordinary" event as defined by Bulletin 17C at some but not all locations. Figure 3 shows how variable the depths of rain plus snowmelt were for the event. The USACE subject matter expert (SME) noted that March 2019 is an extraordinary event as defined in Bulletin 17C and should not be removed from the analysis. It provides indispensable information on what can occur in the Basin. The high uncertainty in the estimated peak flow of this event was captured as a flow interval in the PFF analysis for gage locations with flow interval information.

The March 2019 event was extraordinary because:

- Historically cold February
- Deep front depth. This reduces infiltration into the soil which produces more runoff into channels.
- Record snowpack in January and February.
- Rapid snowmelt.
- Rain on snow.
- Large precipitation depths in the mid-Basin from winter storm Ulmer. Ulmer was a bomb cyclone event.

Mixed-population analysis becomes important when a PFF does not fit all the events of record realistically. This was the case for the Elkhorn study. It is also important to consider when it is well known that floods are created by different mechanisms. In the case of the Glendive study, ice-jams were known to be an issue to flooding. Therefore, mixed-population analysis was considered early in the Glendive study. However, the mixed-population PFF ended up being almost equivalent to the single-population PFF in the case of Glendive because the highest peak flows occurred in the summer. The ice-affected flows (estimated by the U.S. Geological Survey) were relatively low, but their stages were high. Due to this finding and guidance from an SME, a stage-frequency analysis was used along with the ice-season PFF in calibration of a hydraulic model later in the study.

Figure 4 shows the PFF at the Elkhorn River at West Point gage before and after the historic March 2019 event. Both PFFs are single population where the entire annual peak flow record was included in a single Bulletin 17C (B17C) analysis. The preliminary update with the 2019 event included shows that a realistic fit using a single-population analysis was not reasonable for this gage.

Figure 5 shows the updated analysis using a mixed population where events were separated into rainfall and snowmelt populations/series and then recombined using the joint probability theorem to produce a combined mixed-population curve.

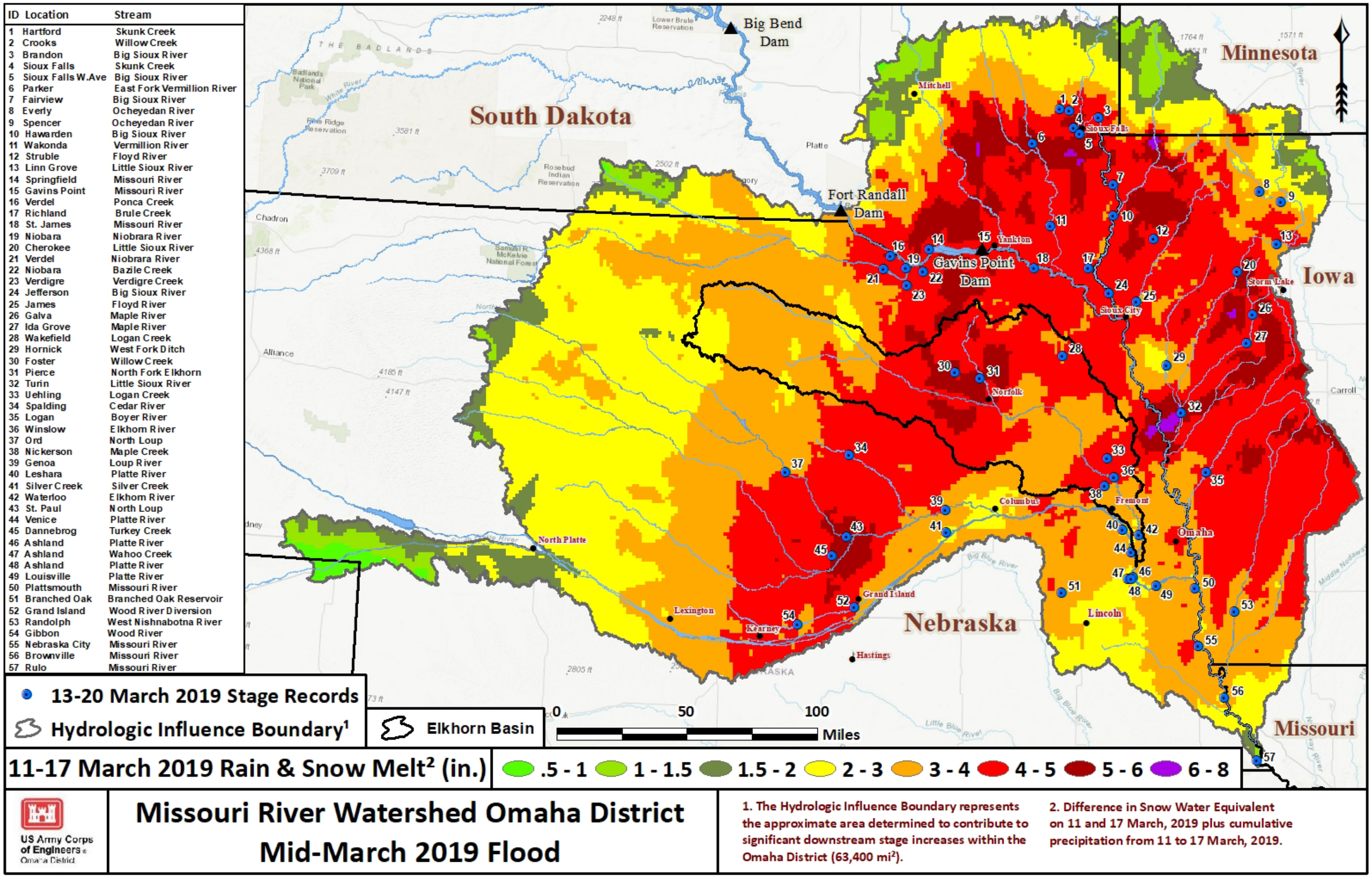

Figure 3. Elkhorn River Bain March 2019 Event

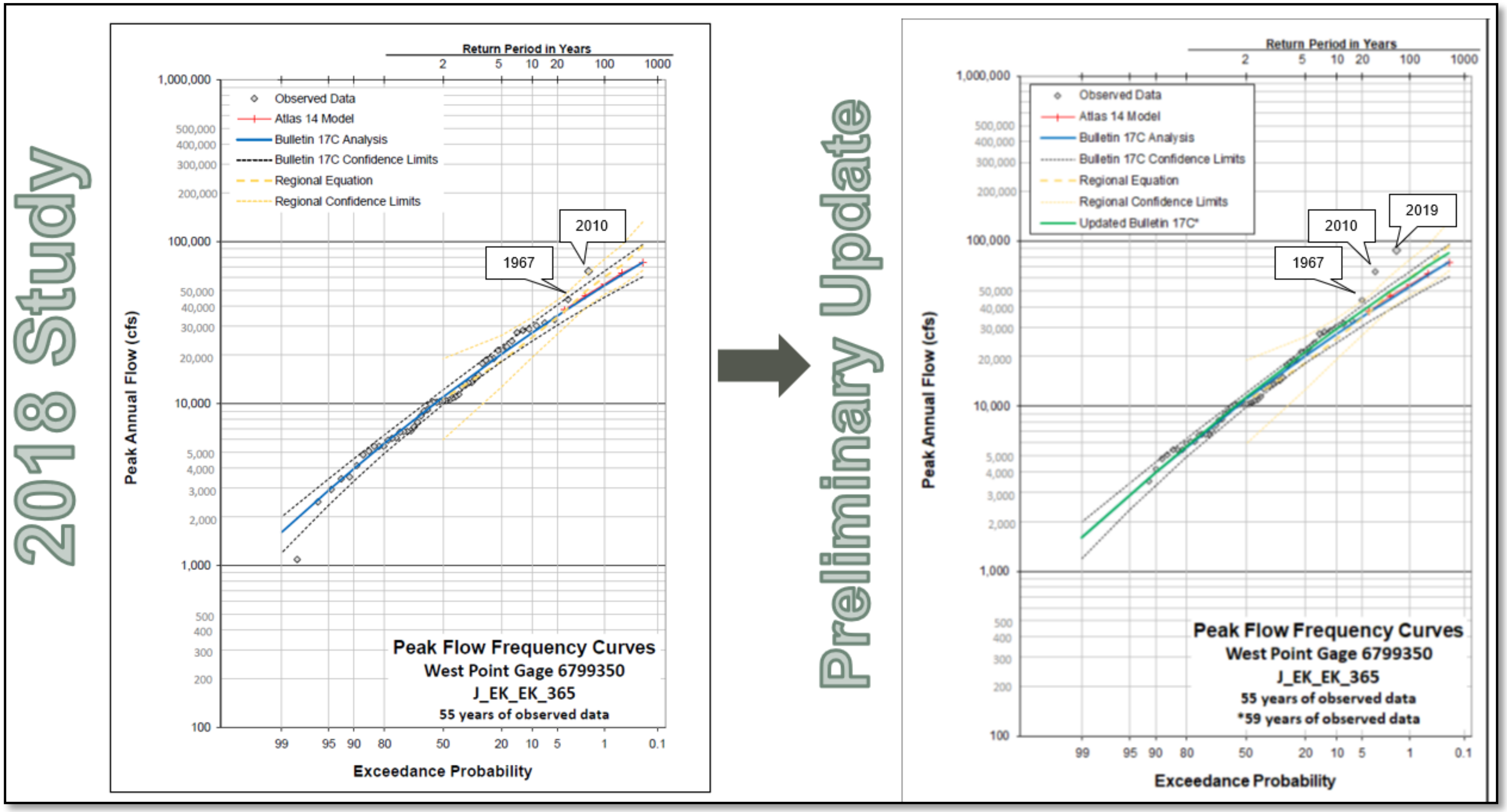

Figure 4. Elkhorn River at West Point – Single Populations

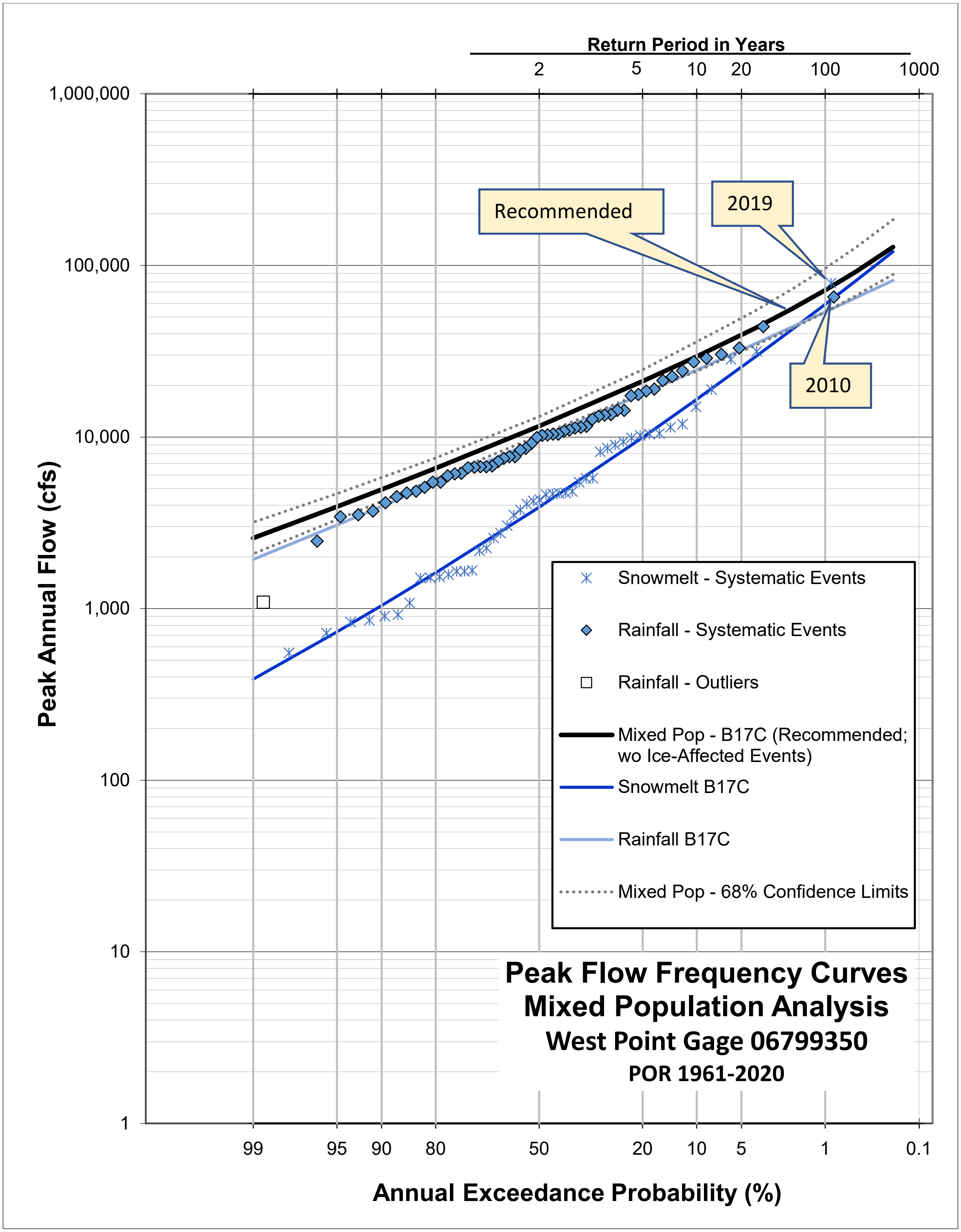

Figure 5. Elkhorn River at West Point – Mixed Population

Event Separation

Ideally, each annual peak flow event should be investigated using hydrometeorological data to estimate its mechanism. This involves collecting temperature, precipitation, and snow depth records at or very near the site and using engineering judgement to separate events into separate populations (e.g., rainfall, snowmelt, rain-on-snow, etc.).

A challenge with the Elkhorn study was that the number of gages included in the study made separation by mechanism on an event-by-event basis time and budget prohibitive. For this reason, an informed seasonal approach was used leveraging information from a past study and the accumulative freezing degree day (AFDD) method. A past study by JEO Consulting Group, Inc. (2016) used a seasonal separation of January through March for snowmelt (including rain-on-snow) and April through December for rainfall driven events in the Elkhorn River Basin. To validate this seasonal separation, events were sampled near the boundary and the AFDD method used to validate if the events were driven by snowmelt or rainfall. For this study, rainfall on snow events were considered in the snowmelt population.

Of the 14 February and March events analyzed at the Elkhorn River at West Point gage using the AFDD method, 50% were snowmelt, 36% were rain-on-snow, and 14% were rainfall driven. Since 86% of these sampled events were either snowmelt or rain-on-snow, the seasonal assumption of January through March for snowmelt was accepted.

Figure 6 shows the AFDD method used in the classification of the March 2019 event. This event was a rain-on-snow event.

Another challenge in the Elkhorn study was separating out the ice-affected events. These ice-affected events were identified through plotting the rating curve for the gage site against the observed events. Figure 7 shows this for the West Point gage on the Elkhorn River. Events above the channel rating curve with high stages for small discharges could then be investigated farther and removed from the record if found to be ice-affected.

In the case of the Glendive study, all the annual peak flow events occurred in May, June, or July except in the case of the March 29 event. Thus, most of the peak flow events occurred in the open-water season not affected by ice-jams. Event separation was not an issue with Glendive due to almost all the events being too far into the open-water season to be affected by ice.

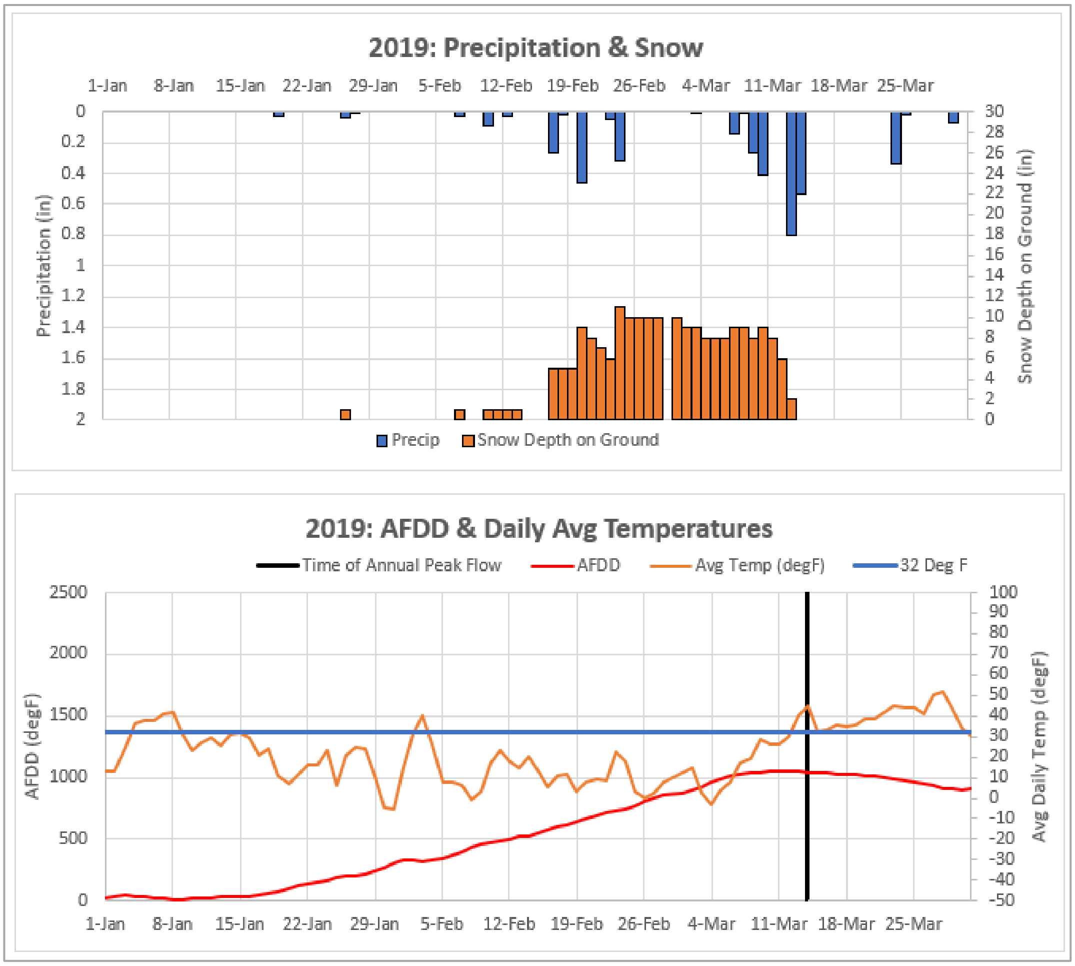

Figure 6. Elkhorn River at West Point – Mixed Population

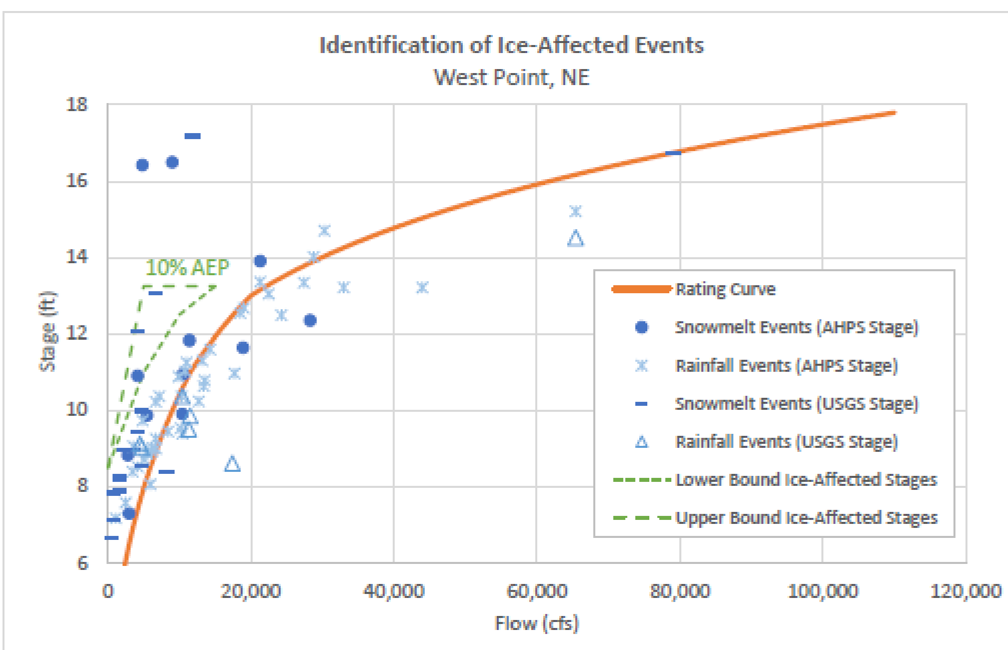

Figure 7. Identification of Possible Ice-Affected Events

Record Extension

The event separation needed for a mixed-population analysis creates an additional challenge – data gaps. Record extension methods were used to use additional data to fill in the gaps. Record extension when used in this paper means filling in missing data within the POR of the observed data, not extrapolating the data forward or backwards in time.

In the case of the two example studies, data gaps were filled using instantaneous 15-minute flows for the missing event type, daily data transformed to an instantaneous peak flow, and Bulletin 17C perception thresholds. Daily data were transformed into an instantaneous peak flow estimate using a peak flow to daily flow ratio developed from overlapping peak and daily flow data.

In the case of the Glendive study, the Bulletin 17C MOVE3 method in the Hydrologic Engineering Center Statistical Software Package (HEC-SSP) was also used. HEC-SSP is the program used by the U.S. Army Corps of Engineers for Bulletin 17C analysis, as well as other statistical analysis. It is publicly available from the HEC website (HEC, 2023). The MOVE3 record extension could be used in the case of the Glendive gage because it had a high enough correlation in its overlapping record with the Sidney gage downstream. The open water events had correlation of 0.988 for the 20 overlapping events and the ice-affected events has a correlation of 0.964 for the 20 overlapping events. The MOVE3 analysis replaces the two-station comparison methodology used in Bulletin 17B. Figure 8 shows an example record extension in HEC-SSP.

Bulletin 17C perception thresholds were used in cases where missing data in either the rainfall or snowmelt populations could not be completed with instantaneous (15-minute gage) records, daily data transformed to instantaneous peaks, or through the MOVE3 analysis. A perception threshold (PT) is the range of flows that would have been measured or recorded had they occurred (England Jr. et. al., 2018). Perception thresholds apply to events that could have occurred (but didn't) during a period when a stream gage was not in operation (e.g., before the establishment of the stream gage, gage discontinued, etc.). Thought of another way, a PT is the range at which a flow would have been considered significant enough to be noticed and estimated had it occurred.

In the case of the Elkhorn study, record extension was less of an issue because multiple agencies have collected both annual peak flows, instantaneous data, and daily data on the Elkhorn River and its tributaries. Most of the record data gaps were filled in using instantaneous 15-minute flows for the missing event type, daily data transformed to an instantaneous peak flow, and Bulletin 17C perception thresholds.

Figure 9 shows the perception threshold for West Point gage for the snowmelt events. This perception threshold was estimated using stage impact information for the Advanced Hydrologic Prediction Service (AHPS; NOAA, 2021) at the West Pont gage converted to a flow using the channel rating curve.

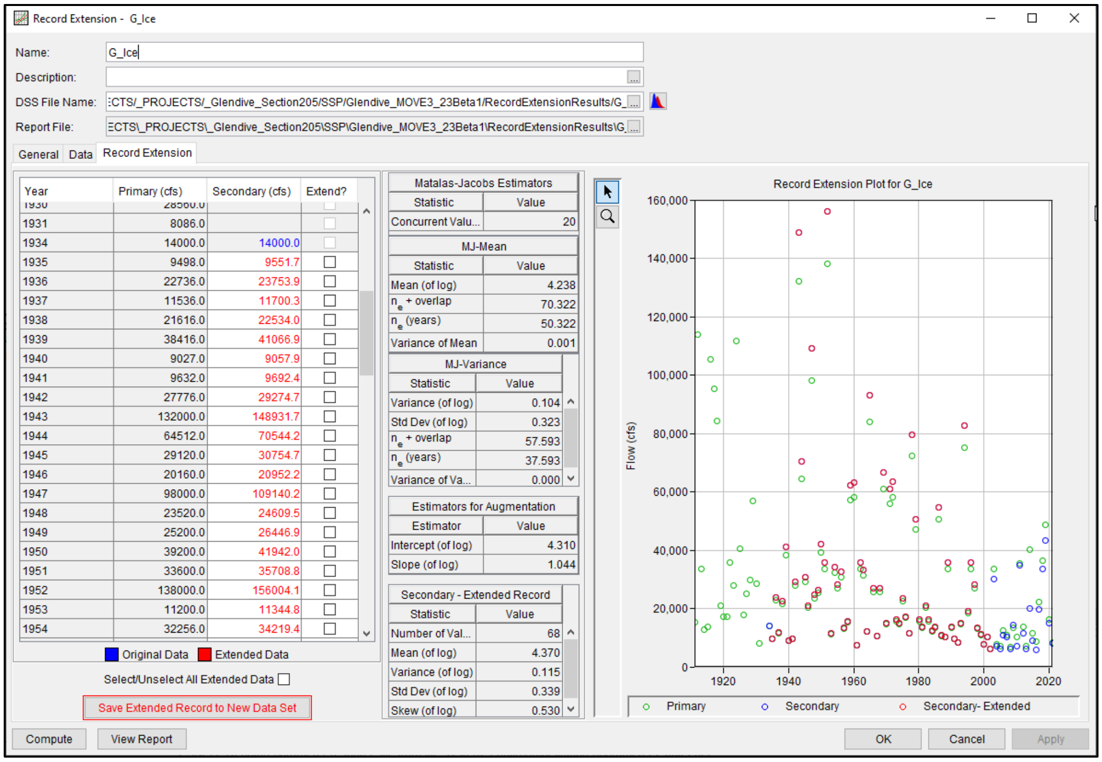

Figure 8. MOVE3 analysis in HEC-SSP

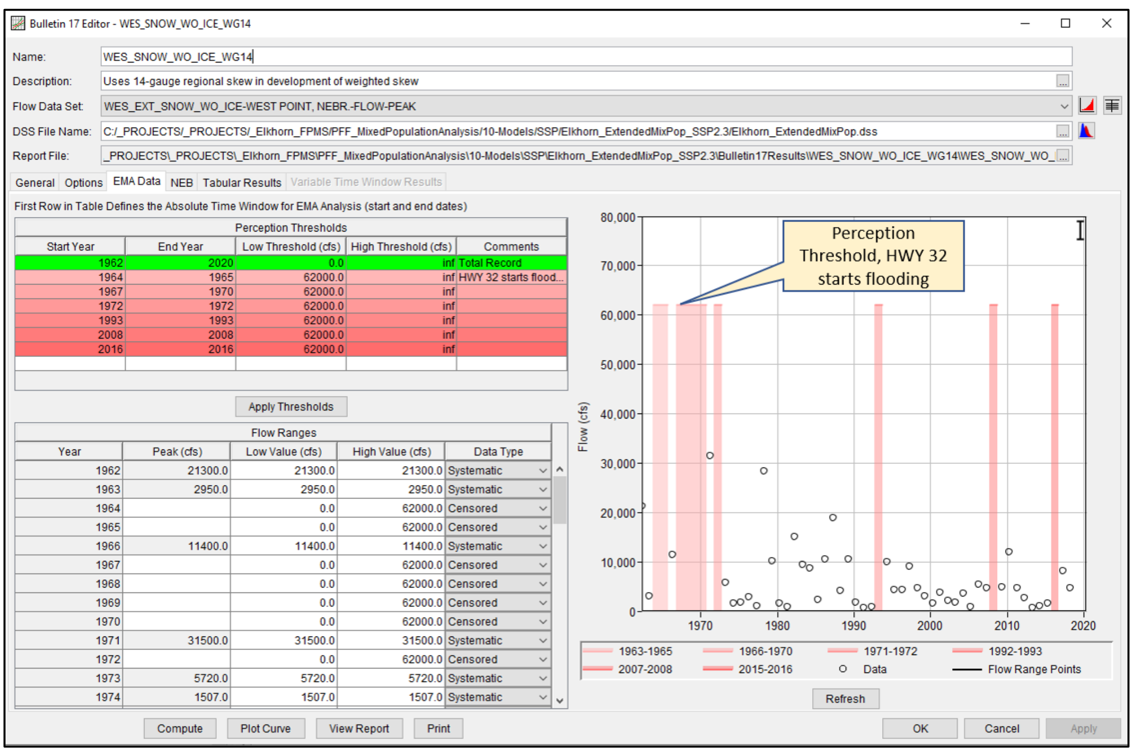

Figure 9. West Point Gage Snowmelt Events PFF Perception Threshold

Skew

Another challenge was determining a seasonal regional skew that could be used to determine a weighted skew for the B17C analysis. The general/regional skews reported in B17B were developed from single-population data. Separate seasonal regional skews were developed in the Elkhorn study by using stations skews and record lengths from the 14 gages considered in the study for the two different populations of events (rainfall and snowmelt).

The seasonal regional weighted average skew for each population (rainfall and snowmelt) was calculated through Equation 1 where G_{SR} is the weighted average seasonal regional skew, x is the total number of gages used (14 in the case of Elkhorn) to determine the seasonal regional skew, n is the number of events for an individual gage i, and G_{si} is the station skew of an individual gage i.

$$G_{SR} = \frac{\sum_{i=1}^{x} n_i * G_{Si}}{\sum_{i=1}^{x} n_i} \quad (1)$$

The mean square error (MSE) for each of the seasonal regional skews was determined through Equation 2 where n is the total number of gages used to determine the seasonal regional skew, G_{SRi} is the weighted average seasonal regional skew for an individual gage i, and G_{si} is the station skew of an individual gage i.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(G_{SRi} - G_{Si})^2 \quad (2)$$

It is recommended that a regional/generalized skew determined from single-population analysis not be used in mixed-population analysis to determine weighted skews. Once the events are separated into separate populations and record extension used to help fill in data gaps, it is the author’s experience that they will have significantly different skews. If a seasonal regional skew cannot be estimated from other gages in the area, the station skew is likely best.

Mixed-Population Flow Frequency

Figures 10 and 11 show the mixed-population results for the sampled gages from the Elkhorn and Glendive studies. An interesting note about the mixed-population PFF developed for Glendive is that it is nearly equivalent to the single-population PFF. This is because nearly all the annual maximum flows in the Glendive record occur during the open-water season. However, the ice-affected PFF was needed in the study for ice-jam modeling in the hydraulic model, so the additional effort of a mixed-population analysis was not wasted.

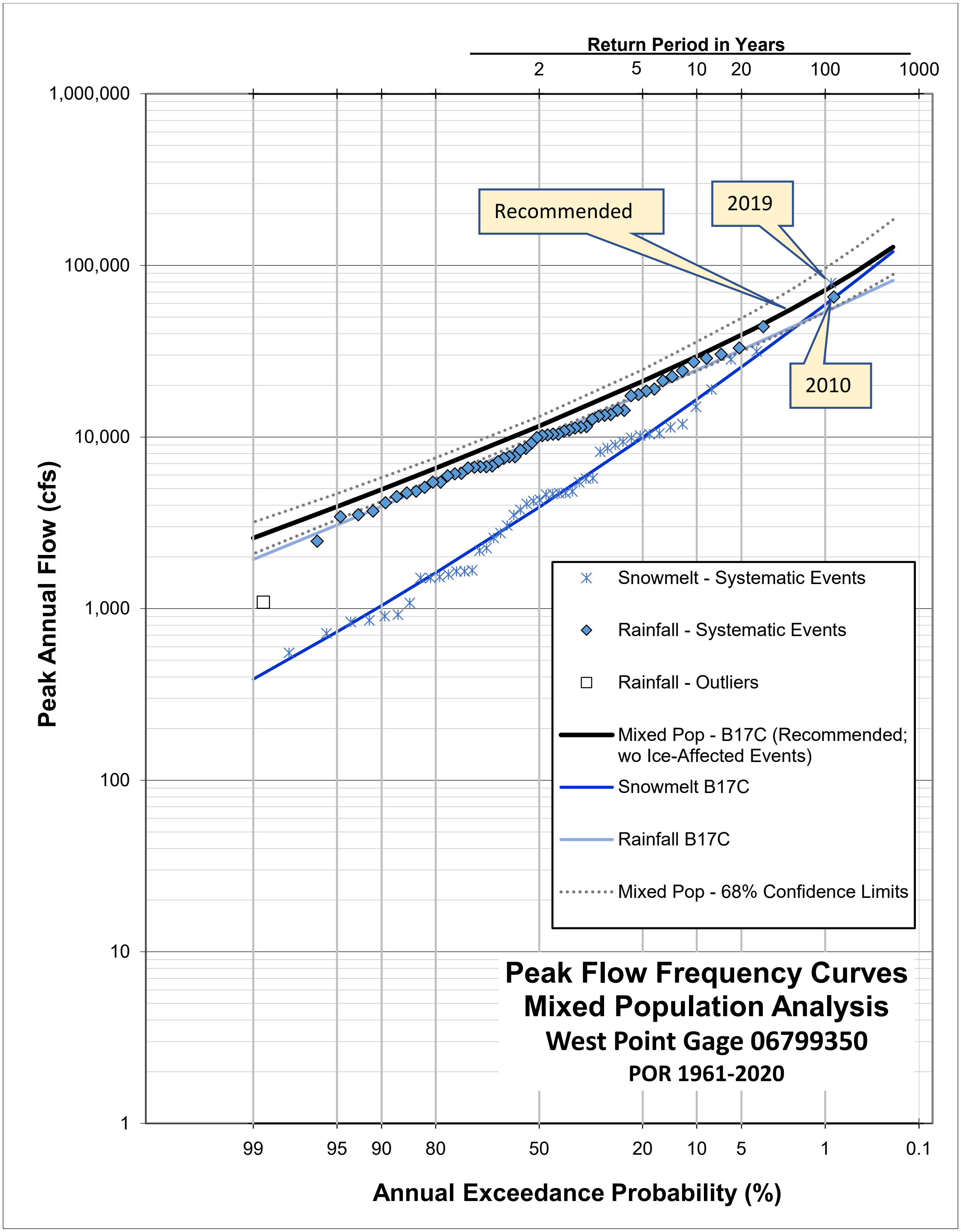

Figure 10. Mixed-Population PFF from Elkhorn Study

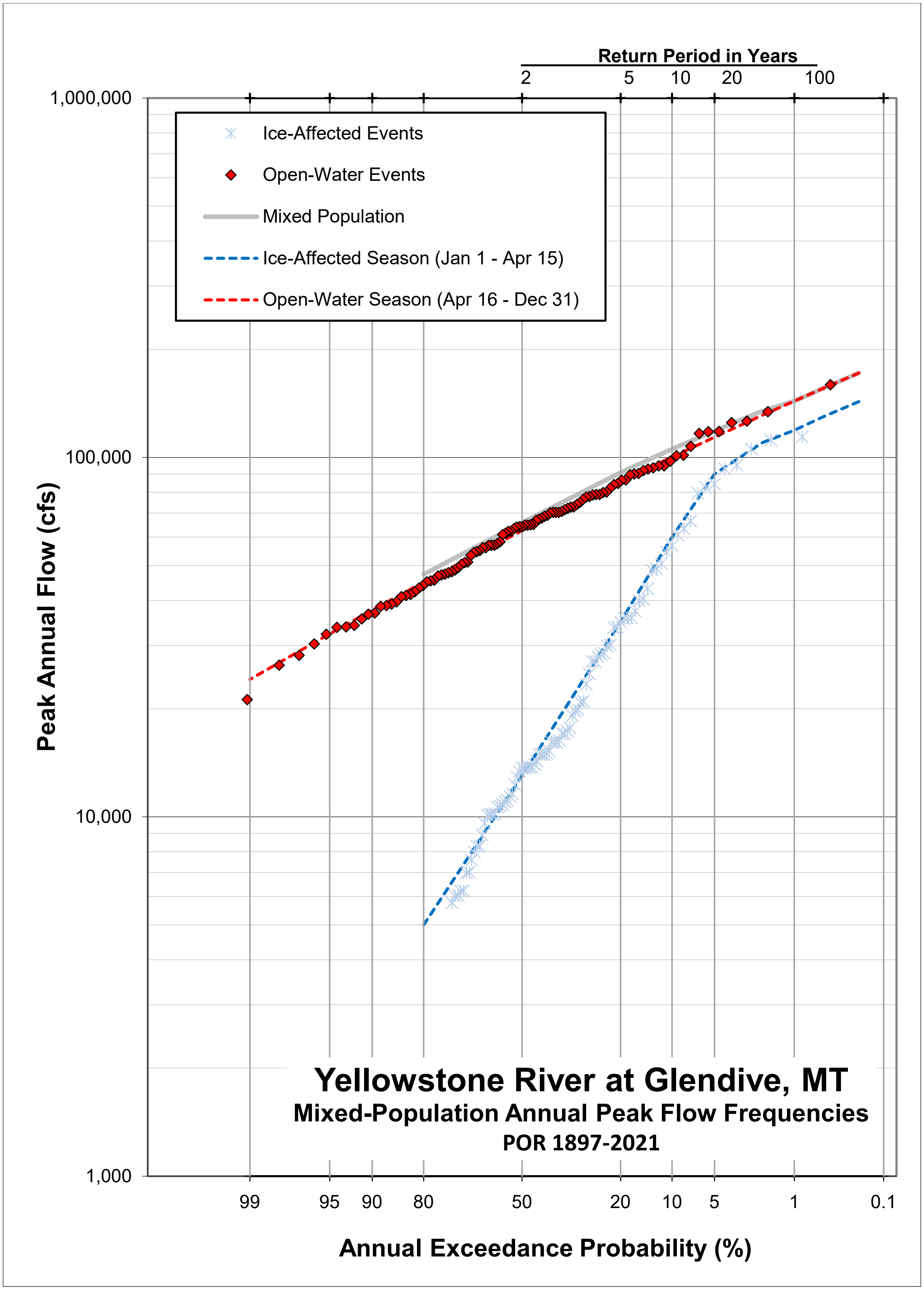

Figure 11. Mixed-Population PFF from Glendive Study

Conclusions

This paper discussed some of the challenges encountered in the mixed-population PFF analysis of gages from two studies. One study, the Glendive study, it was known ahead of time that a mixed-population analysis was needed to help quantify ice-jam flooding. In the other study, the Elkhorn River study, it was determined that a mixed-population analysis was needed after an extraordinary event occurred in the basin.

Challenges presented in this paper related to mixed-population analysis include (1) when to use mixed-population analysis, (2) how to separate annual peak flow events efficiently by mechanism, (3) how to identify ice-affected events, (4) how to extend the record or fill in the data gaps left by assigning the annual peak flow to a mechanism, (5) how to find information for perception thresholds, and (4) how to develop seasonal regional skews for the separate population curves.

Mixed-population analysis becomes important when a single-population PFF does not fit all the events of record realistically, when it is known that flooding is created by different mechanisms, when the record includes an extraordinary event, or when the basin is impacted by hurricanes or ice-jam flooding. One of the important methods of the study is that the estimated impact stages from the Advanced Hydrologic Prediction Service (AHPS) can be used to estimate perception thresholds given a channel rating curve.

The recommended method for a mixed-population analysis includes: (1) develop B17C results assuming a single-population analysis to determine if a mixed-population analysis is necessary, (2) consider project needs for hydraulic analysis and if ice-affects are needed in the study, (3) collect additional streamflow data (instantaneous data, daily data), (4) determine a daily to instantaneous flow data transform ratio if needed, (5) use record extension methods (data transformation, MOVE3 analysis, etc.) to fill in data gaps, (6) determine seasonal regional skews if possible.

References

England Jr., et. al. 2019. Guidelines for Determining Flood Flow Frequency Bulletin 17C. Techniques and Methods 4-B5 Version 1.1, May 2019. Accessed September 2021 at: https://pubs.usgs.gov/tm/04/b05/tm4b5.pdf.

HEC. 2023. U.S. Army Corps of Engineers Hydrologic Engineering Center. HEC-SSP. Accessed April 2023 at: https://www.hec.usace.army.mil/software/hec-ssp/.

NOAA. 2021. National Weather Service Advanced Hydrologic Prediction Services. Accessed on September 2021 at: https://water.weather.gov/ahps2/index.php?wfo=oax.

JEO Consulting Group, Inc. 2016. West Point Levee System: Levee Certification Feasibility Evaluation Phase 2: Elkhorn River Hydrology and Hydraulics Report. JEO Consulting Group, Inc. August 2016.

USACE. 2022a. Elkhorn River Flood Plain Management Services (FPMS) Peak Flow Frequency Improvement with Mixed Population Hydrology and Subject Matter Expert Elicitation. U.S. Army Corps of Engineers Omaha District. August 2022.

USACE. 2022b. Glendive Section 205 Flood Frequency and Stage Frequency Analysis Using Single and Mixed-Population Analysis. U.S. Army Corps of Engineers Omaha District. August 2022.