Fall 2018

Flood frequency and mixed populations in the western United States

Nancy A. Barth
University of Iowa

FLOOD FREQUENCY AND MIXED POPULATIONS IN THE WESTERN UNITED STATES

by

Nancy A. Barth

A thesis submitted in partial fulfillment
of the requirements for the Doctor of Philosophy
degree in Civil and Environmental Engineering in the
Graduate College of
The University of Iowa

December 2018

Thesis Supervisor: Associate Professor Gabriele Villarini

To Dr. Timothy A. Cohn (1957-2017) whose career and passion related to the methodological research in this thesis is best described by his former colleague.

"He [Tim Cohn] also developed a much-improved approach for eliminating the influence of "low outliers" (years with no real flood events) from the analysis of flood frequency. As was typical for Tim's work, he stressed the importance of quantifying the uncertainty, and he reminded practitioners of just how little we actually know. The analysis methods he developed are at the core of the new interagency manual [B17C] for flood frequency analysis (now in the final stages of review), the first update of official flood frequency methods in more than 40 years." (Hirsch, 2017).

"Ithaka:

Pray that your journey be long,
full of many summer mornings
when with much pleasure and much joy
you anchor in harbors never seen before;
...eager and open to receive learning.
Keep Ithaka always in your mind.
Your arrival there is your destiny.
But do not hurry at all; *BE PATIENT*!
Better that it lasts for many years—
longer than you can ever imagine.
So that finally, when you reach this
sacred isle, you will be a wise woman,
abundantly filled by all you have gained along the way;...
Ithaka offered you the profound journey,
the chance to discover the woman you have always been.
Without Ithaka as your inspiration, you
never would have set out in search of Wholeness."

~Sarah Ban Breathnach
Simple Abundance

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere appreciation to my graduate advisor, Dr. Gabriele Villarini. I began my doctorate studies four years ago after moving from northern California to Iowa in the winter of 2015, giving up my flip flops for my knee-high Merrell "Iowa winter boots." I sincerely thought the hardest part about getting acclimated to living and going to school in Iowa would be the daily three mile roundtrip walk to school and work. Little did I know it would be the very steep learning curve of having to learn to write my own codes for my research? Dr. Villarini knew my lack of any experience writing scripts and was quite patient with me over the years. I truly value the attention to detail and passion Gabriele has for his research, students and colleagues. I know working with you was the best thing to happen to both of us. And I know we both grew in our own ways through this four-year collaborative experience. I would like to thank my committee members for serving on my committee, and fellow coauthor Dr. Kate White.

I would like to thank my colleagues and mentors at the U.S. Geological Survey. Chuck Parrett who fostered my interest in flood hazards and statistics and who was also my master's thesis advisor, Janet Curran who taught me about flood hydrology in Alaska, and Nick Paretti who helped me understand more about desert flood hydrology. I also would like to thank Robert Mason who supported my work on important studies at the national level, Dr. Julie Kiang who always made herself available for questions and advice, and Dr. Andrea Veilleux. And I would sincerely like to thank my colleagues who worked tirelessly on the HFAWG committee to improve the federal flood frequency standards that have now been updated in B17C: Dr. Tim Cohn, Dr. John England and Dr. Beth Faber, and Dr. Jery Stedinger. You are all such an

inspiration and your passion fueled my desire to go back to get my doctorate [at my young] to learn more about the broader topics in applied statistical hydrology.

I would also like to thank my incredibly supportive colleagues and fellow research group members: Dr. Wei Zhang, Dr. Munir Nayak, Dr. Iman Mallakpur, Beda Luitel, Yog Aryal, Dr. Kaustubh Salvi, Dr. Louise Slater, Dr. Abdou Khouakhi, Dr. Manuel Rios Gaona and Munsung Keem.

I would like to acknowledge the support of my research that was supported by an appointment to the U.S. Army Corps of Engineers (USACE) Research Participation Program administered by the Oak Ridge Institute for Science and Education (ORISE) through an interagency agreement between the U.S. Department of Energy (DOE) and USACE. And the support of the Ballard Seashore Dissertation Fellowship by the Graduate College.

I want to thank all of the wonderful people who have shared my Sojourn Path over the past six years since 2012 and most especially the people who have become part of my Iowa City family, Margaret and Paul Heidger, my neighbors in University Heights and the people at St. Andrew Presbyterian Church. I would not have been able to realize this dream being so far away from my family and friends back home in California.

And most especially I would like to thank my loving, caring and supportive parents who helped me realize this dream. With their unconditional support, they continually affirmed my dedication to strive for a degree that represents my vocare (my calling). And I would like to thank my sister and brother-in-law and my much-needed source of laugher, my dear friend D'An Fuller.

ABSTRACT

Flood frequency analysis over the western United States is complicated by annual peak flow records that frequently contain annual flows generated from distinctly different flood generating mechanisms. Bulletin17B (B17B) and its update Bulletin 17C (B17C) recognized the difficulties in determining flood frequency estimates with streamflow records that contain a mixed population of flood generated peaks, and recommend developing separate frequency curves when the hydrometeorologic mechanisms that generated the annual peak flows can be separated into distinct populations. Yet challenges arise when trying to consistently quantify the physical process that generated the observed flows. This thesis examines the role played by different flood producing mechanisms in generating annual maximum floods throughout the western United States using process-driven mixed populations.

First I evaluate the impacts of hydrometeorological processes on flood frequency in the western United States, with emphasis on the spatial and fractional contributions of atmospheric rivers (ARs) and eastern North Pacific tropical cyclones and their remnants (TC events) on annual maximum flows throughout this area. Six main areas in which flooding are impacted by ARs at varying degrees are found throughout the western United States. The Pacific Northwest and the northern California coast have the highest fraction of AR-generated peaks (~80–100%), while eastern Montana, Wyoming, Utah, Colorado, and New Mexico have nearly no impacts from ARs. The individual regions of the central Columbia River Basin in the Pacific Northwest, the Sierra Nevada, the central and southern California coast, and central Arizona all show a mixture of 30–70% AR-generated flood peaks. Analyses related to the largest flood peaks on record highlight the strong impact of ARs on flood hydrology in this region. Conversely, TC events play a limited role in controlling the upper tail of the flood peak distributions across the

western United States. Southern California, Arizona, southernmost Nevada and Utah, southern and western New Mexico, central Colorado, and Texas have the highest fractional contributions of TC-event-generated annual maximums flows (~5-14%).

I then build on these insights to develop a statistical framework to perform a process-driven flood frequency analysis using the AR/non-AR-generated annual peak flows identified at 43 long-term U.S. Geological Survey (USGS) streamgages in the western United States. I use a simulation framework to perform flood frequency analyses in terms of mixed distributions and quantify the corresponding uncertainties by accounting for mixed populations. Sites with notably different quantile estimates in the upper tail of the distribution between the single (homogeneous) and the weighted (heterogeneous) population methodologies are found when (i) potentially influential low floods (PILFS) are identified and/or (ii) when the composite distribution contains markedly different at-site log-unit skews (shape parameter) among the AR/non-AR subpopulations compared to the single homogeneous population.

PUBLIC ABSTRACT

The need for reliable flood estimates are essential for flood insurance studies, floodplain management, and the design of transportation and water-conveyance structures, such as roads and levees. Federal, state and local officials rely on this information to protect lives and properties and effectively plan and manage water resources. The federal guidelines of Bulletin 17B (B17B) were established more than 35 years ago (and recently updated; Bulletin 17C (B17C)) to provide a consistent framework for determining flood frequency estimates throughout the United States.

Flood hydrology throughout the western United States covers several hydrologic regimes, from the rainforests of the Pacific Northwest to the arid desert regions in the Southwest. Flood records often contain combinations of multiple zero flows and/or low outliers, and extreme flood events produced by winter rainfall, rain-on-snow, snowmelt, atmospheric rivers or tropical cyclones. Thus streamflow records throughout the western United States frequently contain flows generated from distinctly different flood generating mechanisms.

B17B/B17C recognize the difficulties in determining flood frequency estimates with streamflow records containing mixed populations of flood peaks. They recommend developing separate frequency curves when the hydrometeorologic mechanisms that generated the flows can be separated into distinct populations.

The overall objective of my thesis is to determine the role played by different flood generating mechanisms among annual maximum floods throughout the western United States using process-driven mixed populations.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

INTRODUCTION

> "[T]he main emphasis in stochastic analysis of hydrological processes…has been on the fitting of various preconceived mathematical models to empirical data rather than on arriving at a proper model from the physical nature of the process itself…Thus what we usually find is not, in fact, statistical hydrology but merely an illustration of statistical and probabilistic concepts by means of hydrologic data." (Klemeš, 1974, p.2)

The need for reliable estimates of the magnitude and frequency of floods are essential for flood insurance studies, floodplain management, and the design of transportation and water-conveyance structures, such as roads, bridges, culverts, dams and levees. Federal, state and local officials rely on these estimates to protect lives and properties, as well as effectively plan and manage land use and water resources. The federal guidelines of Bulletin 17B (B17B; Interagency Advisory Committee on Water Data (IACWD),1982) were established more than 35 years ago to provide a consistent framework for determining flood frequency estimates used for flood design structures throughout the United States. The B17B framework recommends the use of the log-Pearson Type III (LP3) distribution to describe the populations of floods for all hydrologic regimes found throughout the United States. The conventional assumptions for performing flood frequency analyses are that the annual time series is a representative time sample of random homogeneous events and that the stochastic processes that generate floods are assumed to be stationary or invariant in time *(*IACWD, 1982). Essentially the streamgage records are assumed to be independent and identically distributed (i.i.d) and to be generated by a single flood generating mechanism. Yet B17B and the recent updates to B17B (Bulletin 17C (England et al., 2018); herein referred to as B17C) recognize these fundamental assumptions are not always

valid, and describe locations throughout the western United States, for example, where these assumptions are most likely violated.

Flood hydrology throughout the western United States covers several extreme hydrologic regimes, from the rainforests of the Pacific Northwest to the arid desert regions in the Southwest. Flood series may often contain combinations of multiple zero flows and/or potentially influential low floods (PILFs; low outliers that substantially deviate from the overall pattern in the data), as well as extreme flood events produced by winter rainfall, rain-on-snow, snowmelt, or remnants of hurricanes and tropical cyclones, among others (e.g., Smith et al., 2011; Villarini, 2016). Thus annual peak flow records throughout the western United States frequently contain flows generated from distinctly different flood generating mechanisms (e.g., Hirschboeck, 1988; Webb and Betancourt, 1992; Smith et al., 2011; Berghuijs et al., 2016; Villarini, 2016; Barth et al., 2017; Villarini and Slater, 2017). It is common for the empirical distribution of the annual peak flows in this region to show sharp breaks in the slope or an s-shaped curve that reverses direction when plotted with respect to the LP3 distribution. Figure 1.1 illustrates this behavior at the high-elevation site in California (U.S. Geological Survey (USGS) streamgage station Fall Creek near Hetch Hetchy, California (11275000)) with the s-shaped curvature in the upper right hand tail of the distribution. Rather than being an issue with the selected parametric distribution (i.e., LP3), this result suggests that the observational record could be stratified into at least two groups, one that mostly encompasses the low to medium flood peaks, and one that mostly includes large and extreme flood events from the upper tail of the distribution.

What is found at Fall Creek near Hetch Hetchy (Figure 1.1) is not an isolated case, but rather is representative of flood frequency results that are often encountered across the western United States. This points to the potential presence of different flood generating mechanisms associated

with flood peaks that characterize different parts of the distribution. In the western United States, atmospheric rivers (ARs) and eastern North Pacific tropical cyclones (TCs) are primarily responsible for extreme precipitation and flooding. ARs are long, narrow corridors of vertically-integrated water vapor transport (IVT) that are responsible for nearly 90% of the extratropical poleward water vapor transport despite only covering 10% of the circumference at a given latitude (Zhu and Newell, 1998). ARs have been frequently associated with heavy rainfall, large snowfall totals, and destructive flooding mostly in the November-April period. During these cold-season months, the storms interact with topographic barriers (e.g., the Cascade and Sierra Nevada ranges) leading to orographically-enhanced precipitation. Heavy precipitation events associated with these storms have been studied at the watershed or regional scale in northern Washington (Neiman et al., 2011), northern and central California (Ralph et al., 2006; Ralph and Dettinger, 2011; 2012; Guan et al., 2013), Arizona (Neiman et al., 2013), and Montana (Bernhardt, 2008). However, much less is known about the role ARs play as a process-driven mechanism in the flood hydrology throughout the western United States. On the other hand, TCs tend to affect the southwestern United States mostly during the summer months. Wood and Ritchie (2013) found that on average TCs over Arizona, New Mexico and Texas contributed to about 10% of the precipitation during 15 June to 30 September over the 21-year period of 1989-2009, with higher fractions in California. Similar regional results were found by (Barlow, 2011), although he looked at hurricanes (TCs with maximum sustained winds in excess of 64 knots) rather than TCs more broadly.

More generally, mixed populations have been documented at USGS streamgaging stations in the high elevations of the Sierra Nevada (Parrett et al., 2010; Gotvald et al., 2012), western Oregon (Copper, 2005), Washington (Neiman et al., 2011; Mastin et al., 2016), northeastern

Idaho (MacDonald and Hoffman, 1995), northwestern Montana (MacDonald and Hoffman, 1995; Parrett and Johnson, 2003; Sando et al., 2015), and Arizona (Hirschboeck, 1987, 1988; Webb and Betancourt, 1992; Thomas et al., 1997; Ailia and Mtiraoui, 2002; Paretti et al., 2014).

B17B and B17C recognize the difficulties in determining flood frequency estimates with streamflow records that contain a mixed population of flood generated peaks such as those found in the Sierra Nevada region (e.g., Figure 1.1). They recommend developing separate frequency curves when the hydrometeorologic mechanisms that generated the annual peak flows can be separated into distinct populations, and discourage the separation of annual peaks solely based on calendar periods (e.g., peaks in October-February vs. peaks in June-September). Yet in practice many of the above-mentioned studies have documented the challenge of quantifying the physical processes that generated the observed flows and the limited availability of long-term hydrometeorologic datasets required to determine meaningful flood frequency estimates. With limited data readily available, hydrologists are often forced to employ inconsistent hydrologic judgement to qualify the physical mechanisms that generated the peaks. The separation between stochastic hydrology and physical processes, and the fact that flood frequency analysis has become a mere application of statistics without accounting for any physical insights have been well summarized by Klemeš (1974, p.2; see quote at the beginning of the chapter), and Hirschboeck (1988):

"In some circles, however, the obvious fact that these [annual peak flow] values represent a response to varying processes in the physical world has tended to become less important than the urge to statistically model flood values in search of the best fit of the observed data and therefore (ideally) the best predictive capability of future flows…".

Despite the recognized importance of properly identifying the physical processes that generate annual floods among mixed populations to obtain more accurate flood frequency estimates (e.g., Klemeš, 1974; IACWD, 1982; Hirschboeck, 1987; 1988, Webb and Bentancourt, 1992; Alila and Mtiraoui, 2002; Smith et al., 2011; Barros et al., 2014), the topic has regretfully received little attention.

The Work Group that updated the guidelines for performing flood frequency in Bulletin 17C 'did not conduct an evaluation of [mixed population] procedures…'. In the 'Future Studies' section of Bulletin 17C, the Work Group identified the need for (among several other) important topics of study: '[1] the identification and treatment of mixed distributions, including those based on hydrometeorological or hydrological conditions'(England et al, 2018).

With the diverse and complex flood hydrology found throughout the western United States as described in this chapter, the overall objective of my thesis is to determine the role played by different flood producing mechanisms in generating annual maximum floods throughout the western United States using process-driven mixed populations. The structure of my thesis is as follows:

- In Chapter 2, I evaluate the spatial and fractional contributions of atmospheric rivers (ARs) on annual maximum flows throughout the western United States at long-term USGS streamgages; I also examine both the role of AR-generated peak flows in the upper tail of the flood peak distribution and on flood quantile estimates.
- In Chapter 3, I evaluate the spatial and fractional contributions of eastern North Pacific tropical cyclones (TCs) and their remnants (TC-events) on flooding in the western United States.

- Chapter 4 provides the methodological developments to account for mixed populations in flood frequency analysis. The results of this chapter build on the physical insights developed in Chapter 2. It provides the statistical framework to perform process-driven flood frequency analysis using long-term stationary streamflow records with AR and non-AR-generated annual maxima as well as accounting for both sampling and mixing uncertainties.
- Chapter 5 describes future studies and summarizes the results of the above-listed research and concludes it.

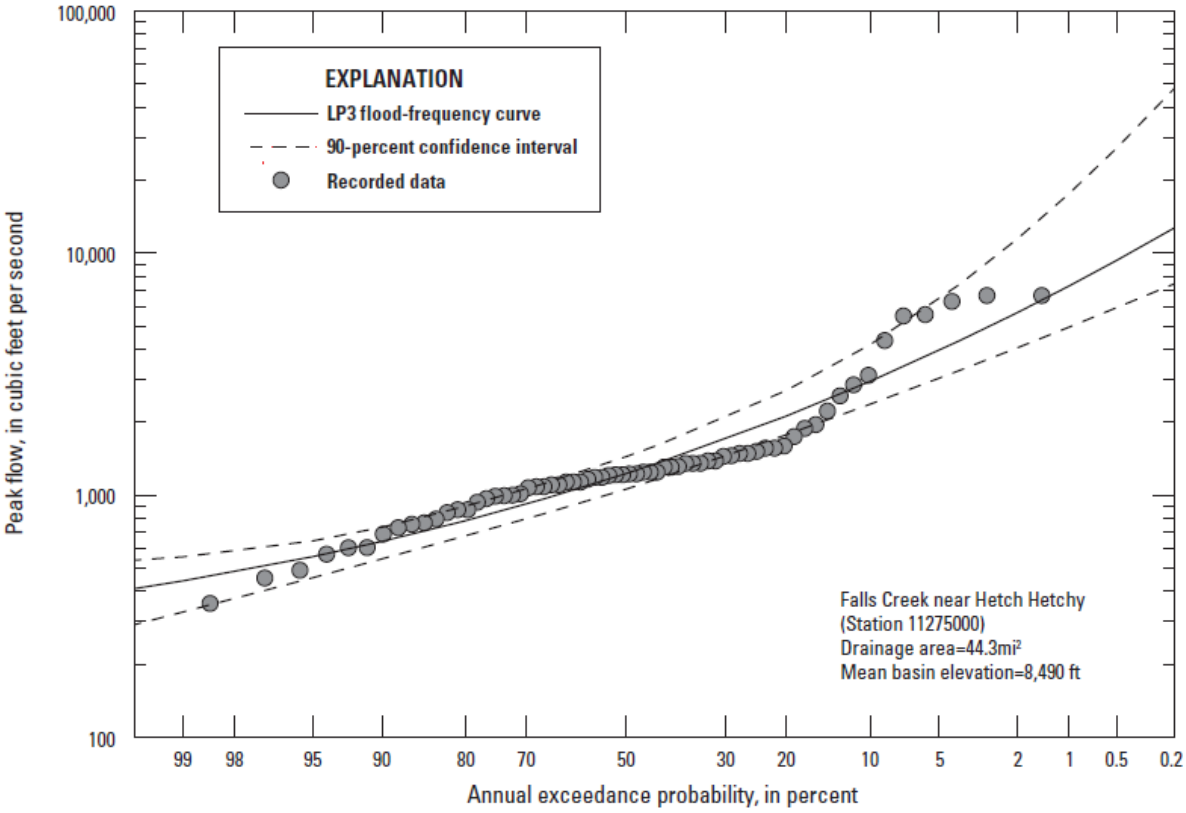

Figure 1.1: Flood frequency curve for Falls Creek near Hetch Hetchy, California (station 11275000). Modified from Gotvald et al. (2012).

Chapter 2

MIXED POPULATIONS AND ANNUAL FLOOD FREQUENCY ESTIMATES IN THE WESTERN UNITED STATES: THE ROLE OF ATMOSPHERIC RIVERS[1]

Despite the recognized importance of identifying mixtures of different flood generating mechanisms on flood frequency (B17B and B17C) (Hirschboeck, 1987; Smith et al., 2011), the topic has received little attention. In this chapter I focus on annual peak flows at long-term USGS stations throughout the western United States that are produced by atmospheric rivers (ARs).

The motivation for several of the local or regional-scale studies described in Chapter 1 has been to highlight the need for water resources managers to balance critically needed water supplies and at the same time provide flood hazard reduction in parts of the western United States frequently influenced by the variability in precipitation. In light of the historic drought in California from 2011 to 2015, the presence or absence of ARs during the cold-season precipitation months can play a large role in water resources management in California (Dettinger et al., 2011; Lavers and Villarini, 2015). ARs could potentially play a significant role in ending historic droughts along the western Pacific states (Dettinger, 2013), including ending a drought with a flood (Dettinger et al., 2011) as was the case in the cold season months of 2015/2016 throughout the state of California. Much of the focus of the existing literature has been on identifying landfalling ARs primarily along the U.S. West Coast and the contributions of ARs to annual precipitation. Recent studies have examined the precipitation associated with interior and inland penetrating ARs in the central Columbia River Basin in the Pacific Northwest, central Arizona (Rutz and Steenburgh, 2012; Rivera et al., 2014; Rutz et al., 2014; 2015) and

[1] Adapted from: 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, doi:10.1002/2016WR019064.

throughout the conterminous United States (Lavers and Villarini, 2015). While several studies have looked at AR-flooding at the watershed and regional scale along the west coast of the United States (Ralph et al., 2006; Ralph and Dettinger, 2011; 2012; Neiman et al., 2011, 2013; Guan et al., 2013) and throughout the central United States (Lavers and Villarini, 2013), no studies have been published that comprehensively examine the relationship between ARs and peak flows over the western United States.

Therefore, the research questions I will address in this chapter are:

1. What are the spatial and fractional contributions of ARs among annual peak flows throughout the western United States?
2. What is the role of ARs on the upper tail of the flood peak distribution (i.e., among the largest peaks on record)?
3. Considering mixed distributions, what are the effects of ARs on flood frequency estimates used for hazard mitigation, water resources and flood design structures?

This chapter is organized as follows. Section 2.1 summarizes the data and methodology employed to identify AR-generated peaks, stationary sites, low outliers and the log-Pearson Type III (LP3) annual exceedance probability (AEP) estimates using the new methodology of the Expected Moments Algorithm with the multiple Grubbs-Beck test (EMA-MGB). Section 2.3 presents the results, while section 2.4 summarizes the results and concludes the study.

2.1 Data and Methodology

To examine the spatial distribution and fractional contributions of ARs to annual peak flows in the western United States, I use streamgage measurements at 1375 USGS streamgaging stations with at least 30 years of annual peak flow data over the recent period of record (water

years 1979–2010; Figure 2.1a). The USGS streamgages used in this study are comprised of continuous-record streamgages that document the instantaneous annual peak flow or in some cases the maximum daily flow or partial record streamgages that document only the largest annual flow. Regardless of the streamgage, the annual peak flow data included in this study are considered the instantaneous annual peak occurring during a water year (from 1 October to 30 September). Records with at least 30 years of data from 1979 to 2010 were extended back in time, allowing for no more than one 2-year gap prior to the 1979 water year. I used 2010 as the last water year for the analyses because the reanalysis data used to identify ARs ends in 2011 (see text below). More than 1000 USGS streamgages span the second part of the twentieth century and the first part of the 21st century (Figures 2.1b and 2.1c).

A subset of 624 stationary, long-term USGS streamgages was used for flood frequency analysis. I tested the records for the presence of both abrupt and gradual changes. The former were detected using the Pettitt test (Pettitt, 1979), while I used the Mann-Kendall test (Helsel and Hirsch, 1992) to identify the presence of monotonic patterns. I set the significance level to 5%. For the flood frequency analyses, I focused on streamgage records with at least 30 years without: (i) significant abrupt or gradual changes or (ii) slowly varying changes in the period after the year of an identified step change. Moreover, I also excluded years that were qualified by the USGS as being affected by regulation, diversion, urbanization and/or land use changes (codes ‘‘6’’ and ‘‘C’’). These “stationary” records were used to examine the differences in the LP3 quantile estimates for design flows between the mixed population and the homogeneous AR-generated population of annual peak flows.

I used IVT to identify ARs. The magnitude of IVT ($kg \cdot m^{-1}s^{-1}$) was calculated using specific humidity (q, in units of $kg \cdot kg^{-1}$), zonal and meridional winds (u and v, respectively, in

$m \cdot s^{-1}$), the acceleration due to gravity (g, in $m \cdot s^{-2}$), and the pressure difference between two adjacent pressure levels (dp, in hPa) as shown in equation 2.1.

$$IVT = \sqrt{\left(\frac{1}{g}\int_{1000}^{300} qu\, dp\right)^2 + \left(\frac{1}{g}\int_{1000}^{300} qv\, dp\right)^2} \quad (2.1)$$

The variables q, u, and v were obtained from the latest version of the twentieth-century reanalysis (20CR-V2C) project (Compo et al., 2011). The 20CR-V2C project data are available from 1851 to 2011 at 6 hourly temporal resolution, at $2^{\circ} \times 2^{\circ}$ horizontal resolution, and 50 hPa vertical resolution from a pressure level of 1000 hPa to 10 hPa. Because most of the water vapor in the atmosphere is present in the lower troposphere, I use the data from 1000 to 300 hPa. While there are uncertainties in the AR identification using different reanalysis products (Brands et al., 2016; Jackson et al., 2016), I use 20CR-V2C instead of other reanalyses because it allows me to examine the connection between ARs and flooding over the western United States during the entire twentieth century rather than for only the most recent decades.

The AR identification is based on the methodology in Lavers and Villarini (2015) with minor modifications. The U.S. West Coast from 22°N to 50°N is approximated based on the longitude from 20CR-V2C. I expect differences in the IVT magnitudes at higher (e.g., 50°N) and lower latitudes (e.g., 22°N), hence the latitude range is divided into seven bins each with a latitudinal extension of 4°. In each bin, the 85th percentile of the daily IVT distribution is treated as the threshold for the AR identification. The criteria I used to identify ARs are: (i) a total length of at least 2000km as measured starting from the U.S. West Coast to the west, (ii) a maximum latitude separation of 4° (~400 km) between two successive longitude points, and (iii) a persistence criterion of three 6 hourly time steps (18 h). Maximum IVT grid points (latitude-longitude points) for each time step were defined as the AR major axis. If the total length and

width criteria were satisfied for a particular time step, the major axis was extended to the east up to 100°W as long as the IVT threshold criterion was satisfied.

The ARs identified using this algorithm were checked against the storms in Neiman et al. (2008). Seventy-six percent were successfully identified as part of the 416 ARs in Neiman et al. (2008). While the vast majority of the storms that were not identify among those found by Neiman et al. (2008) happened to make landfall outside of our focus region (e.g., British Columbia and areas further north) or did not last at least 18 h, the difference in identified AR days is likely due to the use of IVT rather than integrated water vapor (IWV). An annual peak flow discharge was considered an AR-generated flood peak if the USGS streamgage was located within 2.5° around either side of the AR major axis and the annual peak flow occurred within a 10 day window (2 days prior to and 7 days after) of an identified AR. A sensitivity test was conducted for both the temporal and spatial criteria used to identify AR-generated annual flood peaks. More specifically, I considered an additional time window of 5 days (1 day prior to and 3 days after an AR event), also examined the sensitivity of the results to radii of 2° and 3°, in addition to 2.5°: the results were insensitive to the temporal window selection across all radii examined. For each radius, the same number of AR-generated peaks were identified within the 5 and 10 day windows. Furthermore, the same number of ARs were identified for both the 2° and 2.5° radii. Yet an increase in the number of AR-generated peaks were found when the radii was increased from 2.5° to 3°. Figure 2.2a shows the results of the fraction of AR-generated annual peak flows throughout the western United States located within 2.5° around either side of the AR major axis and the annual peak flow occurred within a 10 day window (2 days prior to and 7 days after) of an identified AR. Figure 2.2b shows the increase in the fraction of AR-generated peaks between the 2.5° and 3° width within a 10 day window.

Furthermore, there is a recognized potential bias toward ascribing an annual peak flow to an AR event if additional storms passed over the basin within the selected temporal window. If, for example, two storms passed over a basin within a 10 day window and one was identified as an AR while the other one was not, and if the annual maximum occurred during those 10 days, then I attributed the annual maximum to ARs.

The at-site magnitude and frequency of floods were determined using the LP3 distribution as recommended in the B17B/B17C guidelines. The new EMA-MGB methodology using at-site skew was used to determine the 50%, 20%, 10%, 4%, 2%, 1%, 0.5%, and 0.2% AEP estimates, corresponding to the 2, 5, 10, 25, 50, 100, 200, 500 year return periods, respectively. The MGB test is a new statistical test used to identify those unusually low floods that can have undue influence on the upper tail of the distribution, the portion of the fitted distribution used for flood design structures (Cohn et al., 2013).

An additional flood frequency analysis was performed for a subset of 242 stationary sites with at least 25 years of AR-generated flood peaks. A minimum of 25 years was selected because B17B recommends sites with at least 25 years of record are used to develop regional skew values due to the sensitivity of the skew parameter to outliers. The relative percent difference (*RPD*) between the LP3-AEP estimates from the single mixed population and the estimates from the separated population of AR-generated flood peaks for each of the 242 USGS sites was calculated using the following relation:

$$RPD = \frac{\hat{Q}_{AEP,mixed} - \hat{Q}_{AEP,AR}}{\hat{Q}_{AEP,mixed}} \text{x} 100\% \qquad (2.2)$$

where $\hat{Q}_{AEP,mixed}$ and $\hat{Q}_{AEP,AR}$ are the P% AEP flows (in $m^3 \cdot s^{-1}$) for the single mixed population and AR-generated flood population, respectively.

2.2 Results

2.2.1 Spatial and Fractional Contribution of ARs in Mixed Populations

Across the western United States, I can identify six main areas in which flooding is impacted by ARs at varying degrees (Figure 2.2a). The annual peak flow records along the Pacific Northwest and northern California coast are dominated by more than 80% AR-generated flood peaks. On the other hand, the annual peak flow records at inland locations in eastern Montana, Wyoming, Utah, Colorado, and New Mexico are comprised of nearly all non-AR-generated peaks. These two distinct patterns of AR and non-AR-generated peaks are similar to the regions with the highest AR frequency and lack of inland or interior-penetrating ARs, respectively (e.g., Rutz et al., 2014; 2015; Lavers and Villarini, 2015). Yet four smaller regions have a mixture of 30–70% AR and non-AR-generated flood peaks in their period of record. They include the central Columbia River Basin in the Pacific Northwest, the Sierra Nevada, the central and southern California coast, and central Arizona (Figure 2.2a).

The USGS sites in the eastern regions of Oregon and Washington and western Idaho, within the central Columbia River Basin, have a relatively high fractional contribution of AR-generated peak flows (from 35% to nearly 90%) compared to other inland basins. ARs in these areas are able to penetrate further inland than most other U.S. West Coast landfalling ARs due to a low-elevation corridor through the Cascade Range and into the central Columbia River Basin (Rutz et al., 2015). These inland-penetrating ARs not only extend as far inland as northwestern Montana, but they have high mean durations (in excess of 24–36 h) as found in Rutz et al. (2015) and described in Parrett and Johnson (2003).

The annual peaks in the Sierra Nevada region are generated from a mixture of winter rainfall and/or rain-on-snow (AR-generated) as well as spring snowmelt (non-AR generated)

(Gotvald et al., 2012). As shown in Figure 2.2a, the fractional number of ARs is strongly linked to the steep elevation change from the central valley of northern California to the crest and eastern side of the Sierra Nevada mountain range. The peak flow records at lower-elevation sites have a higher percentage of winter rainfall AR-generated peaks, while the higher-elevation sites have fewer AR-generated peaks due to the more typical springtime snowmelt runoff. This elevation pattern and mixed population was found when developing the new regional skew values for flood frequency analysis in California (Parrett et al., 2010) and when developing regional regression equations for ungaged basins in the Sierra Nevada hydrologic region (Gotvald et al., 2012).

The central and southern California coast has a large number of AR-generated flood peaks (60–80%), similar to the high fractional contribution found along the Pacific Northwest and northern California coast. Yet at many sites south of the Monterey Bay, there is a decrease in AR-generated peaks (40–65%). Unlike sites in the Sierra Nevada, central Columbia River Basin, and central Arizona, the non-AR-generated peaks are not produced by snowmelt or summer monsoonal convective storms. Villarini (2016) found that the central and southern California coast shows an average seasonality of flooding in January and February; however, there is not a strong seasonality in this region. Annual peak flows are instead generated from rainfall distributed throughout the cold-season months. As described in Rutz et al. (2014; 2015), Oregon, Washington, and the northern coast of California have the highest values of landfalling AR frequency while a decrease in their frequency is found along the southern coast of California. With a decrease in AR-frequency along the central and southern coast of California, annual peaks are likely to result from a variety of cold-season rainfall events (e.g., AR and non-AR rainfall events). Furthermore, this variation in precipitation explains the high concentration of

sites with peak flow records with multiple zero flows and/or significantly smaller floods that deviate from the overall pattern in the data.

Central Arizona has a distinct northwest trending region with 35–65% AR-generated peaks in the flood records. As described in House and Hirschboeck (1995), Rivera et al. (2014), and Villarini (2016), the annual peaks are more typically generated from winter rainfall and/or rain-on-snow, or springtime snowmelt while peaks in southeastern Arizona are mainly generated from summer convective storms. As further shown in Neiman et al. (2013) and Rutz et al. (2014; 2015), inland penetrating ARs that originate further south along the Baja Peninsula produce heavy rainfall and high-elevation snow events from enhanced orographic precipitation as the ARs encounter the Mogollon Rim in central Arizona. While the number of inland-penetrating ARs in central Arizona is far smaller than landfalling ARs in the Sierra Nevada, a similar mixture of AR and non-AR-generated peaks is found in central Arizona as in the Sierra Nevada.

2.2.2 Identification of Low and High Outliers

Employing a consistent approach to determining the magnitude and frequency of floods throughout the United States is critical for the National Flood Insurance Program. Yet the flood hydrology throughout the western United States covers several extreme hydrologic regimes, e.g., the rainforests of the Pacific Northwest to the arid, desert regions in the Southwest. Moreover, flood frequency analysis over the western United States is complicated by peak flow records that often contain multiple zero flows and/or low outliers as referred to in B17B. While using different distributions to fit observed annual peak flows across such diverse hydrologic regions may provide better at-site estimates, the LP3 distribution has been the recommended distribution in the Federal Guidelines of B17B and its new updates in B17C. Much of the work by the

Hydrologic Frequency Analysis Work Group (HFAWG) under the Advisory Committee on Water Data (ACWII) to update Bulletin 17B focused on addressing the problematic low outliers and potentially influential low floods (PILFs) in the western United States that reflect different physical processes than those from which larger floods are generated (Cohn et al., 2013). Furthermore, PILFs may not always deviate from the overall pattern of the data, yet may in turn have high leverage that may exert a large influence on the upper right-hand tail, the tail of the distribution comprised of large floods for which we are concerned (Lamontagne et al., 2016). While the use of the multiple Grubbs-Beck test to identify PILFs represents a new key addition to the updates in B17C, this new test has also already been used in six catchments in eastern Australia (Rahman et al., 2014).

Using the new generalized MGB test to identify PILFs, Figure 2.3a shows the ratio of PILFs in the mixed population of peak flows for the 624 stationary USGS sites in the western United States. Figure 2.3b shows the ratio of PILFs identified in the 242 stationary USGS streamgage records with at least 25 years of AR-generated peaks. The central and southern coast of California has the highest ratio of PILFs among both the mixed population and the homogeneous AR-generated peak population. Up to 50% of the annual peaks were identified as PILFs using the MGB test in both populations. As described earlier, this unique pattern of PILFs is most likely influenced by the variation in precipitation from cold-season rainfall events (e.g., AR and non-AR rainfall events). Furthermore Idaho, Montana, and Colorado also show a large number of sites with a similar ratio of PILFs distributed throughout each state in the mixed population of annual peaks. In this region, Villarini (2016) found the strongest seasonality of annual peaks flows throughout the conterminous United States; yet several sites within this

region also showed a weaker seasonality. Although not explored in this study, PILFs in these regions are most likely the result of the mixed populations in the non-AR-generated peaks.

Similar to the influence of PILFs on the upper tail of the distribution, flood series may also have high outliers, representing extreme flood events that deviate from the pattern in the data. They, too, may have a substantial influence on the overall LP3 fit. We examined the contribution of ARs among the largest peaks on record at the 624 stationary sites throughout the western United States. Figure 2.4a shows the location of USGS sites whose largest annual peak is generated by an AR, while Figures 2.4b and 2.4c show the number of AR-generated peaks among the top-5 and top-10 flood peaks, respectively. Not surprisingly, the largest annual peaks along the Pacific Coast of the western United States are nearly all AR generated, while a mixture of AR and non-AR-generated peaks are found throughout the study region east of the Sierra Nevada and Cascade Range. Compared to the central United States, ARs are responsible for more and bigger flood events in the western United States (Lavers and Villarini, 2013). In particular, the highest average monthly precipitation values of more than 250 mm/month were found in the Pacific Northwest due to enhanced orographic precipitation (Lavers and Villarini, 2015). Some of the largest regional flood events along the west coast of the United States have been caused by warm, high-elevation rain and rain-on-snow events producing the November 2006 floods in Washington (Neiman et al., 2011), the February 1996 floods in Oregon (Copper, 2005), and the 1997 New Year's Flood in central and northern California (Hammond and Harmon, 1998).

Furthermore, nearly all the top-5 and top-10 peaks on record in the central Columbia River Basin, the Sierra Nevada, central and southern California coast, and central Arizona are AR generated (Figures 2.4b and 2.4c), unlike the fraction of AR-generated peaks (30–70%)

throughout the period of record in these regions (Figure 2.2a). Moreover, a small number of sites in far northern and southern Nevada have a relatively large number of AR-generated peaks among the top peaks despite being located east of the Sierra Nevada range. Rutz et al. (2015) found a similar spatial pattern of inland penetrating ARs into Nevada from landfalling ARs off the coast of California. They found that inland penetrating ARs can traverse the lower elevations of the Sierra Nevada range north of Lake Tahoe or less frequently over the southern portion of the high Sierra. Overall, the largest flows on record in these regions and along the west coast of the United States are a result of cold-season AR-generated rainfall peaks and not a mixed population.

Figures 2.5a and 2.5b show the seasonal occurrence of AR and non-AR-generated peaks at the Merced River at Pohono Bridge near Yosemite, California (station 11266500), a high-elevation site in the Sierra Nevada, and in central Arizona at the Salt River near Chrysotile, Arizona (station 09497500), respectively. The mixed population seasonality plots show distinct bimodal distributions of annual flood peaks occurrences at these sites and highlight the substantial difference in the magnitude of floods among the seasons. The largest flows at these sites are AR-generated cold-season peaks. The non-AR peaks at station 11266500 were likely caused by springtime snowmelt, while the summertime non-AR peaks at station 09497500 were likely caused by summer monsoon convective storms.

2.2.3 Comparison of Magnitude and Frequency Estimates between Mixed and AR-Generated Flood Populations

Figure 2.6 shows the *RPD* between the quantile estimates for different AEPs. The *RPD*s, which highlight the role of ARs in the estimation of different discharge quantiles, are generally zero along the Pacific Northwest and northern California as expected because more than 85% of

the peak flow records are comprised of AR-generated flows (Figure 2.2a). Negative *RPD*s are commonly found across the remaining sites for all AEPs (the blue hues in Figure 2.6). This indicates that the quantile estimates from the AR-generated population of flood peaks are higher than those estimated from the mixed population.

Two regions show a mixture of *RPD*s. The central and southern California coast and central Arizona regions have a high concentration of both positive and negative *RPD*s across the smaller AEPs (1% to 0.2% AEPs, corresponding to the 100–500 year return intervals, respectively). As described in section 2.2.2, the largest peaks along the central and southern California coast are AR-dominated (Figure 2.4) which should generally lead to negative *RPD*s. Yet the persistent presence of PILFs among both the mixed and AR-generated populations and the high variability in annual peak flows are most likely contributing to the inconsistent pattern of positive and negative *RPD*s.

In central Arizona, although AR-generated peaks comprise many of the top-10 largest peaks, the largest mixture of both positive and negative *RPD*s is found in this region. Similar to the high-elevation sites in the Sierra Nevada, extreme rainfall, heavy snowfall events, and/or a combination of rain-on-snow during the winter cold-season months have led to large peak flows such as those that caused extreme flooding in Ari- zona during the winter of 1993 (House and Hirschboeck, 1995; Neiman et al., 2013; Rivera et al., 2014). Peak flows in this region are also generated by convective summertime thunderstorms (House and Hirschboeck, 1995; Villarini, 2016). As described in section 2.2.2, central Arizona has a high variability and a lower strength in the seasonality of flooding (Villarini, 2016). Therefore, the magnitude of cold-season AR-generated peaks or springtime snowmelt or summertime convective non-AR-generated runoffs provides a highly inconsistent pattern of *RPD*s in central Arizona.

Figures 2.7a and 2.7c show examples of the LP3 flood frequency curves used to calculate the *RPD* for the high-elevation mixed population Sierra Nevada site in Figure 2.5a (station11266500) and the central Arizona site in Figure 2.5b (station 09497500). Not only are the majority of the top-10 annual peak flows generated by ARs, but the majority of the right-hand tail of the distribution is dominated by these flows. Furthermore, the empirical cumulative distribution function (ecdf) plot of the observed annual peak flows at both sites show a poor fit in the right-hand tail compared to the fitted LP3 curves. Yet Figures 2.7b and 2.7d show the improved flood frequency fit among the AR-generated population.

2.3 Discussion and Conclusions

This study examined the role of ARs on the mixed population of annual peak flows in the western United States, and their impact on the magnitude and frequency estimates used for the design of flood structures. Flood frequency analysis in this region often contains annual peak flows generated from distinctly different flood generating mechanisms representing different hydrometeorologic events. Among the different flood agents, ARs are responsible for large, regional-scale floods. Extreme fluctuations in annual peak flows have recently been seen in the span of the past five years throughout the western United States. As described in Chapter 1, this is highlighted by the historic 2011–2015 drought in California, followed by the current onslaught of strong winter storms, producing a combination of heavy rain and snowfall, flooding, and mudslides across California this 2015/2016 winter season.

I calculated the fractional contribution of AR-generated annual peak flows at 1375 long-term USGS streamgage stations having at least 30 years of data over the most recent water years from 1979 to 2010. More than 80% (and 100% at several locations) of the annual peaks in the

Pacific Northwest and along the northern California coast are generated by ARs; on the other hand, eastern Montana, Wyoming, Utah, Colorado, and New Mexico have peaks dominated by events that are not from ARs. The regions of the central Columbia River Basin in the Pacific Northwest, the Sierra Nevada, the central and southern California coast, and central Arizona all show a mixture of 30–70% AR-generated flood peaks.

Low and high outliers may have a substantial influence on the overall LP3 fit. I used the new MGB to identify PILFs (low outliers) in the annual peak flow series for "stationary" sites (sites with at least 30 years of data without abrupt or gradual changes), throughout the western United States. I found sites with a large percentage of PILFs in the mixed population of annual peaks scattered throughout the study region. However, the central and southern coast of California had the highest concentration of sites with PILFs among both the single mixed population and AR-generated peak population. Moreover, high percentages of PILFs were found distributed throughout the states of Idaho, Montana, and Colorado.

To determine the hydrometeorologic mechanism responsible for the largest peaks on record (potential high outliers), I computed the fractional contribution of ARs among the largest, top-5 and top-10 flood peaks on record. As expected, nearly all of the largest peaks on record along the U.S. West Coast (west of the Sierra Nevada and Cascade Ranges) are generated by landfalling ARs. The central Columbia River Basin, northern and southern Nevada, and central Arizona also have a large fraction of AR-generated peaks from inland and interior-penetrating ARs. The new EMA-MGB methodology was used to determine the 50%, 20%, 10%, 4%, 2%, 1%, 0.5%, and 0.2% AEPs for both the single mixed population and the homogenous AR-generated flood peak population (corresponding to the 2, 5, 10, 25, 50, 100, 200, and 500 year recurrence intervals, respectively). Near-zero *RPD*s were found in the frequency estimates

between the two populations along the Pacific Northwest and northern California due to the high fraction of AR-generated annual peaks in the complete period of record. Negative *RPD*s were commonly found across the remaining sites indicating that the quantile estimates from the AR-generated population are higher than those estimates from the single mixed population. However, a combination of both positive (i.e., AEP estimates are larger for the single mixed population than AR-generated population) and negative *RPD*s for the smaller AEPs (1%–0.2%) were found along the central and southern coast of California and central Arizona.

Reliable estimates are needed to protect life and property, but are also needed for effective planning and management of land and water resources. A fundamental assumption that the observed annual peak flow data are a "representative time sample of random homogeneous events" (B17B/B17C) may not properly apply to those regions in the central Columbia River Basin in the Pacific Northwest, the Sierra Nevada, the central and southern California coast, and central Arizona that all show a mixture of 30–70% AR-generated flood peaks. Quite often the fitted annual peaks to the overall LP3 distribution show sharp breaks in the slope or a curve that reverses direction in these regions. This is indicative of streamflow caused by distinctly different flood generating mechanisms and a mixed population analysis is recommended. As described in Chapter 1, both B17B and B17C recommend developing separate frequency curves when the hydrometeorologic mechanisms that generated the annual peak flows can be separated into distinct populations such as those produced from ARs and non-ARs. More accurate flood frequency estimates can be determined by separating the peak flows using a process-driven mixed population approach and will be explored in Chapter 4.

Much of the recent research related to ARs over the western United States has focused on the contribution of ARs to extreme precipitation, with some studies on flooding events at the

watershed or over a limited regional scale. As shown in this chapter, ARs throughout portions of the western United States contribute to some of the largest annual peak flows on record. Yet at higher-elevation sites, ARs contribute to heavy snowfall accumulations that do not directly produce cold-season AR-generated flood peaks. This may potentially lead to lower estimates of the fractional contribution of AR-generated peaks among the high-elevation mixed population sites because the criteria used in this study do not account for AR-generated snowfall that melts later in the season. Further studies are also needed to evaluate the contribution of ARs to annual precipitation and how that relates to other flow durations for water resource management in the western United States.

2.4 Figures Chapter 2

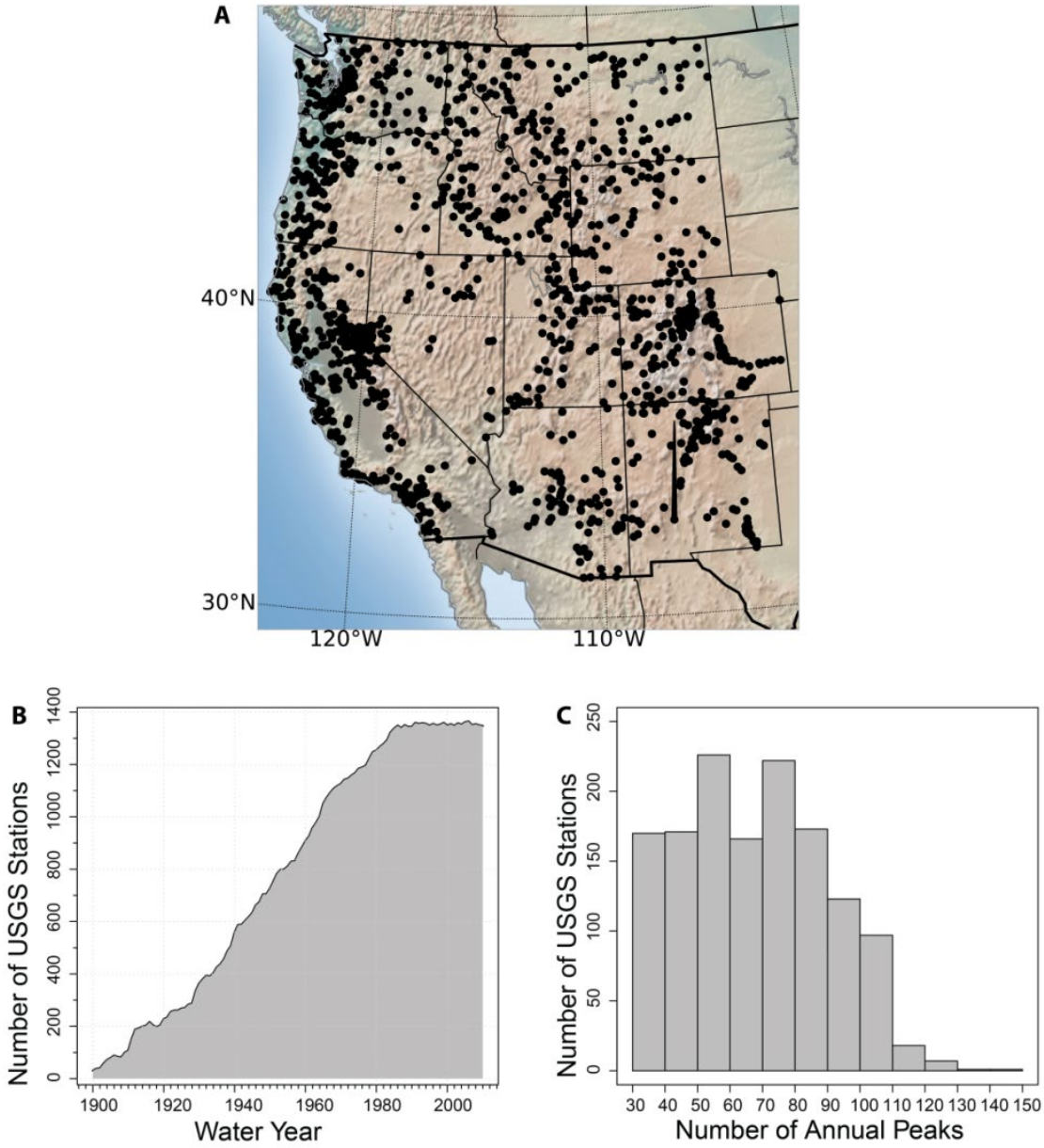

Figure 2.1: (A) Location of 1375 long-term USGS streamgages with at least 30 years of continuous record from water years 1979–2010; (B) span of continuous record (in water years) for the USGS streamgages; (C) histogram of the record lengths of the 1375 stations included in the study.

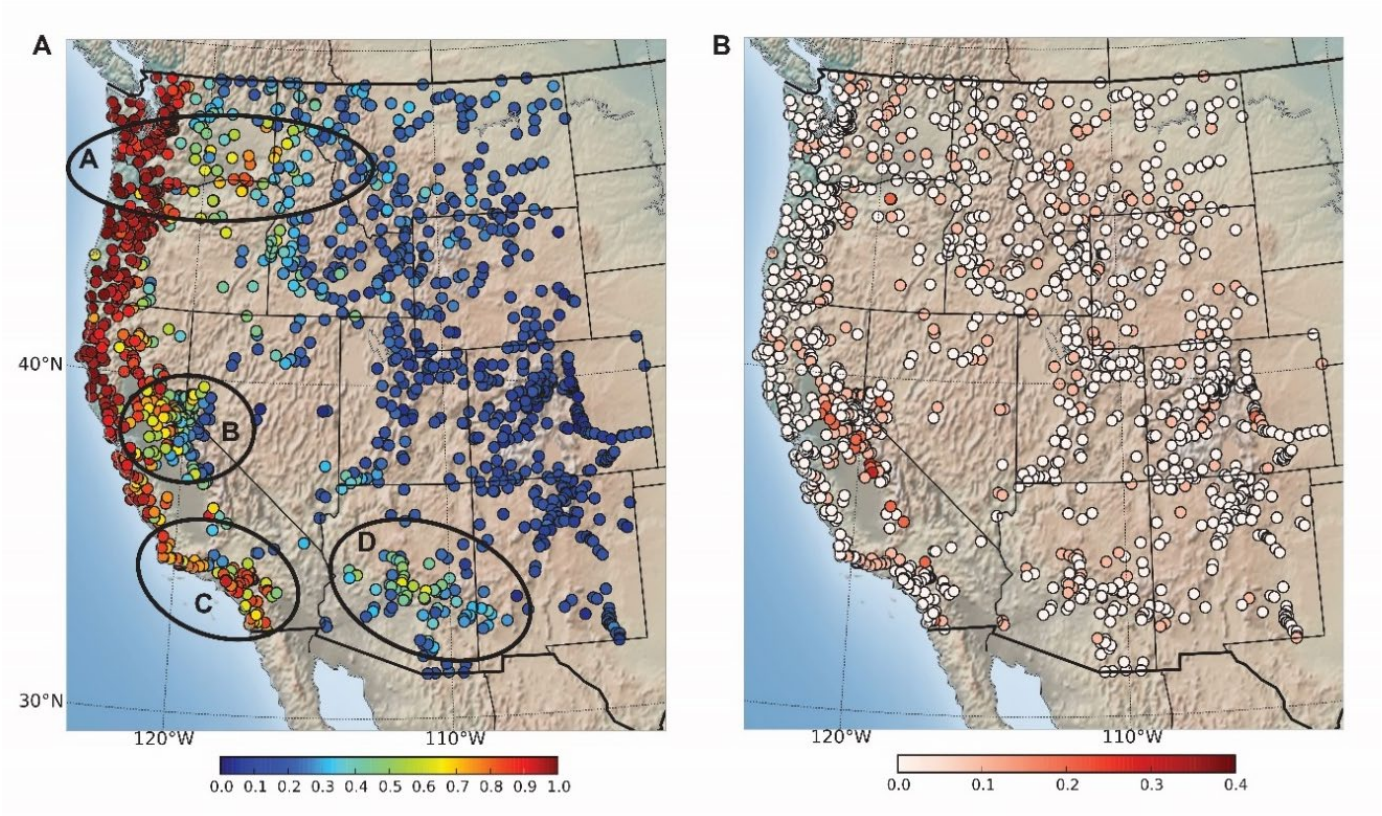

Figure 2.2: (A) The fractional contribution of AR-generated annual peak flows at 1375 long-term USGS streamgaging records throughout the western United States. Location of four smaller regions with a mixture of 30–70% AR and non-AR-generated flood peaks in their period of record are highlighted. Region A is the central Columbia River Basin in the Pacific Northwest, Region B is the Sierra Nevada, Region C is the central and southern California coast, and Region D is central Arizona; (B) map showing the increase in the fraction of AR- generated peaks between the 2.5° and 3° radii widths within a 10-day window.

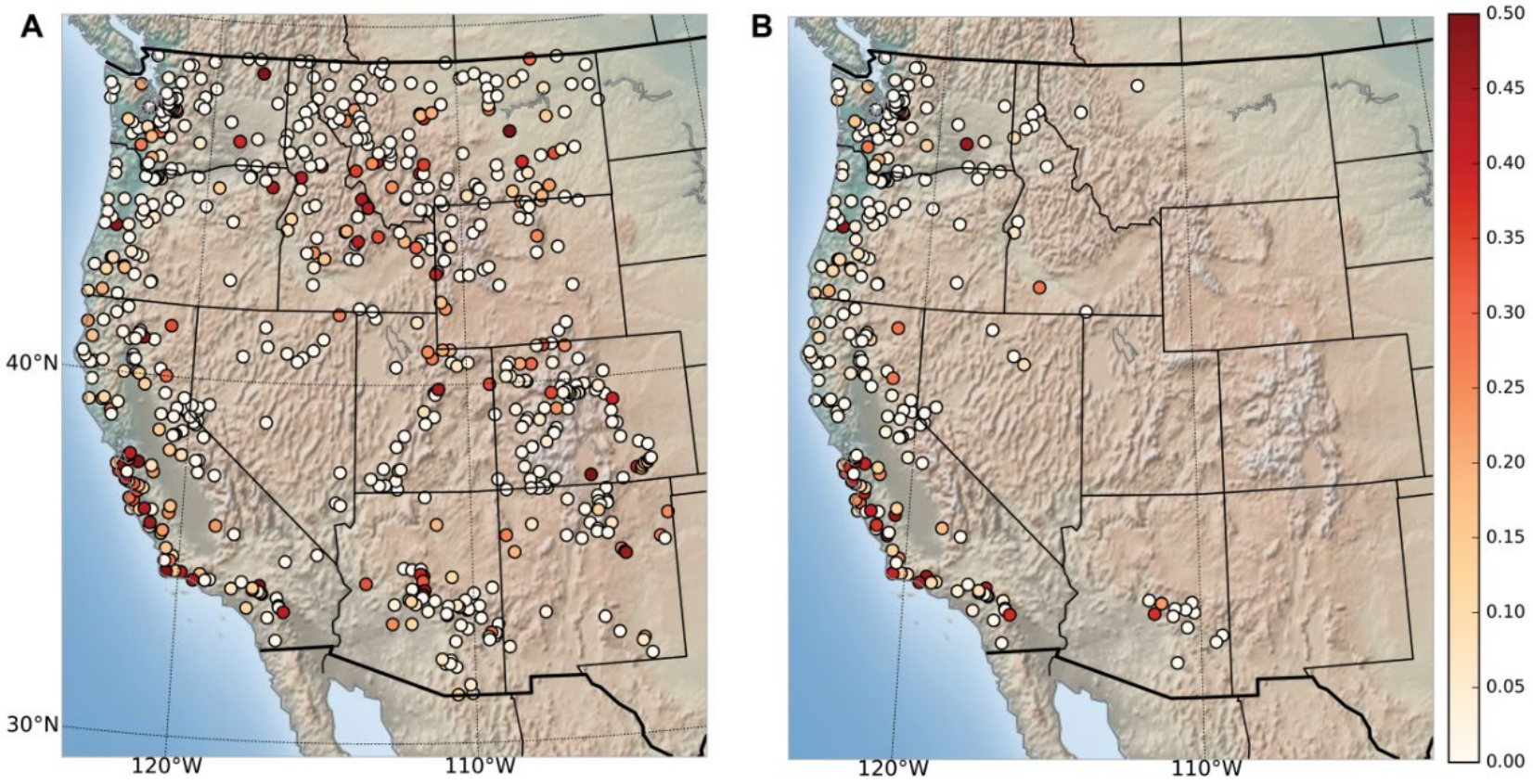

Figure 2.3: Ratio of potentially influential low floods (PILFs) in (A) complete period of record (single mixed population); (B) the population of AR-generated annual flood peaks with at least 25 years of peak flow record.

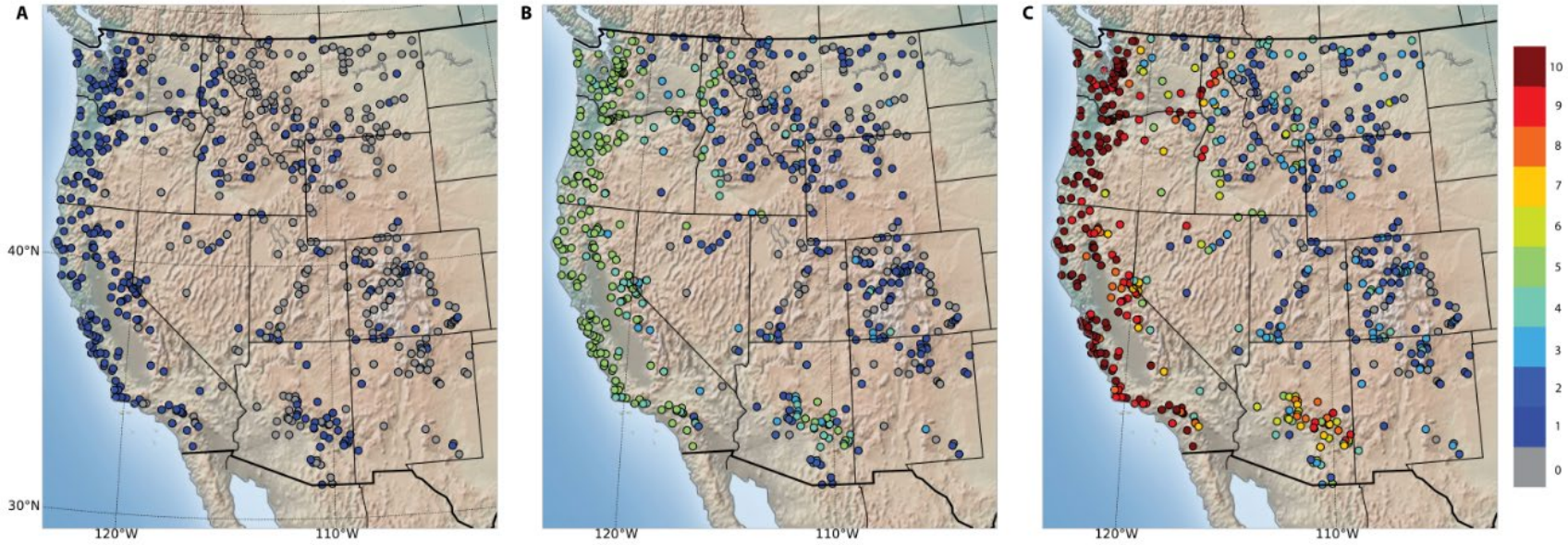

Figure 2.4: Maps showing the number of AR-generated flood peaks among (A) the largest flood peak, (B) the top-5, and (C) top-10 flood peaks during their period of record.

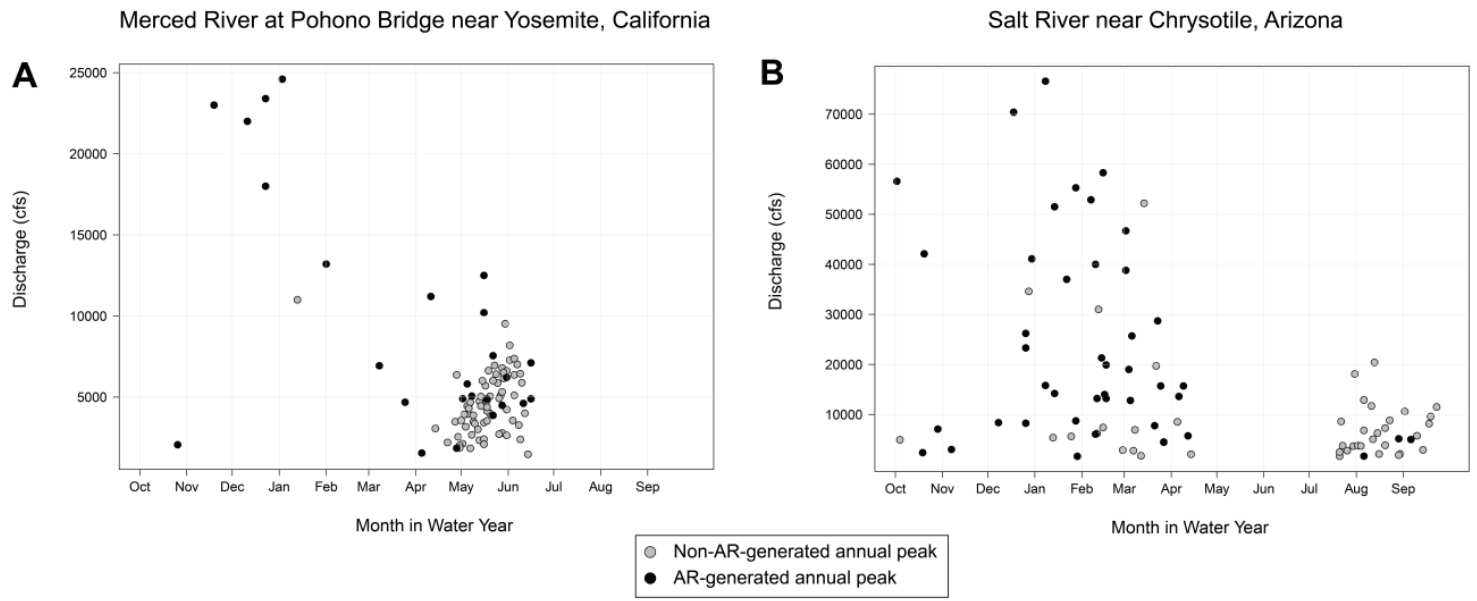

Figure 2.5: The seasonal occurrence of AR and non-AR-generated annual peak flows at (A) the Merced River at Pohono Bridge near Yosemite, California (station 11266500) and (B) at the Salt River near Chrysotile, Arizona (station 09497500).

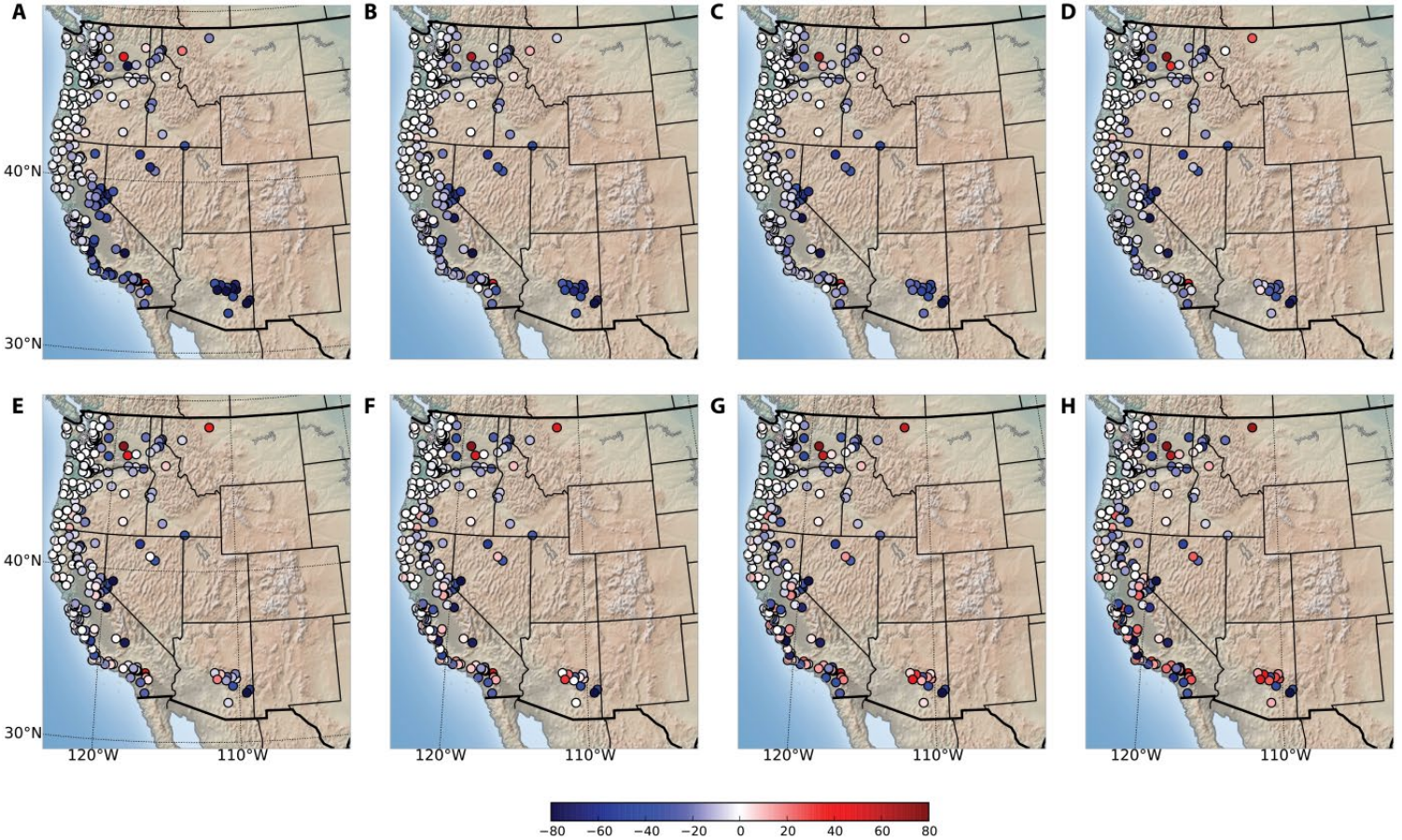

Figure 2.6: Relative percent differences (RPD) between the LP3 AEP estimates from the single mixed and AR-generated population of annual peak flows at 242 stationary, long-term USGS sites. The RPDs for the (A) 50%, (B) 20%, (C) 10%, (D) 4%, (E) 2%, (F) 1%, (G) 0.5%, and (H) 0.2% AEPs. A positive (negative) RPD indicates a higher LP3 estimate from the single mixed (AR-generated) peak population, respectively.

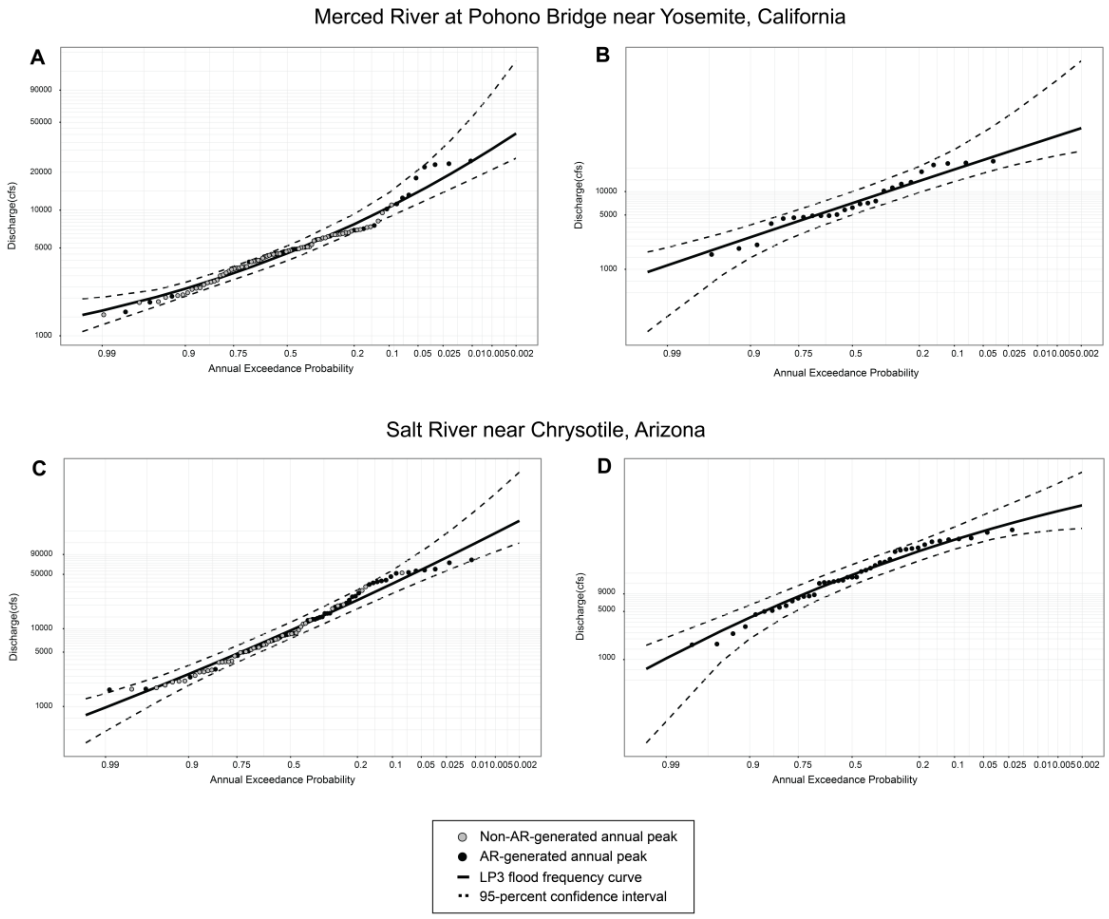

Figure 2.7: The LP3 flood frequency curves for the (left) single mixed population and (right) AR-generated population of annual peak flows at the Merced River (top row) at Pohono Bridge near Yosemite, California (station 11266500) and (bottom row) at the Salt River near Chrysotile, Arizona (station 09497500).

Chapter 3

CONTRIBUTION OF EASTERN NORTH PACIFIC TROPICAL CYCLONES AND THEIR REMNANTS ON FLOODING IN THE WESTERN UNITED STATES[2]

Flood losses due to tropical cyclones (TCs) are responsible for numerous fatalities and significant economic losses estimated in the billions of dollars (e.g., Czajkowski et al., 2013, 2017; Rappaport, 2014). In the United States, the majority of TCs make landfall along the eastern Seaboard, the Florida Peninsula, and the Gulf of Mexico, while far fewer eastern North Pacific TCs make landfall in the western United States. Yet those storms and their remnants, originating in the eastern North Pacific and the Gulf of California, do impact the Southwestern United States (e.g., Corbosiero et al., 2009; Ritchie et al., 2011; Wood and Ritchie, 2013; Khouakhi et al., 2017).

As described in Chapters 1 and 2, flooding throughout the western United States is primarily driven by mixtures of different flood generating mechanisms including winter rainfall and rain-on-snow events associated with atmospheric rivers (e.g., Ralph et al., 2006; Barth et al., 2017), springtime snowmelt (e.g., Thomas et al., 1997; Gotvald et al., 2012), and the summertime North American monsoon (e.g., House and Hirschboeck, 1995; Corbosiero et al., 2009; Ritchie et al., 2011; Paretti et al., 2014) (consult Villarini (2016) for an extensive discussion on the seasonality of flooding across the United States). However throughout the 20th century TCs and their remnants (herein referred to as TC-events) have occasionally caused loss of life and substantial economic losses. For example in southern California and Arizona, moisture from TC-events interacted with midlatitude disturbances such as troughs or cutoff lows and produced locally

[2] Adapted from: Barth, N.A., Villarini, G., and White, K.., (2018), Contribution of tropical cyclones and their remnants on flooding in the Western United States, International Journal of Climatology; 1-6.

heavy rainfall enhanced by the orography in the region (e.g., Smith, 1986; Farfan et al., 2001; Corbosiero et al., 2009; Ritchie et al., 2011; Wood et al., 2013). Most economic losses and fatalities due to TCs in the United States are found near the coast; yet recent studies have shown similar destructive floods occur in inland riverine basins far from the coast as moisture from landfalling TCs produce substantial rainfall in those regions (e.g., Villarini et al., 2014; Czajkowski et al., 2017; Aryal et al. 2018). In the Southwestern United States, heavy rain and devastating flash floods in inland basins were caused by Tropical Storm Kathleen (1976) in southern California and by the entrainment of moisture from the remnants of Tropical Storm Octave (1983) in southeastern Arizona (Smith, 1986; Hjalmarson, 1989; Roeske et al., 1989). While the frequency of widespread flooding related to TC-events is far less in the western United States compared to much of the central portion of the United States and the eastern seaboard, some studies have documented the effects of TC-generated flooding at the basin scale in southern California and southeastern Arizona (e.g., Hirschboeck, 1988, 1991, 2009; Hjalmarson, 1989; Roeske et al., 1989; Webb and Betancourt, 1992; NOAA, 2010).

Various studies have evaluated the climatology of TC-events and their corresponding contribution to heavy rainfall in the southwestern United States (Fors, 1977; Smith, 1986; Hirschboeck, 1991; Webb and Betancourt, 1992; Farfan et al., 2001; Corbosiero et al., 2009; Ritchie et al., 2011; Wood and Ritchie, 2013; Khouakhi et al., 2017), yet little attention has been given to evaluate their impact in terms of flooding, especially at the regional scale. Corbosiero et al. (2009) evaluated TC-generated rainfall from TCs that moved 25°N while still at tropical storm strength (maximum surface winds greater than 17 ms^{-1}). Ritchie et al. (2011) defined TC and TC remnant-generated rainfall based on composite categories comprised of rainfall swaths and synoptic circulation patterns.

Heavy rainfall is an important component of flood generation, however determining its contribution to flooding in the hydrologically diverse western United States requires an approach that captures the complexity of hydrometeorologic interactions commonly found in this region. Villarini et al. (2014a) attributed annual maximum flows to streamgages that were located within 500 km from the center of circulation of a storm and occurred within a 10-day window (2 days prior to and 7 days after) from the passage of the storm, consistent with other studies (e.g., Prat and Nelson, 2013; Khouakhi et al., 2017). As described in Ritchie et al. (2011), TCs and the large quantities of tropical moisture from remnant TCs can produce extreme amounts of orographically-enhanced precipitation as these systems interact with the mountain ranges throughout the western United States; moreover, the moisture from TC remnants can be advected far into the central and eastern United States, given favorable synoptic patterns. Still despite the potential impacts that TC-events can have, a comprehensive evaluation of the role that these storms play in the hydrology of the western United States is still lacking. Here I examine the role played by TC-events on flood events throughout the western United States using 103 storms that occurred between 1958 and 2010.

3.1 Data and Methodology

3.1.1 Streamgage Data

I use 1,429 long-term U.S. Geological Survey (USGS) streamgages in the western United States that are located west of the 100th meridian because the moisture contributing to precipitation from eastern North Pacific TCs and their remnants can be advected well into the central United States (Corbosiero et al., 2009; Ritchie et al., 2011). These qualifying gages have at least 30 years (a year is considered to be "complete" if it has at least 330 daily observations) of

daily discharge observations during the 1958-2010 water years (a water year is defined from 1 October to 30 September). Even though I only considered sites with at least 30 years of record, the majority of them (1,163 sites representing over 80% of the streamgages) have between 40 and 53 years of data, allowing me to frame these results in a longer-term context.

To examine the impacts of TC-events on the flood hydrology in the western United States, I use two approaches, the annual maximum (AM) and peak-over-threshold (POT) both based on daily average discharge values. With the AM approach, I first identify the largest daily average discharge value during a water year. Note that the AM is not based on the instantaneous annual peak discharge data which are also available at many USGS streamgages in the western United States. When using the POT approach, I identify all the daily discharge values exceeding a given threshold level (*S*). Here I set *S* so that, on average, I have two events per year (POT2; e.g., I set a threshold level to identify 60 events if the record length is 30 years). I also performed a sensitivity analysis to determine if the results changed if the number of events per year were increased to, on average, four events per year (POT4). Furthermore, to avoid double-counting among the qualifying POT flows, I only allow one event in an "n-day" window, which represents the amount of time required to produce a peak flow in the basin (n-day window < $|5\ days + log_{10}(A)|$; where A is the contributing drainage area (in mi^2); Lang et al. 1999).

An AM or a POT flood peak is attributed to a TC-event if it meets two criteria. First, the flow is within the window of a TC-event plus an additional five days (defined as TC-days). The additional five days is used to capture potential precipitation induced by TC remnants and coincides with the time needed to generate a flood within the n-day window. The second condition is that the streamgage is located within the TC-event precipitation swaths published in Corbosiero et al. (2009), Ritchie et al. (2011), and the National Oceanic and Atmospheric

Administration (NOAA) Weather Prediction Center (WPC) (http://www.wpc.ncep.noaa.gov/tropical/rain/tcpointofentry.html) (herein referred to as NOAA-WPC).

3.1.2 Tropical Cyclone and Precipitation Data

I consider 103 storms that occurred between 1958 and 2010 to evaluate the contribution of TC-events to annual maximum flows throughout the western United States. These storms represent a combination of TC-events and their attributed rainfall evaluated by Corbosiero et al. (2009), Ritchie et al. (2011), and NOAA-WPC. The TC data are from HURDAT2 (Landsea and Franklin, 2013), which, among other features, provides the latitude and longitude of the center of circulation of any recorded TC in the North Atlantic and eastern North Pacific every six hours. Table 3.1 provides the HURDAT2 serial number, TC name, and TC storm window for each of the 103 storms used in this analysis. Furthermore it provides the precipitation references used to determine if AM and POT flows are attributed to a TC-event.

I use precipitation data from the Climate Prediction Center (CPC) Unified Gauge-Based Analysis of Daily Precipitation (CPC UD United Precipitation data: https://www.esrl.noaa.gov/psd/data/gridded/data.unified.daily.conus.html), which represents the daily accumulated precipitation from 12UTC to 12UTC. The cumulative precipitation for each TC storm event (TC-days) is calculated to further visually verify that the TC-event produced precipitation near the USGS streamgages. Figure 3.1 shows Hurricane Octave and the cumulative precipitation (in inches) across the western United States during the corresponding TC-days. The location of USGS streamgages with annual maxima (AM) (a) and peaks-over-threshold (POT) flows (b-c) attributed to TC-events and those gages that did not have TC-event

attributed flows are also shown. POT threshold level (*S*) corresponding to two (POT2) and four (POT4) are shown in panels b-c, respectively. Similar figures for the characteristics of the 103 TC-events considered in this study are shown in supplementary figures in Barth et al. (2018a).

3.2 Results

The largest fractional contributions of TC-events among maximum flows in the western United States are primarily located in southern California, the "Four Corners" region (consisting of Arizona, Colorado, New Mexico, and Utah), and Texas (Figure 3.2). Of the 1,429 USGS gaging stations that are located west of the 100th meridian, 32% of the gages have at least one AM flow that is attributed to a TC-event, 28% among the POT2, and 40% among the POT4 flows. By lowering the threshold level in the POT4 approach, the spatial footprint of TC-generated flows is increased (Figure 2C). While in many cases more floods are attributed to TC-events by lowering *S*, the overall percentage of TC-related POT4 flood events is smaller than what is observed for the POT2 case due to the increased record length and therefore an overall larger number of events considered. Based on these sensitivity results, I conclude that while there is an increase in the spatial footprint of TC-event-generated flows, the POT2 approach captures the overall pattern of the contribution of TC-events on flooding in the western United States.

The spatial extent of TC-event-generated AM and POT flows closely corresponds to the spatial distribution of extreme precipitation attributed to TCs (e.g., Corbosiero et al., 2009; Ritchie et al., 2011; Khouakhi et al., 2017; NOAA-WPC). Southern California, Arizona, southernmost Nevada and Utah, southern and western New Mexico, central Colorado, and Texas have the highest fractional contributions of TC-event-generated annual maxima (~5-14%), in agreement with the TC-event rainfall patterns described in Ritchie et al. (2011) [e.g., north

recurving (Hurricane Lester), south recurving (Hurricane Kenna), and north/northwest movement (Hurricane Kathleen)]. Furthermore streamgages within Arizona collectively have some of the highest fractional contributions to TC-event-generated flows due to the common entry point of eastern North Pacific TCs and their remnants in the southwestern United States [e.g., Hurricane Octave (Figure 3.1) and Hurricane Norma] and their interaction with the complex topography that can locally produce heavy rainfall enhanced by orography in the region (Ritchie, et al., 2011; NOAA-WPC). Depending on the TC track, the existence of favorable synoptic conditions and sufficient moisture associated with the TC remnant, inland streamgages as far north as Washington, Idaho, and Montana have at least one AM and/or POT associated with TC-events (Figure 3.2).

Up to this point I have focused on determining TC-related AM and POT flows regardless of where they occur within the flood peak distribution. Figure 3.3, instead, focuses on the contribution of TC-events among the top-10 largest flows for the AM (A) and POT2 (B) flood series. In general, there is no overall significant regional contribution of TC-related flows among the largest flows on record in the western United States. Six percent of the sites in this study had at least one of the flows among the top-10 floods on record using both the AM and POT methods. These sites are primarily located in southern California, the Four Corners region, and Texas. While TC-events are not a large contributor to the upper tail of the flood distribution, there is a tendency for streamgages in a southwest-northeast recurving path and a southern recurving path to have at least one TC-event-generated flow among the top-10 flows. These swaths coincide with TC-event paths described above.

These results provide empirical evidence of the limited overall role that TC-events play among the largest flows on record in the western United States; however they highlight smaller

regions containing a handful of streamgages for which TCs and their remnants have a larger impact. For example, in southern California and southern Nevada, precipitation from the remnants of Hurricane Kathleen produced extensive flash flooding that generated flows among the top-10 largest AMs and POT2s. Moreover, the remnants of Tropical Storm Octave (Figure 3.1) produced major floods throughout several large basins in central and southeastern Arizona due to the prolonged period of rainfall as it interacted with a high-altitude low pressure trough from September 27 to October 3 (Roeske et al., 1989). The prolonged rainfall produced the largest instantaneous annual peak discharge on record that was more than 2.2 times the next largest peak at the USGS streamgage 09482500 (Santa Cruz River at Tucson, AZ) (e.g., Roeske et al., 1989).

3.3 Summary

Flood hydrology throughout the western United States is complicated by a mixture of flood generating mechanisms. Recent studies have focused on identifying the spatial coverage of heavy rainfall associated with TC-events throughout the Southwestern United States from eastern North Pacific storms. Yet little is known about the larger spatial contribution of these storms among extreme floods over the western United States. I examined the spatial and fractional contributions of 103 TC-events among annual maximum streamflow at 1,429 long-term USGS streamgages with at least 30 years of daily discharge observations during the 1958-2010 water years. The main findings of this Chapter can be summarized as follows:

(1) Nearly 32% of the gages have at least one AM flow that is attributed to a TC-event, 28% among the POT2, and 40% among the POT4 flows. Southern California, Arizona, southernmost Nevada and Utah, southern and western New Mexico, central Colorado,

and Texas have the highest fractional contributions of TC-event-generated annual maximums flows (~5-14%), consistent with the TC-event tracks and rainfall patterns found in previous studies.

(2) There is no overall significant regional contribution of TC-event-generated flows among the top-10 maximum flows on record. Sites that had at least one of the AM and/or POT2 flows among the top-10 floods on record are primarily located in southern California, the Four Corners region, and Texas.

(3) The results of this study not only identify the spatial extent of TC-generated maximum flows throughout the western United States, but they also highlight the contribution of TC-remnants among these flows as they interact with the complex topography to generate, in some limited cases, the largest flows on record. The results of this study further enhance the understanding of the complex process-driven flood-generating mechanisms found in the western United States and the role of TC-events on the overall flood peak distribution.

3.4 Figures Chapter 3

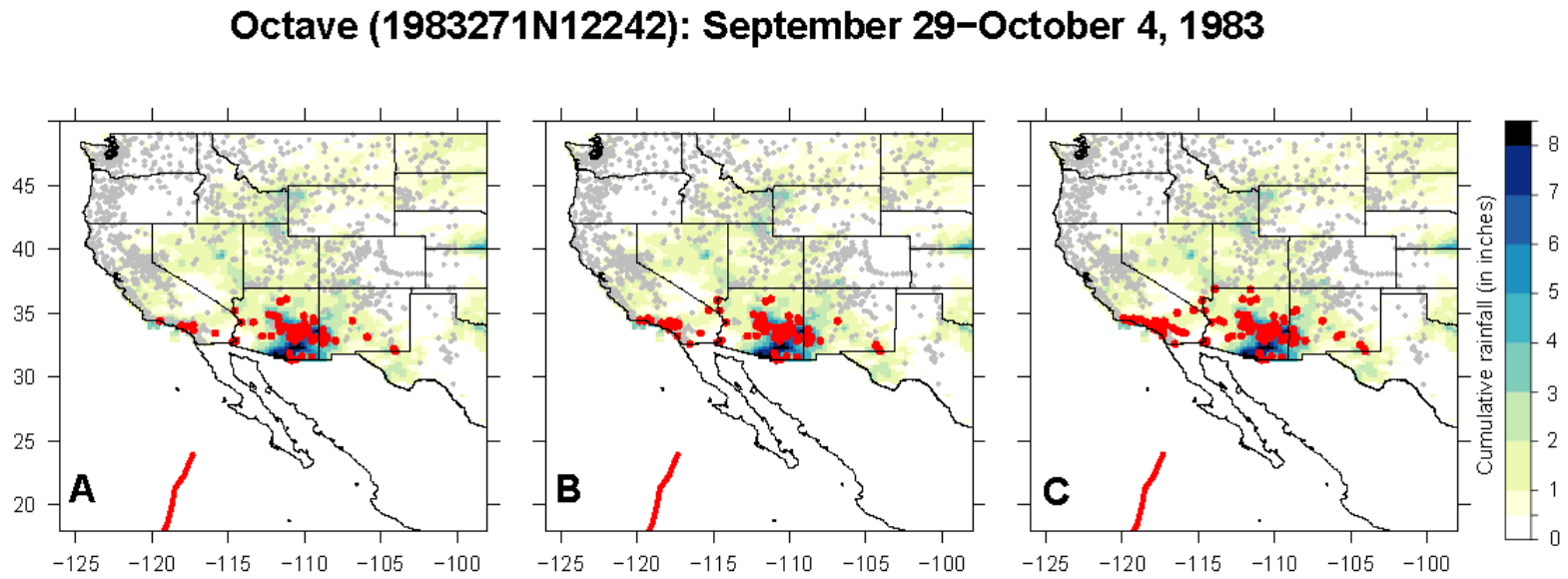

Figure 3.1: Hurricane Octave (the red line represents its track) and the cumulative precipitation (in inches) across the western United States during the TC-event window (defined as TC-days). The location of USGS streamgages with annual maximums based on daily discharge values (AM) (A) and peaks-over-threshold (POT) flows (B-C) attributed to TC-events (red circles) and those gages that did not have TC-event attributed flows (grey circles). POT threshold level (S) corresponding to two (POT2) and four (POT4) are shown in panels b-c, respectively.

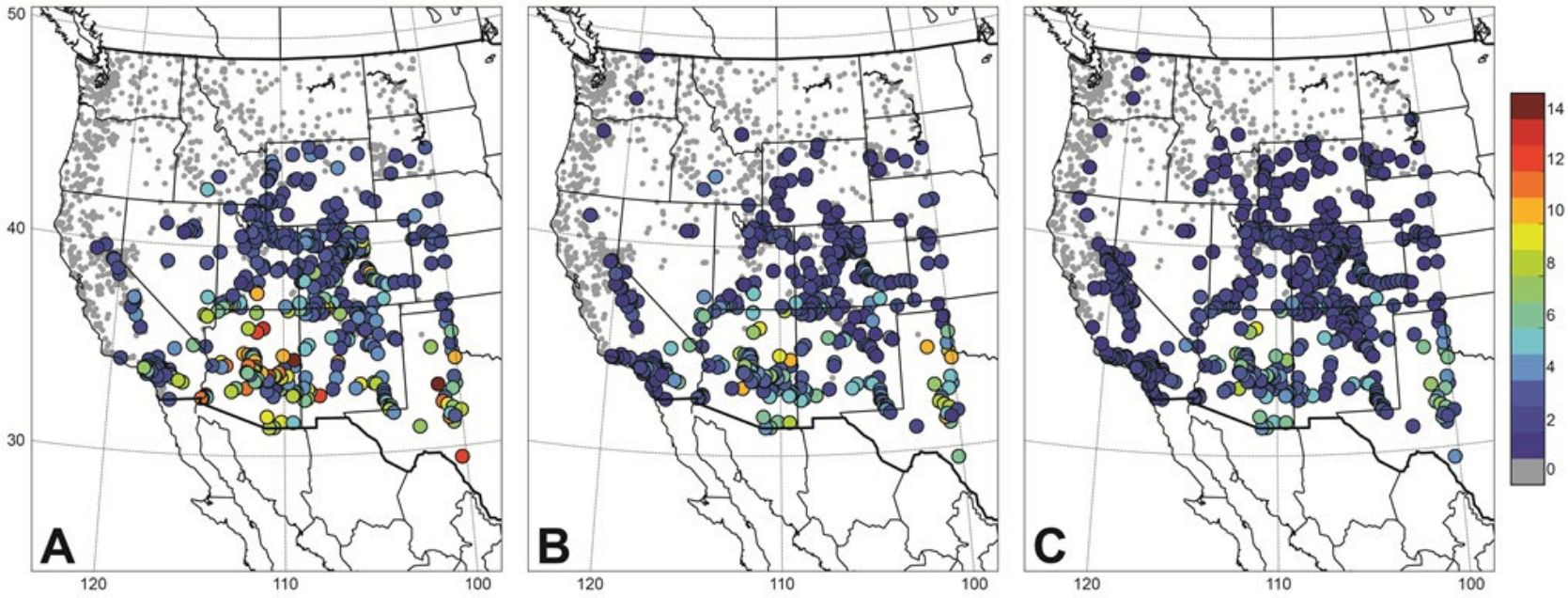

Figure 3.2: Percentage of annual maximum based on daily discharge values (AM) (A) and peak-over-threshold (POT) (B-C) flood events associated with TCs and their remnants. POT threshold level (S) corresponding to two (POT2) and four (POT4) are shown in panels B-C, flood events associated with TCs and their remnants. POT threshold level (S) corresponding to two (POT2) and four (POT4) are shown in panels B-C, respectively. The smaller grey circles represent USGS streamgages with no TC contributions.

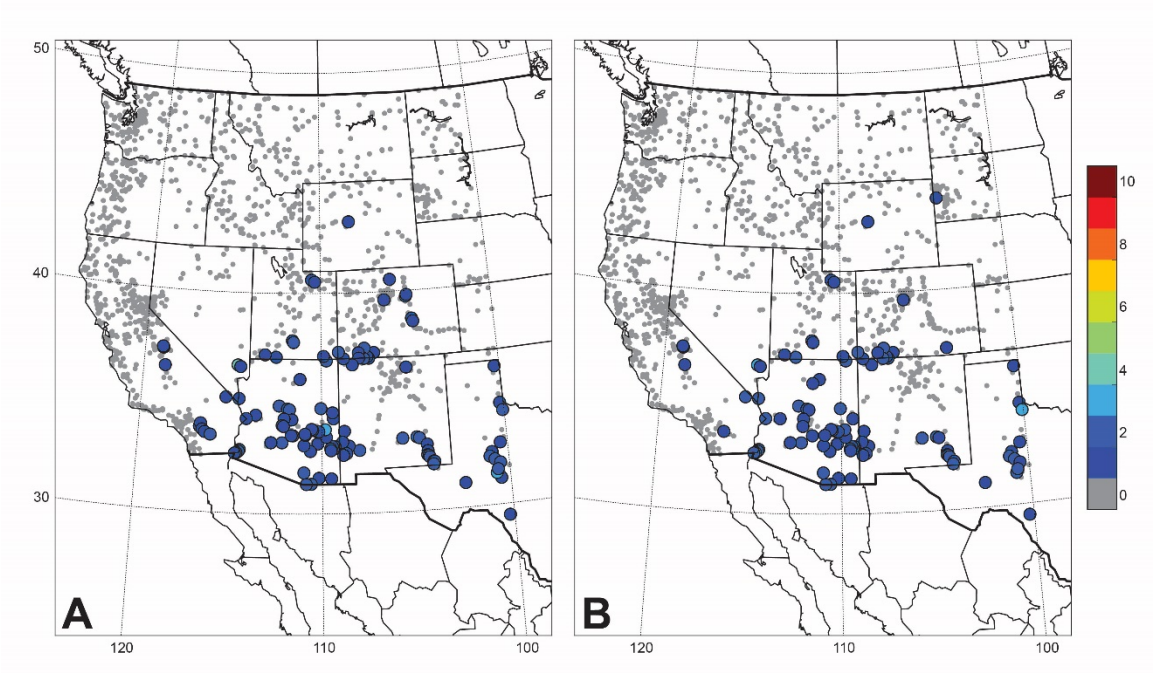

Figure 3.3: The number of TC-event-generated floods among the top-10 flood events for (A) the annual maxima based on daily discharge values (AM) and (B) the POT2 events.

3.5 Tables Chapter 3

Table 3.1: Characteristics of the 103 Tropical Cyclones (TC) and their remnants (TC-events) used to evaluate their contribution to annual maximum streamflow in the western United States from 1958-2010.

[Plot Number, corresponds to supplemental figures (S1-S103); Precipitation Reference,1, the National Oceanic and Atmospheric Administration (NOAA) Weather Prediction Center (WPC) (http://www.wpc.ncep.noaa.gov/tropical/rain/tcpointofentry.html); 2, Richie et al., 2011; 3, Corbosiero et al., 2009; HURDAT2, the Best Track Data from the National Hurricane Center).]

Plot Number	Precipitation Reference	HURDAT2 Serial Number	Tropical Cyclone Name	Date	Plot Number	Precipitation Reference	HURDAT2 Serial Number	Tropical Cyclone Name	Date
1	1	1958165N21266	Alma	June 14–16, 1958	53	1,2	1993252N11265	Lidia	September 8–15, 1993
2	3	1958254N21252	Tropical Storm 10	September 10–14, 1958	54	2	1994219N19252	Hector	August 7–11, 1994
3	3	1958273N16260	Hurricane 11	October 4–6, 1958	55	2	1994223N15253	Ileana	August 12–13, 1994
4	3	1960230N15260	Diana	August 19–23, 1960	56	1,2	1994282N17247	Rosa	October 9–15, 1994
5	3	1962263N17253	Claudia	September 22–26, 1962	57	2	1995220N15256	Flossie	August 11–13, 1995
6	3	1963261N13264	Jennifer–Katherine	September 17–21, 1963	58	1	1995244N16256	Henriette	September 1–7, 1995
7	3	1966269N19245	Kristen	September 28–October 2, 1966	59	1,2,3	1995256N15253	Ismael	September 12–16, 1995
8	3	1967242N18253	Katrina	September 1–5, 1967	60	1	1996172N13261	Alma	June 20–27, 1996
9	3	1968230N17252	Hyacinth	August 19–21, 1968	61	1	1996179N13266	Boris	June 26–July 1, 1996
10	3	1968270N13261	Pauline	October 2–5, 1968	62	1	1996184N09271	Cristina	July 1–6, 1996
11	3	1970243N16257	Norma	September 3–7, 1970	63	1,2,3	1996254N14258	Fausto	September 8–14, 1996
12	1,3	1972274N13255	Joanne	October 3–8, 1972	64	1,2	1996274N13261	Hernan	September 27– October 6, 1996
13	1	1973173N14262	Bernice	June 19–24, 1973	65	2	1997176N16252	Carlos	June 27–28, 1997
14	1	1973177N11264	Claudia	June 25–30, 1973	66	1,2	1997229N19242	Ignacio	August 19–21, 1997
15	1	1973265N13258	Irah	September 23–27, 1973	67	1,2,3	1997259N13258	Nora	September 14–27, 1997
16	1	1973266N13246	Jennifer	September 24–28, 1973	68	1,2	1997270N11265	Olaf	October 10–16, 1997
17	1	1974264N16263	Orlene/Fifi	September 17–25, 1974	69	1	1997279N12263	Pauline	October 3–10, 1997
18	1	1975295N14252	Olivia	October 21–26, 1975	70	1	1997311N09258	Rick	November 7–10, 1997
19	1,3	1976251N15251	Kathleen	September 7–12, 1976	71	2	1998216N09260	Frank	August 7–10, 1998
20	1,3	1976270N13253	Liza	September 30–October 3, 1976	72	1	1998233N25268	Charley	August 20–29, 1998
21	1	1976273N10270	Madeline	October 6–9, 1976	73	1,2,3	1998242N14252	Isis	August 31–September 5, 1998
22	1	1976299N17244	Naomi	October 28–31, 1976	74	1,2	1998249N16256	Javier	September 4–14, 1998
23	1,3	1977225N18254	Doreen	August 12–18, 1977	75	2	1998283N13268	Madeline	October 16–20, 1998
24	3	1977267N11253	Glenda	September 26–28, 1977	76	1,2	1999248N17256	Greg	September 1–10, 1999
25	1,3	1977277N15253	Heather	October 5–8, 1977	77	2,3	1999260N15255	Hilary	September 20–24, 1999
26	1	1978243N13262	Norman	September 4–7, 1978	78	2	2000164N12256	Bud	June 14–19, 2000
27	1	1980265N15283	Hermine	September 21–29, 1980	79	2	2000170N10268	Carlotta	June 23–27, 2000
28	1	1981280N20250	Lidia	October 5–8, 1981	80	2	2000226N16258	Ileana	August 14–15, 2000
29	1	1981282N15256	Norma	October 9–14, 1981	81	2	2000249N15258	Lane	September 11–15, 2000
30	1,3	1982262N11258	Olivia	September 23–28, 1982	82	1,2	2000260N19253	Miriam	September 14–18, 2000
31	3	1983255N12261	Manuel	September 18–20, 1983	83	1,2	2000264N17257	Norman	September 16–24, 2000
32	1	1983271N12242	Octave	September 29–October 4, 1983	84	1	2000309N10270	Rosa	November 3–8, 2000
33	1	1983285N09260	Tico	October 13–24, 1983	85	2	2001238N19252	Flossie	August 27 –September 2, 2001
34	1,3	1984259N19245	Norbert	September 24–27, 1984	86	2	2001254N15261	Ivo	September 13–14, 2001
35	1	1984262N14259	Odile	September 16–24, 1984	87	1,2	2001264N13269	Juliette	September 25–October 3, 2001
36	1	1984270N13262	Polo	October 1–4, 1984	88	2	2002238N15250	Genevieve	August 29–September 1, 2002
37	1	1985280N15254	Waldo	October 5–12, 1985	89	2	2002242N14257	Hernan	September 7–8, 2002
38	1,3	1986262N12266	Newton	September 17–26, 1986	90	1	2002268N15258	Julio	September 23–27, 2002
39	1	1986271N12267	Paine	September 26–October 4, 1986	91	1,2	2002295N11261	Kenna	October 24–26, 2002
40	1	1986289N10267	Roslyn	October 18–26, 1986	92	1	2003177N14263	Carlos	June 25–28, 2003
41	1	1987203N11263	Eugene	July 21–27, 1987	93	1,2,3	2003235N21253	Ignacio	August 16–20, 2003
42	1	1989169N13265	Cosme	June 17–24, 1989	94	1,2,3	2003262N17254	Marty	September 21–26, 2003
43	1	1989238N22253	Kiko	August 23–29, 1989	95	1,2	2003275N16252	Nora	October 3–8, 2003
44	1,3	1989268N14261	Raymond	October 3–7, 1989	96	1,2	2003276N12262	Olaf	October 3–8, 2003
45	1	1990154N11260	Boris	June 4–11, 1990	97	2	2004195N15255	Blas	July 13–19, 2004
46	1	1990216N13281	Diana	August 4–14, 1990	98	2	2004244N12257	Howard	September 2–7, 2004
47	1	1990270N14261	Rachel	September 30–October 3, 1990	99	1,2	2004255N11267	Javier	September 16–20, 2004
48	2	1992185N13265	Darby	July 6–9, 1992	100	2	2005271N17255	Otis	September 29–October 4, 2005
49	1,2,3	1992233N16254	Lester	August 19–26, 1992	101	1	2008203N18276	Dolly	July 22–28, 2008
50	1	1992281N11263	Winifred	October 3–11, 1992	102	1	2010247N15266	Hermine/Tropical Depression 11_E	September 3–11, 2010
51	1,2	1993186N13262	Calvin	July 3–9, 1993	103	1	2010264N20252	Georgette	September 19–23, 2010
52	2,3	1993229N13265	Hilary	August 24–30, 1993					

Chapter 4

ACCOUNTING FOR MIXED POPULATIONS IN FLOOD FREQUENCY ANALYSIS: A BULLETIN 17C PERSPECTIVE[3]

As described in Chapter 1, the need for reliable estimates of the frequency and magnitude of floods are essential for flood insurance studies, floodplain management, and the design of transportation and water-conveyance structures, such as roads, bridges, culverts, dams and levees. The federal guidelines of B17B and its update (B17C) recommend the use of the LP3 distribution to describe the populations of floods for all hydrologic regimes found throughout the United States. The conventional assumptions for performing flood frequency analyses are that the annual time series is a representative time sample of random homogeneous events, that the stochastic processes that generate floods are assumed to be stationary or invariant in time and are be generated by a single flood generating mechanism.

B17B and B17C recognize the difficulties in determining flood frequency estimates with streamflow records that contain flood peaks coming from different flood-generating mechanisms, such as those found in the Sierra Nevada region (Figure 1.1 and Figure 4.1). This high-elevation site illustrates the classic s-shaped curvature in the upper right hand tail of the distribution associated with mixed distributions. Rather than being an issue with the selected parametric distribution (i.e., LP3), this result suggests that the observational record could be stratified into at least two groups, one that mostly encompasses the low to medium flood peaks, and one that mostly includes large and extreme flood events from the upper tail of the

[3] Adapted from: Barth, N.A., Villarini, G., and White, K., (2018), Accounting for mixed populations in flood frequency analysis: A Bulletin 17C perspective, Journal of Hydrologic Engineering, (in press).

distribution. This site is not unique, but rather is representative of flood frequency results that is often encountered across the western United States.

As described in Chapters 1 and 2, atmospheric rivers (ARs) are responsible for extreme precipitation and flooding. In Chapter 2, I determined not only the spatial and fractional contributions of AR-generated peak flows at long-term USGS stations in this region, but their contribution to flood peaks among the top-10 largest annual maxima. I found that areas of the central Columbia River Basin in the Pacific Northwest, the Sierra Nevada, the central and southern California coast, and central Arizona all showed a mixture of 30–70% AR-generated flood peaks. Therefore, the fundamental assumption that the observed annual peak flows are representative of random homogeneous events may not properly apply to sites in these regions.

This Chapter provides a statistical framework to perform a process-driven flood frequency analysis using the AR/non-AR-generated annual peak flows identified in Chapter 2. I provide both visual and quantitative comparisons of the goodness-of-fit to assess the improvements in flood frequency estimates obtained using a weighted (heterogeneous) mixed population approach ($LP3_{mixed}$) against estimates based on an assumed single homogeneous flood distribution ($LP3_{single}$). I also account for sampling and mixing uncertainties for the weighted mixed population approach.

4.1 Data and Methodology

4.1.1 Location of Study Area

I use 43 long-term USGS streamgage records throughout the western United States with at least 50 years of annual peak flow data, with at least 25 of which are generated by ARs (Figure 4.2). A minimum of 25 years was selected due to the sensitivity of the skew parameter to outliers

in small sample sizes (IACWD, 1982). These streamgage records are a subset of the 1375 sites used in Chapter 2 to determine the spatial and fractional contribution of ARs among annual peak flows. Each of these sites has at least 30 years of data from 1979 to 2010 and were extended back in time, allowing for no more than one 2 year gap prior to water year 1979 (the water year starts on October 1 and ends on September 30). The average record length is 80 years among the qualifying long-term USGS sites. As described in Section 2.1, an annual peak flow is considered an AR-generated flood peak if it occurs within a 10-day window (2 days prior to and 7 days after) an AR event and the USGS streamgage is located within a distance of 2.5 degrees on either side of an AR major axis. Further details about the criteria used for AR identification is also described in Section 2.1.

These 43 streamgages did not present statistically significant abrupt or gradual changes and were verified to not be affected by land-use changes, urbanization, regulation or diversions. Figure 4.2 (top panel) shows the spatial locations of these sites and the location of three broad physiographic regions where the mixed population sites are located—(A) the Cascade/Sierra Nevada mountain ranges, (B) the intermountain west (IMW) that includes the Columbia Plateau, northern Rocky Mountains and portions of the Basin and Range in northeast Nevada, and (C) the Basin and Range and Transition Zone in central Arizona (e.g., Paretti et al., 2014; Archfield et al., 2016). Figure 4.2 (bottom panels) shows the number of AR/non-AR-generated annual peaks for each streamgage separated by the physiographic regions (A-C).

4.1.2 Weighted Mixed Population Flood Frequency Analysis Using a Heterogeneous Distribution Approach

A weighted mixed population approach is used to determine flood frequency estimates for a composite heterogeneous distribution. The additive rule of probability (e.g., Alila and Mtiraoui,

2002) is used to separate the complete period of record into two mutually independent AR/non-AR-generated subpopulations. Each subpopulation is fitted with an LP3 distribution and a weighted heterogeneous distribution is determined as follows:

$$LP3_{mixed} = \hat{\theta} \cdot LP3_{AR} + (1 - \hat{\theta}) \cdot LP3_{nAR} \quad (4.1)$$

where $LP3_{mixed}$ is the composite exceedance probability for the heterogeneous distribution and $\hat{\theta}$ and (1-$\hat{\theta}$) are the weighting factors corresponding to the fractional contribution of AR and non-AR-generated peaks, respectively. The ˆ symbol indicates that we are dealing with an estimate from our sample.

An important contribution to the updates of B17B is the identification of multiple potentially influential low floods (PILFs; low outliers) in a flood series. It has been long recognized that performing flood frequency in the western United States is complicated by flood series that often contain zero flows and/or PILFs (Cohn et al., 2013; Lamontagne et al., 2016). The original Grubbs-Beck (GB) test for outliers, described in B17B, does a poor job identifying the full set of annual peaks that may be considered to be low outliers. Because of the limitations of the original B17B procedure, hydrologists have often employed a subjective threshold below which flows are deemed low outliers. Cohn et al. (2013) published a new method for identifying all low outliers, called the multiple Grubbs-Beck (MGB) test. For consistency, I use the latest USGS PeakFQ software (version 7.1) that employs the new EMA-MGB methodology, as recommended in B17C, to estimate the parameters of the single population and for each of the AR/non-AR-generated subpopulations. In cases when PILFs are detected, the EMA-MGB procedure re-codes these PILFS as censored interval flows to estimate more accurately the sample moment estimates.

As recommended by B17B and B17C, the LP3 distribution parameters are estimated using the method-of-moments from the logarithms of the observed discharge values. The parameters of the LP3 distribution are the scale (α), shape (β), and location (τ). These parameters are functions of the three sample statistics (mean ($\hat{\mu}$), standard deviation ($\hat{\sigma}$), and skew ($\hat{\gamma}$)) through the following relations:

$$\hat{\alpha} = \frac{4}{\hat{\gamma}^2}, \tag{4.2}$$

$$\hat{\beta} = sign(\hat{\gamma}) \left(\frac{\hat{\sigma}^2}{\hat{\alpha}}\right)^{0.5} \tag{4.3}$$

$$\hat{\tau} = \hat{\mu} - \hat{\alpha}\hat{\beta}. \tag{4.4}$$

By taking into account the relative fraction of AR/non-AR peaks, I determine $LP3_{mixed}$ through Monte Carlo simulations with the following procedure:

1. Separate the complete period of record of size N_{PK} into AR (sample size N_{AR}) and non-AR (sample size N_{nAR}) peaks; determine the fractional contribution of each subpopulation (N_{AR}/N_{PK} corresponds to $\hat{\theta}$, while N_{nAR}/N_{PK} corresponds to (1-$\hat{\theta}$)).
2. Fit each separated subpopulation of AR/non-AR flood peaks with a corresponding LP3 distribution from the parameter estimates calculated using equations 4.2-4.4 based on the sample moments from the EMA-MGB methodology ($LP3_{AR}(\mathrm{X} \geq \mathrm{x}|\hat{\tau}_{AR}, \hat{\alpha}_{AR}, \hat{\beta}_{AR})$ and $LP3_{nAR}(\mathrm{X} \geq \mathrm{x}|\hat{\tau}_{nAR}, \hat{\alpha}_{nAR}, \hat{\beta}_{nAR})$).
3. Generate 10,000×N_{AR} variates from the $LP3_{AR}$ distribution and 10,000×N_{nAR} with parameters from the $LP3_{nAR}$ distribution. The derived combined record consists of 10,000×N_{PK} "synthetic peaks" with the observed fractional contribution of AR/non-AR events.

4. Estimate the empirical cumulative density function (ecdf) from the 10,000×N_{PK} synthetic peaks, which represents the mixed heterogeneous distribution ($LP3_{mixed}$) for that specific site.

One of the acknowledged disadvantages to this subpopulation approach (equation 4.1) is the increase in sampling uncertainties of the distribution parameters due to the reduced sample sizes (e.g., Alila and Mtiraoui, 2002; Villarini and Slater, 2017). There is also an additional source of uncertainty due to the fraction of mixing that needs to properly be taken into account. This is due to the fact that I estimate θ from a sample of 80-100 years. Therefore, I have developed a simulation framework to account for the uncertainties in the estimation of the LP3 parameters and of the mixing ratio θ.

The subpopulations of AR/non-AR-generated annual peaks are assumed to be independent and the probability that the number of observed AR-generated peaks (Y) in a flow series of length N_{PK}, is best described by the binomial probability mass function,

$$Y \sim Binomial\left(N_{PK}, \hat{\theta}\right) \tag{4.5}$$

Given $\hat{\theta}$, I generate 100 variates from a binomial distribution, with parameters N_{PK} and $\hat{\theta}$. Each of these 100 binomial variates is then used as the initial fractional contribution of AR-generated peaks. These new fractional contributions of AR/non-AR-generated peaks are used to calculate a synthetic peak record of size N_{PK}.

To illustrate how the mixing ratio and sampling uncertainties for each site are accounted for to estimate the overall uncertainties for the weighted mixed population approach, I describe the procedure below.

1. For the i^{th} binomial variate (equation 5) (i =1, 2,…, 100), determine the number of AR (N_{AR}^{i}) and non-AR (N_{nAR}^{i}) peaks. Overall, $N_{AR}^{i} + N_{nAR}^{i} = N_{PK}$.

2. Randomly generate a record of AR peaks of size N_{AR}^{i} using the LP3 parameters estimated for the original AR record ($S_{i,j,AR}$). Similarly, randomly generate a record of non-AR peaks of size N_{nAR}^{i} using the LP3 parameters estimated for the original non-AR record ($S_{i,j,nAR}$).
3. Determine the new LP3 parameter estimates from $S_{i,j,AR}$ and $S_{i,j,nAR}$ using the maximum likelihood method.
4. Generate long-term synthetic records with lengths $10{,}000 \times N_{AR}^{i}$ and $10{,}000 \times N_{nAR}^{i}$ based on the new LP3 parameter estimates in step 3, respectively. The derived combined record consists of 10,000×N_{PK} "synthetic peaks" with the fractional contribution of AR/non-AR events from the randomly generated values in step 1.
5. Estimate the ecdf from these 10,000×N_{PK} synthetic peaks, which represents the mixed distribution ($LP3_{mixed}[i,j]$) for that specific site.
6. For each i^{th} value (i =1, 2, …, 100), repeat steps 2-5 100 times.

This simulation approach allows me to capture the uncertainties in both the estimation of the LP3 parameters based on a limited sample, and in the fractional contribution of the two subpopulations. Based on the 10,000 synthetic records, I compute the median and the 95% confidence intervals (CIs) for different annual exceedance probabilities (AEPs). Table 4.1 provides the number of annual peaks (N_{PK}), AR-generated peaks (N_{AR}) and non-AR-generated peaks (N_{nAR}) used in these simulations.

4.1.3 Goodness-of-fit of Flood Quantile Estimates between the Homogeneous and Heterogeneous Distribution Approaches

I evaluate the goodness-of-fit of the fitted LP3 distributions using both visual and quantitative assessments. For each site, the visual assessment includes plotting the ecdf of the observed flood flows (using the Hirsch-Stedinger plotting positions (Hirsch and Stedinger (1987) to properly account for censored data (i.e., PILFs in this study; B17C)), together with the fitted LP3 distributions (assuming that all the peaks come from a single population and accounting for AR/non-AR peaks). In terms of quantitative evaluation, I compute the relative percent differences ($RPD^1{}_{AEP}$) at the 4% and 2% AEP corresponding to the 25- and 50-year return periods between the observations and the LP3 fits based on a single population or a mixture:

$$RPD^1{}_{AEP,single/mix} = \frac{\hat{Q}_{AEP,obs} - \hat{Q}_{AEP,single/mix}}{\hat{Q}_{AEP,obs}} \times 100\%, \quad (4.6)$$

where $RPD^1{}_{AEP,single/mix}$ is the relative percent difference for a given AEP based on the single population ("single") or accounting for AR/non-AR events ("mix"); $\hat{Q}_{AEP,obs}$ is the observed annual peak discharge at a given AEP level, while $\hat{Q}_{AEP,single/mix}$ represents the corresponding value based on the fitted LP3 for the single population or weighted mixed population approach, respectively. The values of $RPD^1{}_{AEP,single/mix}$ that are positive (negative) indicate that the quantile estimates are lower (higher) than the observed data at the corresponding probability of exceedance. A $RPD^2{}_{AEP,single/mix}$ is additionally calculated to quantify the differences between the quantile estimates for the single and weighted mixed population approaches at the 1% and 0.5% AEPs corresponding to the 100 and 200-year return periods, respectively:

$$RPD^2{}_{AEP,single/mix} = \frac{\hat{Q}_{AEP,single} - \hat{Q}_{AEP,mix}}{\hat{Q}_{AEP,single}} \times 100\%. \quad (4.7)$$

A positive (negative) $RPD^2{}_{AEP,single/mix}$ indicates that the quantile estimates for the single population are higher (lower) than the weighted mixed population approach. The main difference between the metrics in equations 4.6 and 4.7 is that in the latter I do not use observations as the

reference, but I only consider the relative difference between single and mixed populations in terms of the estimates of discharge with low AEPs (or high return periods).

4.1.4 Sensitivity Analysis with Frequency Curve Comparisons

B17B and its update, B17C, recommend the LP3 distribution be used as a federal guideline to provide a consistent flood frequency framework. However other distributions also exist for hydrologic extremes and have been used to determine flood frequency relations, with the generalized extreme value (GEV) distribution that is often used because of theoretical reasons (e.g., Coles, 2001). Therefore I conducted a sensitivity analysis to compare the fitted distributions for the homogeneous single population and the heterogeneous weighted mixed population using the GEV distribution with the same statistical framework described in Section 4.1.2. The estimation of the three parameters of the GEV distribution is performed by means of the L moments method (Hosking, 1990). The results of the corresponding fitted GEV frequency curves are shown with the LP3 fitted curves to provide a qualitative visual comparison of the goodness-of-fit between the probability distributions and the observed peak flow data. Because the focus of this study is accounting for mixed populations from a B17C perspective, the results and discussion primarily focus on the LP3 distribution using the EMA-MGB approach. Yet a comparative analysis based on the GEV distribution is presented in the Discussion section and in the Supplemental Material (Chapter 6).

4.2 Results

4.2.1 Qualitative Comparison of Flood Frequency Estimates Based on the Mixture of Flood-generating Mechanisms

Figures S.1-S.43 in the Supplemental Material section show the flood frequency estimates based on the LP3 probability distribution for: (i) a single homogeneous population ($LP3_{single}$) (top panel) comprised of annual peak flows generated from AR/non-AR events; (ii) the subpopulation of AR-generated peaks ($LP3_{AR}$) (middle panel); and (iii) the subpopulation of non-AR-generated peaks ($LP3_{nAR}$) (bottom panel). The flood frequency plots not only highlight the location of the flood generating-mechanisms throughout the distribution (top panel) but also provide a visual goodness-of-fit of the LP3 frequency curve fitted to the individual subpopulations of flood flows generated by AR/non-AR events (middle and bottom panels, respectively). The plots provide additional information about not only the number of PILFs identified in each population (complete, AR or non-AR), but the flood-generating mechanism that produced these flows. Table 4.1 provides the number of PILFs and the corresponding critical threshold below which flows were censored among the single population ($N_{PILF,PK}$, $Q_{MGB,PK}$) and AR/non-AR subpopulations ($N_{PILF,AR}$/$Q_{MGB,AR}$ and $N_{PILF,nAR}$/$Q_{MGB,nAR}$, respectively).The fits of the LP3 curves for each population and the number of PILFs identified play a key role in the visual goodness-of-fit comparison between the single homogenous and weighted mixed population fits to the observed data (Section 4.2.2).

The LP3 frequency curves for the AR/non-AR-generated subpopulations indicate regional patterns in the overall fit to the recorded peak flow data and regional patterns in the skew (shape) parameters (i.e., the curvature of the distribution) among the subpopulations. The log-unit skews of the LP3 distribution for single population ($\hat{g}_{PK}$) and AR/non-AR subpopulations ($\hat{g}_{AR}$ and $\hat{g}_{nAR}$, respectively) using the recommended EMA-MGB approach (B17C) are provided in Table 4.2. Most of the sites in the Cascade/Sierra region display notable breaks in the slope of the observed flood peaks in the upper tail of the distribution, the classic s-shaped curvature

indicative of mixed populations, and positive skews (concave in the tails). However, the lower elevation sites in California are characterized by negative skews (convex) and a substantial number of PILFs. This is consistent with the new USGS regional skew equations developed for California which are based on mean basin elevation (Parrett et al., 2010). Parrett et al. (2010) found that high elevation sites have large positive log-unit skews (from +0.13 to +0.61, for mean basin elevations ranging from 6,000 to 11,000 feet) and large negative skews at lower elevation sites (from -0.21 to -0.61, spanning from 4,000 to 0 feet). These skew patterns correspond to the transition from snowmelt generated peaks at the high elevations to rainfall generated peaks at the lower elevations.

Sites in the IMW show a wide range of skews. For example, eastern Oregon and northeastern Nevada generally have near zero and negative skews with AR-generated peaks commonly found among the largest flows (e.g., gaging station 13333000: Grande Ronde River at Troy, OR (Figure S.40)), while gages in Idaho more commonly have multiple breaks in slope throughout the distribution, and have both negative and positive skews with some high visual outliers generated by ARs (e.g., gaging stations 12414500: St. Joe River at Calder, ID and 13258500: Weiser River near Cambridge, ID (Figures S.34 and S.39)). Two sites in western Montana (gaging stations 05014500: Swiftcurrent Creek at Many Glacier, MT and 06099500: Marias River near Shelby, MT (Figures S.1 and S.2)) have extremely large positive skews (>+0.8; Table 4.2) and were not used in developing regional skew equations due to their problematic mixed populations (Parrett and Johnson, 2003; Sando et al., 2015). Furthermore, more than 50% of the sites in Idaho and one in northeast Nevada had PILFs among the single population, AR and/or non-AR subpopulations.

In central Arizona, a more consistent pattern of an overall negative skew is found among the single population and AR-generated subpopulation, while both positive and negative skews are found among the non-AR subpopulation. Furthermore the observed flows at most of the sites fall into three general shapes: (1) an s-curve in the upper tail of the distribution (e.g., gaging stations 09448500: Gila River at Head of Safford Valley, near Solomon, AZ and 09484000: Sabino Creek near Tucson, AZ (Figures S.4 and S.6)); (2) asymptotic thresholds above and below which the observed flow magnitudes are bounded (e.g., 09496500: Carrizo Creek near Show Low, AZ and 09497500: Salt River near Chrysotile, AZ (Figures S.8 and S.9)); and (3) a pronounced upper bound that is indicative of highly negative skewed distributions (e.g., 09499000: Tonto Creek above Gun Creek, near Roosevelt, AZ (Figure S.11)). Despite the overall tendency for negatively skewed homogenous populations, only one site had one PILF identified in both the single population and the non-AR subpopulation and one site had multiple PILFs in the AR subpopulation.

Figure 4.3 shows a comparison of the log-unit skew (shape parameter) for the single and the AR-generated subpopulations for the 43 long-term USGS streamgages. Three notable patterns are found among the Cascade/Sierra Nevada and Arizona physiographic regions. First, the low elevation California sites are clustered on or above the 1:1 line in the third quadrant indicating the single population at-site skew is more negative than the AR-generated subpopulation (e.g., 11317000: Middle Fork Mokelumne River at West Point, CA (Figure S.69); 11348500: Pit River near Canby, CA (Figure S.71)). The differences are a function of the number of PILFs identified in the single and/or AR subpopulations. Second, the high elevation California sites cluster almost exclusively below the 1:1 line in the first quadrant indicating that the single population at-site skew is greater than the AR-generated subpopulation (e.g., 10336660: Blackwood Creek near

Tahoe City, CA (Figure S.63) and 112665000: Merced River at Pohono Bridge near Yosemite, CA (Figure S.67)). Third, the majority of the central Arizona sites are clustered in a region below the 1:1 line indicating the AR-generated subpopulation at-site skews are all negative and smaller than the single population (e.g., 09497500: Salt River near Chrysotile, AZ (Figure S.52) and 09499000: Tonto Creek above Gun Creek, near Roosevelt, AZ (Figure S.54)). Not only are the above-listed sites representative of the differences found in the at-site skew patterns among the two populations, but many are outliers from the 1:1 line. As will be discussed in Section 4.2.3, many of these sites have the largest differences in quantile estimates among the 1% and 0.5% AEPs.

4.2.2 Quantitative Goodness-of-fit

In the previous section I focused on the fit of the observed data based on the LP3 distribution, and highlighted differences in the role that ARs play in the flood frequency analyses for these sites. Here I apply the methodology described in Section 4.1.2 to account for the mixed distributions. Figures S.44-S.86 show the flood frequency curves based on: (i) the single homogenous population ($LP3_{single}$) the same as shown in the top panel of Figures S.1-S.43) with the corresponding 95% CIs based on the $LP3_{single}$ and (ii) the weighted mixed population approach ($LP3_{mixed}$) with the corresponding 95% CIs based on the $LP3_{mixed}$ that account for sampling and mixing uncertainties.

Up to this point, the discussion of the results was based on a qualitative analysis of the fitted distribution, relating the observed behavior to the PILFs and the role played by different flood agents. Here I use the $RPD^1{}_{AEP}$ to quantitatively compare quantile estimates from the LP3 distributions (Table 4.2) to the observations at the 4% and 2% AEPs. Figure 4.4 shows the

RPD^1s for the single and weighted mixed populations among (a) all the 43 USGS streamgages, (b) the 17 gages in the Cascade/Sierra Nevada, (c) the 16 gages in the IMW, and (d) the 10 gages in central Arizona physiographic regions. In general, the two approaches share similar quantile estimates and RPD^1s at the 4% AEP (25-year return interval) across the study region (Figure 4.4A). However larger differences in the RPD^1s are found at the 2% AEP (50-year return interval) for both frequency curves and particularly for the $LP3_{single}$ approach (Figure 4.4A-D), as indicated by the increased spread in the boxplots. The overall median ${RPD^1}_{2\%}$ values are nearly the same and positive (Figure 4A), indicating that the quantile estimates are below the observed values. However, there is more variability in the $LP3_{single}$ estimates at the 2% AEP due to five low outliers all located in central Arizona.

Figure 4.4 (panels B-D) show the results of the ${RPD^1}_{AEP}$ based on physiographic regions to further evaluate and highlight the increased RPD^1 differences between the two approaches at the 4% and 2% AEPs. Large positive RPD^1s are found among the 4% and 2% AEPs in the Cascade/Sierra Nevada mountain region (Figure 4.4B). The large positive RPD^1s indicate that both frequency curves are below the observed value, even though the variability in the $LP3_{mixed}$ fits are reduced compared to the $LP3_{single}$. This pattern is consistent with the large positive skews and s-shaped curves that are common in the right-hand tail at these high-elevation, mixed population sites. Furthermore the breaks in slope in the observed flows are found near the 4% AEP at over 50% of the sites (e.g., Figures S.56-S.59, S.73-S.74, S.78-S.79). In the IMW, more pronounced differences in the RPD^1s are found at the 2% AEP (Figure 4.4C). In general, the ${RPD^1}_{2\%}$ for both frequency curves are positive which is primarily due to the positively skewed distributions and visual high outliers. In both the Cascade/Sierra Nevada and IMW

physiographic regions, the $LP3_{mixed}$ estimates are closer to the observed values than the $LP3_{single}$ estimates, and exhibit less variability in the estimates.

The sign and magnitude of the RPD^1s at the 4% and 2% AEP in central Arizona are distinctly different from the other two hydrologic regions. The ${RPD^1}_{4\%}$ values at 50% of the gaging sites have large positive values while 70% of the sites have large negative ${RPD^1}_{2\%}$ values. This reversal in ${RPD^1}_{AEP}$ sign coincides with the location of changes in slope, s-shaped curves, the highly negative at-site skew among the AR subpopulation, and/or horizontal asymptotes found among the observed flows in the right-hand tail (Figures S.3-S.12 and S.46-S.55). Similar to the other two physiographic regions, the $LP3_{mixed}$ estimates at the 2% AEP are closer to the observed values than the $LP3_{single}$ estimates.

4.2.3 Frequency Curve Comparisons for Low Annual Exceedance Probabilities

Based on the qualitative and quantitative comparisons between the $LP3_{single}$ and $LP3_{mixed}$ frequency curves, 34 (80%) of the 43 sites have similar flood frequency curves from both methodological approaches. Yet nine (20%) of the sites have different quantile estimates in the upper tail of the distribution. Two important factors contribute to the overall differences found in the flood frequency estimates among these sites, regardless of their physiographic locations. The best visual goodness-of-fit in the upper tail of the distribution, the portion of most concern when designing flood flow structures, is found when: (i) PILFS are identified and/or (ii) when the composite distribution contains notably different at-site log-unit skews (shape parameter) among the AR/non-AR subpopulations compared to the single homogeneous population (Figure 4.3).

Figure 4.4 shows the quantitative goodness-of-fit between the LP3 frequency curves for the single and weighted mixed distributions and the underlying data within the range of the empirical probabilities of the flow data. However, most design flood structures require flood quantile estimates that are beyond the empirical probabilities of the data such as at the 1% and 0.5% AEPs which correspond to the 100 and 200-year return interval, respectively. Because the true values are not known and cannot be estimated directly from the observations, a qualitative comparison is made between the $LP3_{single}$ and $LP3_{mixed}$ quantile estimates (Table 4.2) for the 1% and 0.5% AEPs. Figure 4.5 shows the results of the ${RPD^2}_{AEP}$ between the single and weighted mixed population estimates among (a) all the 43 USGS streamgages, (b) the gages in the Cascade/Sierra Nevada, (c) the gages in the IMW, and (d) the gages in central Arizona physiographic regions. Overall the median RPD^2s between the two frequency curves are close to zero and the interquartile range is between +3% and -10%. However, there are five notable large outliers (Figure 4.5A) which are all highly negatively skewed central Arizona sites (09490500: Black River near Fort Apache, AZ; 09496500: Carrizo Creek near Show Low, AZ; 09497500: Salt River near Chrysotile, AZ; 09498500: Salt River near Roosevelt, AZ; 09499000: Tonto Creek above Gun Creek, near Roosevelt, AZ; Figures S.50-S.54).

If I further examine the RPD^2s within the three physiographic regions (Figure 4.5B-D), three sites in the Cascade/Sierra Nevada mountain region (111895000: South Fork Kern River near Onyx, CA; 11317000: Middle Fork Mokelumne River at West Point, CA; and 12459000: Wenatchee River at Peshastin, WA (Figures S.66, S.69, S.79)) have a ${RPD^2}_{0.5\%}$ difference of -15%, -17% and -14%, respectively and the IMW site 12414500 (St. Joe River at Calder, ID (Figure S.77)) has a -13% and -19% RPD^2 for the 1% and 0.5% AEPs, respectively. These negative RPD^2 values indicate that the quantile estimates for the $LP3_{mixed}$ are higher compared

to the $LP3_{single}$. In sharp contrast, the median RPD^2s at the 1% and 0.5% AEPs in central Arizona are 12% and 15% and the corresponding interquartile ranges are between zero and 22%. Additionally, site 0949900 (Figure S.54) has a RPD^2 of 30% and 35% at the 1% and 0.5% AEPs, respectively. These RPD^2s in central Arizona reflect a substantially higher estimate for the single population compared to the mixed population based on large differences in the at-site log-space skew between the single population and the AR-generated subpopulation (Table 4.1 and Figure 4.3). The above-listed nine USGS sites represent the 20% of the flood frequency curves that have visually different quantile estimates in the upper tail of the distribution.

4.2.4 Heterogeneous Distribution Confidence Intervals for Quantile Estimates

Among the updates in B17C is the improved method for computing confidence intervals about the LP3 flood frequency curve. Notable differences may be found between the confidence intervals using the B17B procedure versus the methods in B17C. The intervals tend to be wider with the B17C procedure (EMA-MGB) because it properly incorporates the uncertainties of the estimated skew, censored data and historical information (England et al., 2018). However, as previously described, no updates were made for treating mixed populations in B17C and therefore no guidelines are recommended for quantifying the uncertainties with a weighted mixed population approach. Using the simulation framework described in Section 4.1.2, the $LP3_{mixed}$ CIs tend to be wider than the $LP3_{single}$ CIs in both tails of the distribution. This is primarily due to the reduced sample size from separating the observed flow series into AR/non-AR subpopulations and the contributions from the mixing fraction of ARs. Yet I found similar interval widths throughout the remaining distribution, implying that the simulation framework can capture the improved procedures of B17C in addition to the mixing ratio uncertainties. In

cases where the flood frequency estimates are nearly the same between both approaches, the CIs have the same general shape, yet the CIs are consistently wider especially in the upper tail (e.g., USGS gaging site 13337000: Lochsa River near Lowell, ID; Figure S.84).

There are notable differences in the CI widths when PILFs are identified in the single population ($LP3_{single}$). In the EMA-MGB procedure, observed flows below the PILF threshold are re-coded as interval (censored) flows (B17C) and the resulting uncertainties in the left-hand tail of the distribution are considerably wider (e.g., site 11317000: Middle Fork Mokelumne River at West Point, CA (Figure S.69)). Because no PILFs were identified among the AR/non-AR subpopulations at site 11317000, the CIs for $LP3_{mixed}$ are much narrower. Similar differences in the interval widths in the left-hand tail are found at sites where PILFs are identified among the single population and AR and/or non-AR subpopulations. While the $LP3_{mixed}$ flood frequency curve may provide an overall better visual goodness-of-fit to the observed flood data than the $LP3_{single}$ curve at many sites with PILFs, the narrow interval width in the left-hand tail should be interpreted with caution. A similar discrepancy between the interval widths in the left-hand tail is found for the highly positive skewed USGS gage 05014500: Swiftcurrent Creek at Many Glacier, MT (Figure S.44). This mixed population Montana site has a single population at-site skew of +1.71 and the AR (non-AR) subpopulation skews of +1.29 (+1.99).

4.3 Discussion

As described in Section 4.1.4, a sensitivity analysis was conducted to compare the fitted distributions for the homogeneous single population and the heterogeneous weighted mixed population using the GEV distribution based on L moments. Table 4.3 provides the magnitude

and frequency estimates based on the GEV distribution for the 4%, 2%, 1%, and 0.5% AEPs for the 43 long-term USGS streamgaging records. In general, the LP3 and GEV probability distributions have visually similar frequency curve fits among the underlying data within the range of the empirical probabilities of the flow data (Figures S.1-S.86). Asquith et al. (2017) found similar results between the fits of the LP3 (EMA) and GEV (L moments) probability distributions at two positively skewed, long-term USGS sites in the northeastern United States. However at several sites in this study, the two flood frequency curves diverge in the right-hand tail of the distribution when the at-site skews have large positive (negative) values (e.g., 111895000: South Fork Kern River near Onyx, CA (Figure S.66) and 09497500: Salt River near Chrysotile, AZ (Figure S.52), respectively).

Figure 4.6 further highlights the quantitative goodness-of-fit ($RPD^1{}_{AEP}$) for the single and mixed populations of both the LP3 and GEV probability distributions at the 4% and 2% AEP. Overall, the LP3 single and weighted mixed population quantile estimates are closer to the observed flow data at both the 4% and 2% AEPs than the corresponding GEV estimates. And the $LP3_{mixed}$ consistently improves the frequency curve fit to the observed data among all physiographic regions compared to the GEV_{mixed} (Figure 4.6, panels B-D). As previously mentioned in Section 4.2.2, the variability in the boxplot range for the $LP3_{mixed}$ is substantially smaller than the $LP3_{single}$. Conversely, the boxplot range remains the same or in some cases is slightly wider for the GEV_{mixed} than the GEV_{single}. Yet most notably, the differences in quantile estimators between the LP3 and GEV distributions are small compared to the uncertainty in the quantile estimators themselves (Stedinger and Griffis, 2008). Therefore based on the visual and quantitative comparisons of the goodness-of-fit between the LP3 and GEV distributions in this

sensitivity analysis, the LP3 (EMA-MGB) distribution remains reasonably consistent with the observed flow data.

As previously mentioned, because the true distribution is not known and large differences in the visual goodness-of-fit between the LP3 and GEV distributions are found among some sites, this further highlights the need for (1) a consistent, unified approach that utilizes one probability distribution (e.g., continued use of the LP3 distribution) (Stedinger and Griffis, 2008; B17C), as well as (2) incorporating the physical flood-generating mechanisms into a mixed populations analysis when appropriate. One of the major updates to B17B in B17C is the new MGB test to identify PILFs. When zero flows and PILFs are identified by the MGB test and are in turn recoded as censored flows, their influence in the upper tail of the distribution is lessened. If the true population is negatively skewed, then many of the smallest values will be identified as PILFs (Cohn et al., 2013). PILFs are identified, for example, in the single population or non-AR subpopulation at sites (11317000: Middle Fork Mokelumne River at West Point, CA (Figure S.69); 12414500: St. Joe River at Calder, ID (Figure S.77); and 12459000: Wenatchee River at Peshastin, WA (Figure S.79)). The corresponding $LP3_{mixed}$ flood frequency fits at the 1% and 0.5% AEPs in the right-hand tail of the distribution all show an improved visual goodness-of-fit. Yet when the MGB test is employed among the highly negatively skewed sites in central Arizona, PILFs are rarely found. The most improved $LP3_{mixed}$ frequency fits in central Arizona do not come from the identification of PILFs, they come from separating the observed flows based on flood-generating mechanisms.

Finally, Parrett et al. (2010) found that the regional skew in California is a function of mean basin elevation (ELEV). They found that ELEV best described the variability in at-site skews among sites whose flows are generated by low elevation rainfall events and high elevation

snowmelt. And Gotvald et al. (2012) and Paretti et al. (2014) recently developed regional regression equations for ungaged basins in the complicated hydrologic regions of the Sierra Nevada and the central Highlands region of Arizona, respectively. In these mountainous regions, the regression equations are functions of drainage area, mean annual precipitation and ELEV. These USGS regional studies highlight the need to include the ELEV basin characteristic to best describe the variability in regional skew and flood flows in these regions that contain sites with mixed populations. Similarly, I found the most improved $LP3_{mixed}$ visual and quantitative goodness-of-fits when PILFS are identified and/or when the composite distribution contains notably different at-site log-unit skews among the AR/non-AR subpopulations compared to the single homogeneous population throughout our study region.

4.4 Conclusions

This study provides a statistical framework to perform a process-driven flood frequency analysis using the AR/non-AR-generated annual peak flows identified in Chapter 2 at 43 long-term USGS streamgages in the western United States. I use a simulation framework to perform flood frequency analyses in terms of mixed distributions and quantify the corresponding uncertainties. The gaging stations are located in three broad physiographic regions: the Cascade/Sierra Nevada mountain ranges, the intermountain west (IMW), and central Arizona. The flood frequency estimates from the mixed population approach ($LP3_{mixed}$) is compared with the frequency estimates from a single homogeneous population ($LP3_{single}$) using visual and quantitative goodness-of-fit assessments. The primary findings of this study are summarized below:

1. The visual goodness-of-fit of the LP3 frequency curves for the single population and the AR/non-AR subpopulations indicate regional patterns in the overall fits to the observed streamflow data and in the log-unit skew (shape parameter). The observed data among the high-elevation sites in the Cascade/Sierra Nevada region show sharp breaks in slope, s-shaped curves in the right-hand tail of the of the distribution and positive skews. The low elevation sites in California are characterized by negative skews and PILFs that are identified in the single population and/or AR/non-AR subpopulations. The sites in the IMW show a wide range of skews, yet gages in Idaho commonly have multiple breaks in slope, are positively skewed and have high outliers while the sites in Montana have large positive skews. Many of the gage sites in central Arizona have large negative skews in the single population which are characterized by lower and upper horizontal asymptotic bounds in the observed streamflow. The AR-generated flows in central Arizona are commonly found in the upper tail of the single population and the AR subpopulation has a consistent negative skew while the non-AR subpopulation have near-zero or positive skews.
2. Thirty-four (80%) of the 43 sites have similar flood frequency curves from both the homogeneous (single) and heterogeneous (weighted mixed) population methodological approaches. Yet nine (20%) of the sites have notably different quantile estimates in the upper tail of the distribution. Two important factors contribute to the overall differences found in the flood frequency estimates among these sites, regardless of their physiographic locations. The best quantitative and visual goodness-of-fit in the upper tail of the distribution, the portion of most concern when designing flood flow structures, is found when: (i) PILFS are identified and/or (ii) when the composite distribution contains

markedly different at-site log-unit skews (shape parameter) among the AR/non-AR subpopulations compared to the single homogeneous population. Sites in the Cascade/Sierra Nevada and IMW that have the largest differences between the 4%, 2%, 1% and 0.5% AEP quantile estimates ($RPD^1{}_{AEP}$ and $RPD^2{}_{AEP}$, respectively) in the right-hand tail are a function of the number of PILFs that are identified in the single population versus the AR/non-AR subpopulations and the differences in at-site skews. Yet the largest $RPD^1{}_{AEP}$ and $RPD^2{}_{AEP}$ are found in central Arizona and reflect a substantially higher estimate for the single population compared to the mixed population based solely on large differences in the log-space skew between the single population and the AR-generated subpopulation and not on the identification of PILFs among the populations.

3. I used a simulation framework to quantify the uncertainties in the sampling and mixing ratios for the $LP3_{mixed}$ quantile estimates. The $LP3_{single}$ CIs are computed using the improved procedure described in B17C that incorporates the uncertainties of the estimated skew, censored data and historical information in the EMA-MGB methodology. The $LP3_{mixed}$ CIs tend to be wider than the $LP3_{single}$ CIs in both tails of the distribution. This is primarily due to the reduced sample size from separating the observed flow series into AR/non-AR subpopulations and the contributions from the mixing fraction of ARs. Yet I found similar interval widths throughout the remaining distribution, implying that the simulation framework can capture the improved procedures of B17C in addition to the mixing ratio uncertainties.
4. As a general statement, beside the improvements in the estimation of the quantiles for different AEPs, I believe that this methodology addresses the issue raised by Klemeš (1974): our effort serves to move us back to treating the observed peak values as

representative of physical processes rather than as a "collection of abstract numbers that could pertain to anything or to nothing at all." By doing so, I also address one of the issues left untouched in B17C: mixed distributions. This methodology is general enough that it can be adopted outside of the western United States and for different flood generating mechanisms, and I hope that it can represent a path forward for future revisions of B17C.

4.5 Figures Chapter 4

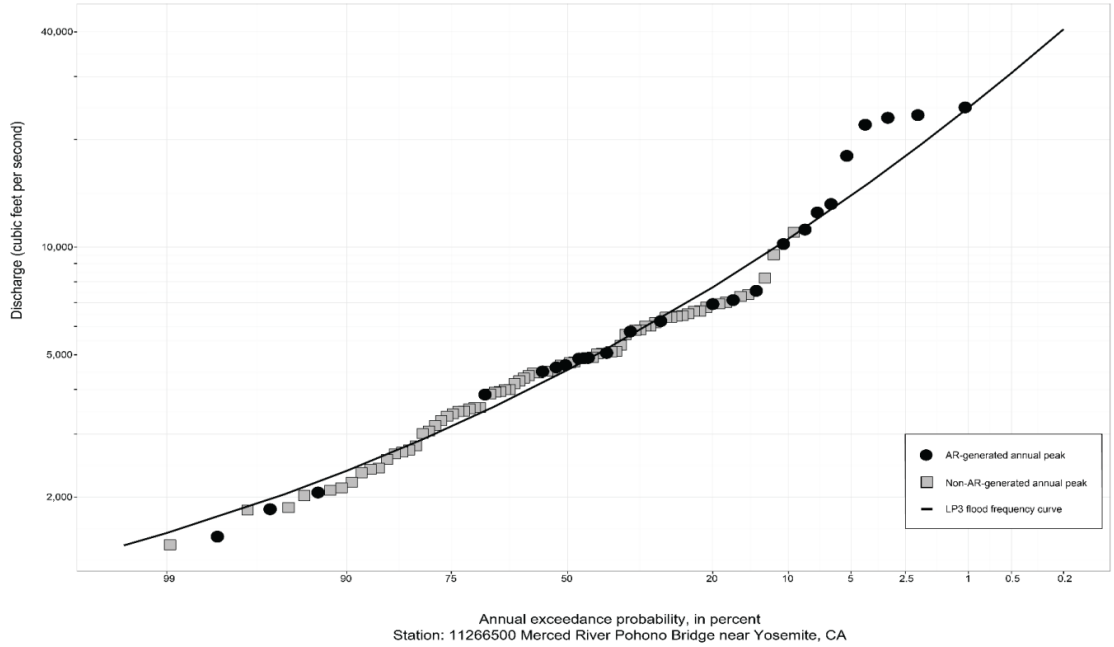

Figure 4.1: Observed streamflow data (symbols), and the fitted LP3 distribution (black line) at Merced River at Pohono Bridge near Yosemite, California (USGS streamgage 11266500). Black circles (grey squares) are the AR (non-AR)-generated annual peaks.

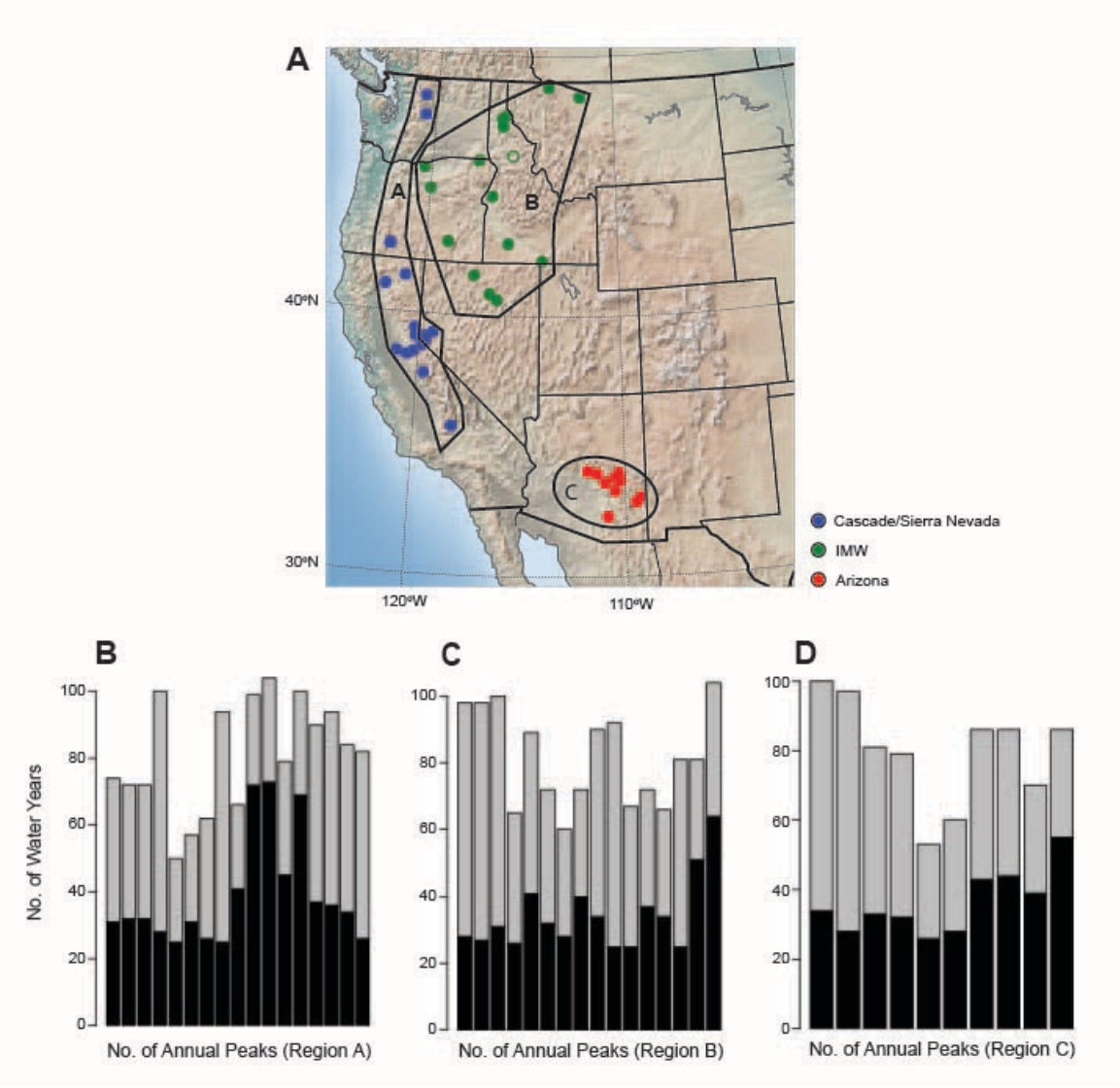

Figure 4.2: (Top panel) Location of the 43 long-term USGS streamgages with at least 50 years of continuous record through 2010. Three broad physiographic regions are identified: region A is the Cascade/Sierra Nevada mountain region (17 sites), region B is the intermountain west (IMW) (16 sites), and region C is central Arizona (10 sites). (Bottom panels) The number of AR (black)/non-AR (grey)-generated annual peaks per streamgage grouped by physiographic regions.

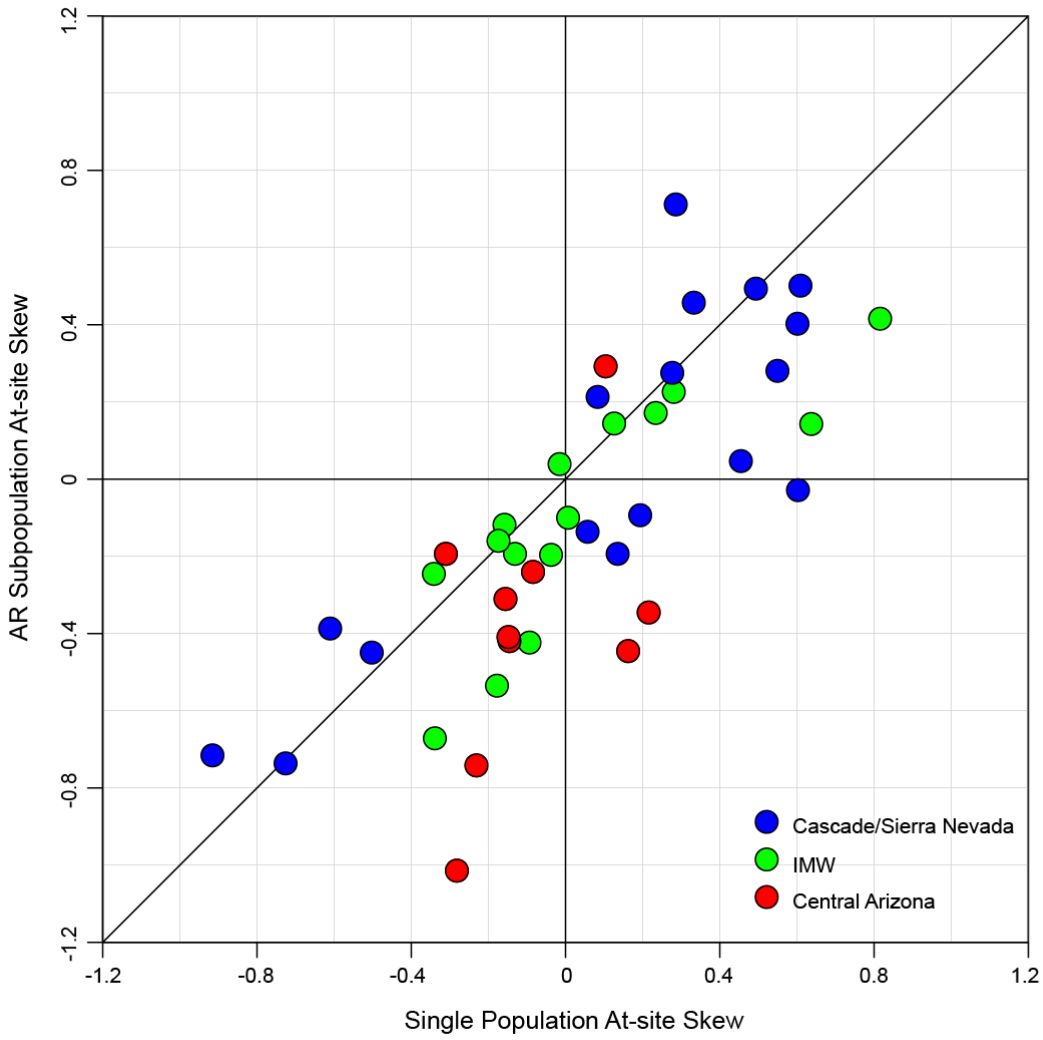

Figure 4.3: Single population versus AR subpopulation at-site log space LP3 skew estimates using EMA-MGB methodology. Region A (blue dots), region B (green dots) and region C (red dots) correspond to USGS streamgages located in the Cascade/Sierra Nevada mountain region, intermountain west (IMW), and central Arizona physiographic regions, respectively.

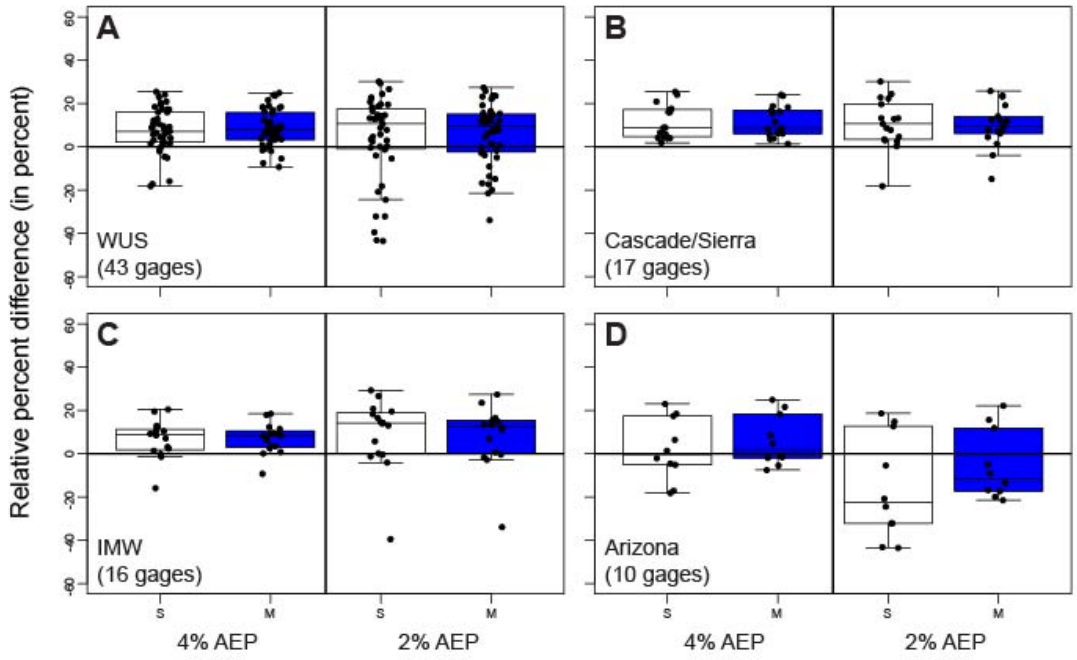

Figure 4.4: Relative percent differences (RPD) between the observed data and the LP3 quantile estimates for the single population (white boxes) and weighted mixed population (blue boxes) at the 4% and 2% AEPs for (A) the entire study region (western United States (WUS)), (B) the Cascade/Sierra Nevada mountain region, (C) the intermountain west (IMW), and (D) central Arizona. A positive (negative) RPD indicates that the quantile estimates are lower (higher) than the observed data at the corresponding probability of exceedance. The boundaries of the box plot represent the 25th and the 75th percentiles, while the line in between is the median. The whiskers extend to the 5th and 95th percentiles, respectively. The black circles represent the sample points off of which the boxplots are built.

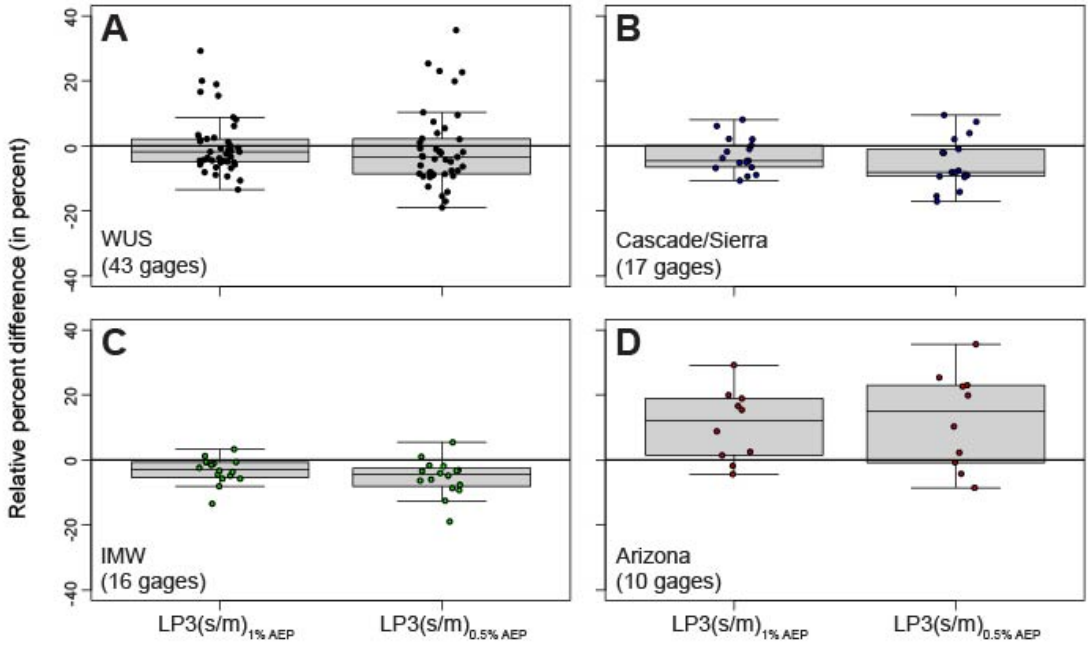

Figure 4.5: Relative percent differences (RPD) between the LP3 quantile estimates for the single population and weighted mixed population at the 1% and 0.5% AEPs for (A) the entire study region (western United States (WUS)), (B) the Cascade/Sierra Nevada mountain region, (C) the intermountain west (IMW), and (D) central Arizona. A positive (negative) RPD indicates the single population quantile estimate is greater (less) than the weighted mixed population quantile estimates. The boundaries of the box plot represent the 25th and the 75th percentiles, while the line in between is the median. The whiskers extend to the 5th and 95th percentiles, respectively. The black circles represent the sample points off of which the boxplots are built.

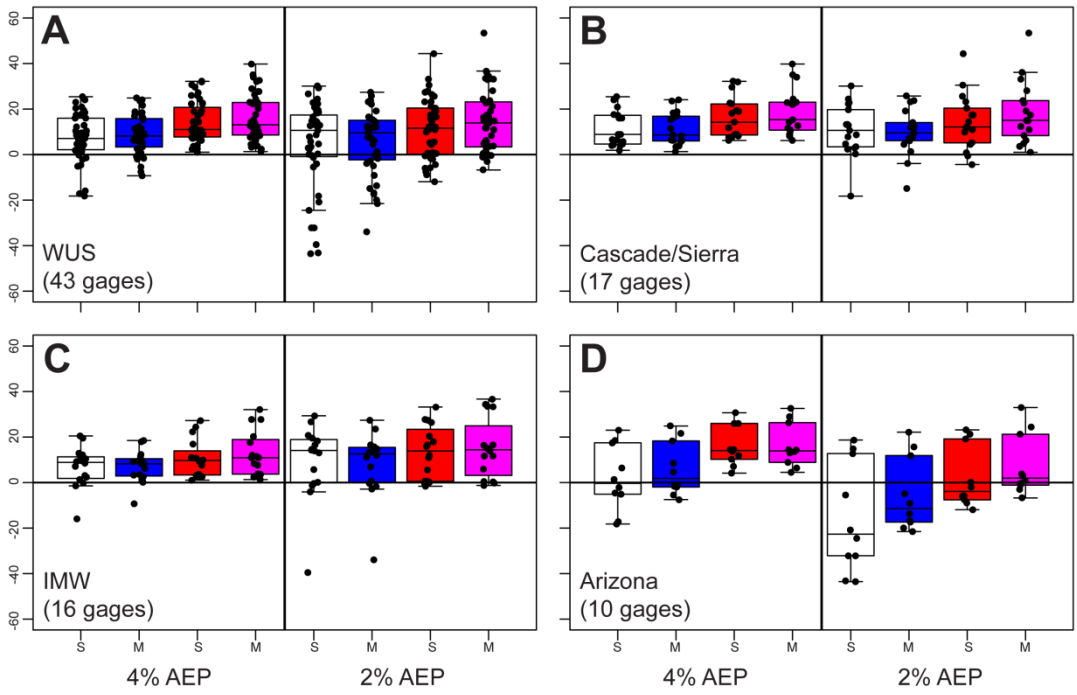

Figure 4.6: Relative percent differences (RPD) between the observed data and the LP3/GEV quantile estimates for the single population (white/red boxes) and weighted mixed population (blue/magenta boxes) at the 4% and 2% AEPs for (a) the entire study region (western United States (WUS)), (b) the Cascade/Sierra Nevada mountain region, (c) the intermountain west (IMW), and (d) central Arizona. A positive (negative) RPD indicates that the quantile estimates are lower (higher) than the observed data at the corresponding probability of exceedance. The boundaries of the box plot represent the 25th and the 75th percentiles, while the line in between is the median. The whiskers extend to the 5th and 95th percentiles, respectively. The black circles represent the sample points off of which the boxplots are built.

4.6 Tables Chapter

Table 4.1: Characteristics of the Characteristics of the 43 long-term USGS streamgaging records.

[**Region A**, Cascade/Sierra Nevada mountain ranges; **B**, Intermountain west (IMW); **C**, central Arizona; N_{PK}, number of annual peaks in single population; N_{AR}, number of AR-generated annual peaks; N_{nAR}, number of non-AR-generated annual peaks; ***PILF***, potentially influential low floods; $N_{PILF,PK}$, number of PILFs in single population; $Q_{MGB,PK}$, multiple Grubbs-Beck (MGB) low outlier criteria (cubic feet per second (cfs)) in single population; $N_{PILF,AR}$, number of PILFs in AR subpopulation; $Q_{MGB,AR}$, MGB low outlier criteria (cfs) in AR subpopulation; $N_{PILF,nAR}$, number of PILFs in non-AR subpopulation; $Q_{MGB,nAR}$, MGB low outlier criteria (cfs) in non-AR subpopulation; $\hat{g}_{PK}$ at-site single population log unit skew; $\hat{g}_{AR}$ at-site AR subpopulation log unit skew; $\hat{g}_{nAR}$ at-site non-AR subpopulation log unit skew; **AZ**, Arizona; **CA**, California; **ID**, Idado; **MT**, Montana; **NV**, Nevada; **OR**, Oregon; **WA**, Washington.]

Site	USGS ID	Region	Station Name	N_{PK}	N_{AR}	N_{nAR}	$N_{PILF,PK}$	$Q_{MGB,PK}$ (cfs)	$N_{PILF,AR}$	$Q_{MGB,AR}$ (cfs)	$N_{PILF,nAR}$	$Q_{MGB,nAR}$ (cfs)	$ghat_{PK}$	$ghat_{AR}$	$ghat_{nAR}$
1	05014500	B	Swiftcurrent Creek at Many Glacier, MT	98	28	70	0	0	0	0	0	0	1.705	1.291	1.99
2	06099500	B	Marias River near Shelby, MT	98	27	71	0	0	0	0	0	0	0.815	0.415	0.995
3	09444500	C	San Francisco River at Clifton, AZ	100	34	66	0	0	0	0	0	0	−0.085	−0.24	−0.199
4	09448500	C	Gila River at Head of Safford Valley, near Solomon, AZ	97	28	69	0	0	0	0	0	0	0.104	0.292	−0.062
5	09468500	C	San Carlos River near Peridot, AZ	81	33	48	1	565	0	0	1	565	−0.311	−0.194	−0.442
6	09484000	C	Sabino Creek near Tuscon, AZ	79	32	47	0	0	0	0	0	0	−0.156	−0.31	0.011
7	09490500	C	Black River near Fort Apache, AZ	53	26	27	0	0	0	0	0	0	−0.145	−0.419	0.214
8	09496500	C	Carrizo Creek near Show Low, AZ	60	28	32	0	0	0	0	0	0	−0.231	−0.741	−0.076
9	09497500	C	Salt River near Chrysotile, AZ	86	43	43	0	0	0	0	0	0	0.163	−0.446	0.577
10	09498500	C	Salt River near Roosevelt, AZ	86	44	42	0	0	0	0	0	0	0.215	−0.346	0.574
11	09499000	C	Tonto Creek above Gun Creek, near Roosevelt, AZ	70	39	31	0	0	9	6080	0	0	−0.282	−1.014	−0.201
12	09508500	C	Verde River below Tangle Creek, above Horseshoe Dam, AZ	86	55	31	0	0	0	0	0	0	−0.148	−0.409	0.189
13	10309000	A	East Fork Carson River near Gardnerville, NV	74	31	43	0	0	0	0	0	0	0.601	0.402	0.693
14	10310000	A	West Fork Carson River at Woodfords, CA	72	32	40	0	0	0	0	0	0	0.609	0.501	0.459
15	10311000	A	Carson River near Carson City, NV	72	32	40	0	0	0	0	0	0	0.549	0.281	0.53
16	10312000	A	Carson River near Fort Churchill, NV	100	28	72	0	0	0	0	0	0	0.057	−0.136	−0.155
17	10322500	B	Humbolt River at Palisade, NV	100	31	69	0	0	3	463	0	0	−0.339	−0.672	−0.075
18	10324500	B	Rock Creek near Battle Mountain, NV	65	26	39	0	0	0	0	0	0	−0.037	−0.196	0.219
19	10329500	B	Martin Creek near Paradise Valley, NV	89	41	48	0	0	0	0	0	0	0.28	0.226	0.065
20	10336660	A	Blackwood Creek near Tahoe City, CA	50	25	25	0	0	0	0	0	0	0.455	0.047	0.945
21	10343500	A	Sagehen Creek near Truckee, CA	57	31	26	0	0	0	0	0	0	0.135	−0.194	0.388
22	10396000	B	Donner Und Blitzen River near Frenchglen, OR	72	32	40	0	0	0	0	0	0	−0.341	−0.245	−0.315
23	11189500	A	South Fork Kern River near Onyx, CA	62	26	36	0	0	0	0	0	0	0.493	0.493	0.415
24	11266500	A	Merced River at Pohono Bridge near Yosemite, CA	94	25	69	0	0	0	0	0	0	0.603	−0.029	−0.243
25	11315000	A	Cole Creek near Salt Springs Dam, CA	66	41	25	0	0	0	0	0	0	0.277	0.275	0.653
26	11317000	A	Middle Fork Mokelumne River at West Point, CA	99	72	27	17	238	0	0	0	0	−0.61	−0.387	0.11
27	11335000	A	Cosumnes River at Michigan Bar, CA	104	73	31	16	2880	3	1700	0	0	−0.503	−0.449	−0.173
28	11348500	A	Pit River near Canby, CA	79	45	34	27	1330	13	1700	0	0	−0.915	−0.716	−0.135
29	11363000	A	Pit River at Big Bend, CA	100	69	31	48	7600	9	2220	2	1420	−0.726	−0.736	−0.355
30	11501000	A	Sprague River near Chiloquin, OR	90	37	53	0	0	0	0	0	0	0.083	0.213	−0.155
31	11502500	A	Williamson River below Sprague River near Chiloquin, OR	94	36	58	0	0	0	0	0	0	0.333	0.457	0.239
32	12411000	B	North Fork Coeur Dam Alene River above Shoshone Creek near Prichard, ID	60	28	32	0	0	0	0	2	3160	−0.131	−0.194	0.15
33	12413000	B	North Fork Coeur Dam Alene River at Enaville, ID	72	40	32	0	0	0	0	0	0	0.126	0.144	−0.152
34	12414500	B	St. Joe River at Calder, ID	90	34	56	0	0	0	0	9	11500	−0.016	0.039	−0.254
35	12451000	A	Stehekin River at Stehekin, WA	84	34	50	0	0	0	0	0	0	0.193	−0.093	0.217
36	12459000	A	Wenatchee River at Peshastin, WA	82	26	56	0	0	0	0	1	8800	0.286	0.712	−0.166
37	13082500	B	Goose Creek above Trapper Creek near Oakley, ID	92	25	67	0	0	0	0	0	0	0.637	0.143	0.833
38	13168500	B	Bruneau River near Hot Spring, ID	67	25	42	0	0	7	1510	1	704	−0.178	−0.535	−0.085
39	13258500	B	Weiser River near Cambridge, ID	72	37	35	9	2600	0	0	6	2600	−0.093	−0.423	−0.519
40	13333000	B	Grande Ronde River at Troy, OR	66	34	32	0	0	0	0	0	0	0.234	0.171	−0.57
41	13337000	B	Lochsa River near Lowell, ID	81	25	56	0	0	0	0	0	0	0.006	−0.1	0.015
42	14046500	B	John Day River at Service Creek, OR	81	51	30	0	0	0	0	0	0	−0.159	−0.118	−0.363
43	14048000	B	John Day River at McDonald Ferry, OR	104	64	40	0	0	0	0	0	0	−0.174	−0.16	−0.313

Table 4.2: Magnitude and frequency estimates based on the LP3 distribution using the EMA-MGB methodology (B17C) for the homogeneous single and heterogeneous weighted mixed populations for the 4%, 2%, 1%, and 0.5% AEPs for the 43 long-term USGS streamgaging records.

[**Region A**, Cascade/Sierra Nevada mountain ranges; **B**, Intermountain west (IMW); **C**, central Arizona; ***LP3***, log-Pearson Type III quantile estimates using the expected moments algorithm with the multiple Grubbs-Beck test (EMA-MGB) to identify potentially influential low floods (PILFs) (in cubic feet per second, cfs), for the single population (***s***) and the weighted mixed population (***m***) for the P% annual exceedance probability (AEP); **AZ**, Arizona; **CA**, California; **ID**, Idaho; **MT**, Montana; **NV**, Nevada; **OR**, Oregon; **WA**, Washington.]

Site	ID	Region	Station Name	$LP3_s$(4%)	$LP3_m$(4%)	$LP3_s$(2%)	$LP3_m$(2%)	$LP3_s$(1%)	$LP3_m$(1%)	$LP3_s$(0.5%)	$LP3_m$(0.5%)
1	05014500	B	Swiftcurrent Creek at Many Glacier, MT	2505	2543	3233	3323	4162	4347	5348	5689
2	06099500	B	Marias River near Shelby, MT	23,150	23,589	31,830	33,170	43,110	46,599	57,750	64,984
3	09444500	C	San Francisco River at Clifton, AZ	44,600	44,130	62,250	61,623	83,820	82,552	109,800	107,328
4	09448500	C	Gila River at Head of Safford Valley, near Solomon, AZ	67,910	66,294	97,030	97,956	134,200	140,080	180,900	196,437
5	09468500	C	San Carlos River near Periidot, AZ	39,510	38,630	51,380	51,143	64,560	65,730	79,040	82,393
6	09484000	C	Sabino Creek near Tuscon, AZ	8767	8428	12,080	11,565	16,030	15,636	20,700	20,865
7	09490500	C	Black River near Fort Apache, AZ	53,170	48,800	75,020	63,665	101,800	81,448	134,100	103,271
8	09496500	C	Carrizo Creek near Show Low, AZ	19,940	19,268	27,320	24,684	36,030	30,450	46,150	36,992
9	09497500	C	Salt River near Chrysotile, AZ	64,500	62,634	91,520	81,299	126,000	102,165	169,400	126,461
10	09498500	C	Salt River near Roosevelt, AZ	111,100	109,910	164,300	149,156	235,000	195,851	328,000	253,726
11	09499000	C	Tonto Creek above Gun Creek, near Roosevelt, AZ	75,990	67,828	103,700	82,103	136,100	96,380	173,200	111,577
12	09508500	C	Verde River below Tangle Creek, above Horseshoe Dam, AZ	108,500	105,257	150,400	141,122	200,800	183,151	260,700	233,700
13	10309000	A	East Fork Carson River near Gardnerville, NV	11,650	11,794	16,080	16,260	21,790	22,010	29,110	29,400
14	10310000	A	West Fork Carson River at Woodfords, CA	3688	3754	5124	5335	6985	7441	9385	10,213
15	10311000	A	Carson River near Carson City, NV	15,760	15,900	23,250	23,167	33,510	32,805	47,420	45,567
16	10312000	A	Carson River near Fort Churchill, NV	9612	9342	12,630	12,745	16,170	17,006	20,300	22,119
17	10322500	B	Humbolt River at Palisade, NV	6914	7072	8466	8786	10,090	10,582	11,780	12,490
18	10324500	B	Rock Creek near Battle Mountain, NV	3692	3480	5506	5284	7876	7935	10,920	11,862
19	10329500	B	Martin Creek near Paradise Valley, NV	3084	3077	4667	4743	6836	7065	9768	10,241
20	10336660	A	Blackwood Creek near Tahoe City, CA	2306	2336	3265	3173	4516	4242	6136	5679
21	10343500	A	Sagehen Creek near Truckee, CA	708	697	1014	964	1405	1292	1901	1720
22	10396000	B	Donner Und Blitzen River near Frenchglen, OR	3734	3651	4331	4279	4924	4953	5514	5685
23	11189500	A	South Fork Kern River near Onyx, CA	7708	7920	12,660	13,451	20,140	22,278	31,270	36,082
24	11266500	A	Merced River at Pohono Bridge near Yosemite, CA	15,140	15,518	19,420	20,801	24,550	26,726	30,690	33,184
25	11315000	A	Cole Creek near Salt Springs Dam, CA	4554	4509	5961	5864	7640	7474	9634	9432
26	11317000	A	Middle Fork Mokelumne River at West Point, CA	3874	3914	4810	5014	5758	6299	6709	7855
27	11335000	A	Cosumnes River at Michigan Bar, CA	46,810	47,079	58,750	59,808	71,190	73,883	84,050	90,485
28	11348500	A	Pit River near Canby, CA	7067	7113	8139	8318	9089	9504	9928	10,747
29	11363000	A	Pit River at Big Bend, CA	35,380	34,568	42,610	41,935	49,590	49,485	56,250	57,460
30	11501000	A	Sprague River near Chiloquin, OR	8098	7903	10,350	10,362	12,930	13,523	15,880	17,396
31	11502500	A	Williamson River below Sprague River near Chiloquin, OR	9420	9384	11,870	12,063	14,700	15,436	17,980	19,653
32	12411000	B	North Fork Coeur Dam Alene River above Shoshone Creek near Prichard, ID	14,190	14,544	16,290	16,811	18,420	19,097	20,570	21,385
33	12413000	B	North Fork Coeur Dam Alene River at Enaville, ID	40,610	40,563	48,450	49,029	56,900	58,238	66,010	68,266
34	12414500	B	St. Joe River at Calder, ID	32,020	32,128	36,390	39,105	40,830	46,296	45,340	53,924
35	12451000	A	Stehekin River at Stehekin, WA	18,100	18,159	20,290	20,551	22,530	22,928	24,830	25,341
36	12459000	A	Wenatchee River at Peshastin, WA	29,800	28,946	33,500	33,748	37,300	39,823	41,260	47,074
37	13082500	B	Goose Creek above Trapper Creek near Oakley, ID	1319	1352	1921	1993	2739	2898	3843	4197
38	13168500	B	Bruneau River near Hot Spring, ID	5267	5205	6153	6069	7059	6973	7988	7912
39	13258500	B	Weiser River near Cambridge, ID	11,020	10,976	12,690	12,490	14,390	13,910	16,130	15,249
40	13333000	B	Grande Ronde River at Troy, OR	34,080	34,409	39,900	41,347	46,110	48,771	52,760	56,776
41	13337000	B	Lochsa River near Lowell, ID	32,260	32,195	35,460	35,591	38,600	38,999	41,720	42,499
42	14046500	B	John Day River at Service Creek, OR	30,750	30,365	35,610	35,600	40,530	41,171	45,560	47,048
43	14048000	B	John Day River at McDonald Ferry, OR	29,360	29,043	33,730	33,599	38,130	38,366	42,570	43,281

Table 4.3: Magnitude and frequency estimates based on the GEV distribution using L moments for the homogeneous single and heterogeneous weighted mixed populations for the 4%, 2%, 1%, and 0.5% AEPs for the 43 long-term USGS streamgaging records.

[**Region A**, Cascade/Sierra Nevada mountain ranges; **B**, Intermountain west (IMW); **C**, central Arizona; ***GEV***, generalized extreme value quantile estimates using L moments (in cubic feet per second, cfs), for the single population (***s***) and the weighted mixed population (***m***) for the P% annual exceedance probability (AEP); **AZ**, Arizona; **CA**, California; **ID**, Idaho; **MT**, Montana; **NV**, Nevada; **OR**, Oregon; **WA**, Washington.]

Site	ID	Region	Station Name	GEV_s(4%)	GEV_m(4%)	GEV_s(2%)	GEV_m(2%)	GEV_s(1%)	GEV_m(1%)	GEV_s(0.5%)	GEV_m(0.5%)
1	05014500	B	Swiftcurrent Creek at Many Glacier, MT	2390	2293	3058	2897	3927	3702	5062	4765
2	06099500	B	Marias River near Shelby, MT	22,348	20,772	31,355	28,452	43,753	39,076	60,852	53,789
3	09444500	C	San Francisco River at Clifton, AZ	39,988	39,809	57,549	55,278	81,605	75,209	114,627	101,220
4	09448500	C	Gila River at Head of Safford Valley, near Solomon, AZ	61,163	62,632	89,862	87,576	130,308	119,116	187,417	160,029
5	09468500	C	San Carlos River near Periidot, AZ	36,432	36,164	48,673	46,939	63,895	59,756	82,874	75,156
6	09484000	C	Sabino Creek near Tuscon, AZ	7963	7251	11,415	9956	16,119	13,587	22,542	18,532
7	09490500	C	Black River near Fort Apache, AZ	42,168	42,482	56,381	52,648	73,895	63,370	95,534	75,644
8	09496500	C	Carrizo Creek near Show Low, AZ	17,299	17,296	23,962	22,869	32,617	29,493	43,884	37,600
9	09497500	C	Salt River near Chrysotile, AZ	55,364	56,186	75,411	71,377	100,965	88,713	133,619	108,550
10	09498500	C	Salt River near Roosevelt, AZ	93,101	94,587	131,547	123,292	182,830	157,247	251,384	198,671
11	09499000	C	Tonto Creek above Gun Creek, near Roosevelt, AZ	61,603	61,375	80,859	77,128	103,979	95,216	131,816	116,430
12	09508500	C	Verde River below Tangle Creek, above Horseshoe Dam, AZ	91,047	89,271	123,207	117,137	163,655	150,629	214,654	192,230
13	10309000	A	East Fork Carson River near Gardnerville, NV	10,951	10,637	15,249	14,295	21,038	18,989	28,851	25,002
14	10310000	A	West Fork Carson River at Woodfords, CA	3478	3210	4901	4403	6850	5988	9521	8088
15	10311000	A	Carson River near Carson City, NV	14,151	13,701	20,948	19,255	30,697	26,729	44,704	36,441
16	10312000	A	Carson River near Fort Churchill, NV	9356	8834	12,846	12,023	17,422	16,155	23,435	21,572
17	10322500	B	Humbolt River at Palisade, NV	6868	6796	8790	8732	11,073	11,063	13,790	13,879
18	10324500	B	Rock Creek near Battle Mountain, NV	2843	2623	3966	3483	5429	4595	7338	6083
19	10329500	B	Martin Creek near Paradise Valley, NV	2654	2538	3968	3644	5850	5134	8552	7114
20	10336660	A	Blackwood Creek near Tahoe City, CA	2075	2046	2881	2660	3948	3400	5362	4331
21	10343500	A	Sagehen Creek near Truckee, CA	612	584	859	793	1185	1051	1618	1380
22	10396000	B	Donner Und Blitzen River near Frenchglen, OR	3642	3591	4227	4212	4823	4897	5433	5654
23	11189500	A	South Fork Kern River near Onyx, CA	6346	5633	10,075	8436	15,801	12,354	24,602	17,922
24	11266500	A	Merced River at Pohono Bridge near Yosemite, CA	14,793	14,853	19,264	19,558	24,891	24,866	31,992	30,945
25	11315000	A	Cole Creek near Salt Springs Dam, CA	4394	4233	5859	5468	7723	6964	10,101	8821
26	11317000	A	Middle Fork Mokelumne River at West Point, CA	3659	3556	4788	4591	6148	5822	7792	7360
27	11335000	A	Cosumnes River at Michigan Bar, CA	44,107	43,483	57,379	56,426	73,245	71,788	92,264	90,199
28	11348500	A	Pit River near Canby, CA	6933	6935	8477	8341	10,168	9816	12,025	11,392
29	11363000	A	Pit River at Big Bend, CA	33,512	33,031	42,362	41,869	52,516	52,167	64,198	64,276
30	11501000	A	Sprague River near Chiloquin, OR	7716	7463	9861	9536	12,408	12,027	15,442	15,093
31	11502500	A	Williamson River below Sprague River near Chiloquin, OR	9229	8907	11,718	11,311	14,700	14,316	18,283	18,076
32	12411000	B	North Fork Coeur Dam Alene River above Shoshone Creek near Prichard, ID	14,340	14,335	16,789	16,664	19,402	19,022	22,199	21,427
33	12413000	B	North Fork Coeur Dam Alene River at Enaville, ID	40,476	39,833	48,994	47,975	58,580	57,004	69,400	66,970
34	12414500	B	St. Joe River at Calder, ID	31,873	31,338	36,535	38,316	41,389	45,963	46,462	54,241
35	12451000	A	Stehekin River at Stehekin, WA	18,146	18,260	20,443	20,812	22,808	23,520	25,251	26,341
36	12459000	A	Wenatchee River at Peshastin, WA	30,018	28,556	34,137	32,883	38,507	38,525	43,161	45,626
37	13082500	B	Goose Creek above Trapper Creek near Oakley, ID	1206	1127	1756	1587	2533	2220	3631	3105
38	13168500	B	Bruneau River near Hot Spring, ID	5223	5196	6160	6145	7155	7177	8216	8318
39	13258500	B	Weiser River near Cambridge, ID	10,715	10,848	12,102	12,525	13,452	14,191	14,772	15,866
40	13333000	B	Grande Ronde River at Troy, OR	33,931	33,847	39,900	40,670	46,383	48,018	53,449	55,931
41	13337000	B	Lochsa River near Lowell, ID	32,074	32,061	34,953	35,077	37,669	38,044	40,239	40,931
42	14046500	B	John Day River at Service Creek, OR	30,382	30,125	35,260	35,314	40,307	40,875	45,550	46,827
43	14048000	B	John Day River at McDonald Ferry, OR	29,153	28,939	33,648	33,718	38,272	38,819	43,046	44,245

Chapter 5

FUTURE STUDIES AND CONCLUSIONS

5.1 Future Studies

In addition to different flood agents, flood frequency in the western United States is complicated by changes in the climate system that occur on decadal and multi-decadal time scales. Extreme fluctuations in annual peak flows have recently been seen in the span of the past five years throughout this area. This is highlighted by the historic 2011-2015 drought in California, followed by the onslaught of strong winter storms (ARs), producing a combination of heavy rain and snowfall, flooding, and mudslides across California during the 2016/2017 winter season. These shifts, when analyzed in a long-term context, highlight the alternating of "flood rich" and "flood poor" periods. Figure 5.1 provides a clear example of the changes in flooding that occur at decadal and multi-decadal time scales. The Gila River at the head of Safford Valley, Arizona (station 09448500) shows the alternation of "flood rich" and "flood poor" decades. This long-term annual peak flow record displays what has been described as a two-state system representing a nonstationary [mixed] population with a longer persistence of wetter and dryer periods (Hirsch, 2011; England et al., 2018). This behavior can be tied to changes in the climate system that take place across different time scales, with El Niño – Southern Oscillation (ENSO), the Pacific Decadal Oscillation (PDO), the Pacific North American teleconnection pattern (PNA) and the Pacific-Japan teleconnection pattern (PJ) likely playing a major role. An improved understanding of the relationship between climatic drivers and flooding would allow for the development of climate-driven flood frequency analysis, leading to more robust and realistic estimates of design values for flood structure and water-related projects.

As mentioned in Chapter 1, the conventional assumptions for performing flood frequency analyses are that the annual time series is a representative time sample of random homogeneous events and that the stochastic processes that generate floods are assumed to be stationary or invariant in time (IACWD, 1982). The Work Group that updated the guidelines for performing flood frequency in B17C 'did not evaluate methods to account for climate variability in flood frequency.' In the 'Future Studies' section of Bulletin 17C, the Work Group identified the need for two (among several other) important topics of study: '[1] the identification and treatment of mixed distributions, including those based on hydrometeorological or hydrological conditions; and [6] guides for estimating dynamic flood frequency curves that vary with time, incorporating climate indices...and addressing potential nonstationary climate conditions' (England et al, 2018). The research in this thesis specifically address topic [1] in the 'Future Studies' section, yet topic [6] represents another type of mixed population that, if addressed, will help move away from the detection of nonstationarities and move towards the attribution of the observed changes. Analyses related to process-driven flood frequency analysis that incorporates climate information to describe the interannual and multidecadal variability should be evaluated.

5.2 Summary and Conclusions

As described in Chapter 1, the conventional assumptions for performing flood frequency analysis as recommended in the B17B federal guidelines framework established more than 35 years ago, are that the annual time series is a representative time sample of random homogeneous events and that the stochastic processes that generate floods are assumed to be stationary or invariant in time. However, the Work Group that recently updated the guidelines for performing flood frequency in B17C "did not conduct an evaluation of [mixed population] procedures…" In

the ‘Future Studies’ section of B17C, the Work Group identified the need for an (among several other) important topic of study: ‘[1] the identification and treatment of mixed distributions, including those based on hydrometeorological or hydrological conditions.’

With the diverse and complex flood hydrology found throughout the western United States (Chapter 1), the need to objectively identify the hydrometeorological drivers responsible for different flood-generating mechanisms is paramount to address the first topic related to mixed population distributions in the ‘Future Studies’ section of B17C. Both B17B and B17C recognize the difficulties in determining flood frequency estimates with streamflow records that contain flood peaks coming from different flood-generating mechanisms. They recommend developing separate frequency curves when the hydrometeorologic mechanisms that generated the annual peak flows can be separated into distinct populations. Yet challenges arise when trying to consistently quantify the physical process that generated the observed flows.

In the western United States, atmospheric rivers (ARs) and tropical cyclones (TCs) and their remnants are primarily responsible for extreme precipitation and flooding. ARs are long, narrow, river-like features that transport atmospheric water vapor poleward from the tropics to the midlatitudes. ARs have been frequently associated with heavy rainfall, large snowfall totals, and destructive flooding mostly in the November-April period. During these cold-season months, the storms interact with topographic barriers (e.g., the Cascade and Sierra Nevada ranges) leading to orographically-enhanced precipitation. I identify six main areas in which flooding is impacted by ARs at varying degrees throughout the western United States (Figure 2.2). The Pacific Northwest and the northern California coast have the highest fraction of AR-generated peaks (~80–100%), while eastern Montana, Wyoming, Utah, Colorado, and New Mexico have nearly no impacts from ARs. The individual regions of the central Columbia River Basin in the

Pacific Northwest, the Sierra Nevada, the central and southern California coast, and central Arizona all show a mixture of 30–70% AR-generated flood peaks. Analyses related to the largest flood peaks on record highlight the strong impact of ARs on flood hydrology in this region (Figure 2.4). Conversely, TC events play a limited role in controlling the upper tail of the flood peak distributions across the western United States. Southern California, Arizona, southernmost Nevada and Utah, southern and western New Mexico, central Colorado, and Texas have the highest fractional contributions of TC-event-generated annual maximums flows (~5-14%) (Figure 3.2).

Using the hydrometeorological information from Chapter 2, I develop a statistical framework to perform a process-driven flood frequency analysis using the AR/non-AR-generated annual peak flows identified at 43 long-term USGS streamgages in the western United States from the mixed population regions in Chapter 2. I use a simulation framework to perform flood frequency analyses in terms of mixed distributions and quantify the corresponding uncertainties by accounting for mixed populations. Sites with notably different quantile estimates in the upper tail of the distribution between the single (homogeneous) and the weighted (heterogeneous) population methodologies are found when (i) potentially influential low floods (PILFS) are identified and/or (ii) when the composite distribution contains markedly different at-site log-unit skews (shape parameter) among the AR/non-AR subpopulations compared to the single homogeneous population.

The new methodological developments described in my thesis specifically address one limitation identified in the 'Future Studies' section in the updated federal guidelines of B17C. The improvements associated with this flood frequency analysis methodology not only leads to a better characterization of the observational records, but it also allows us to incorporate the

physical processes, resulting in process-based flood frequency analysis. By addressing topic [1] of the above-listed 'Futures Studies' section, the studies conducted in this thesis add to a growing body of literature continuing to build a more solid framework based on physical processes to be used in the revisions of Bulletin 17C.

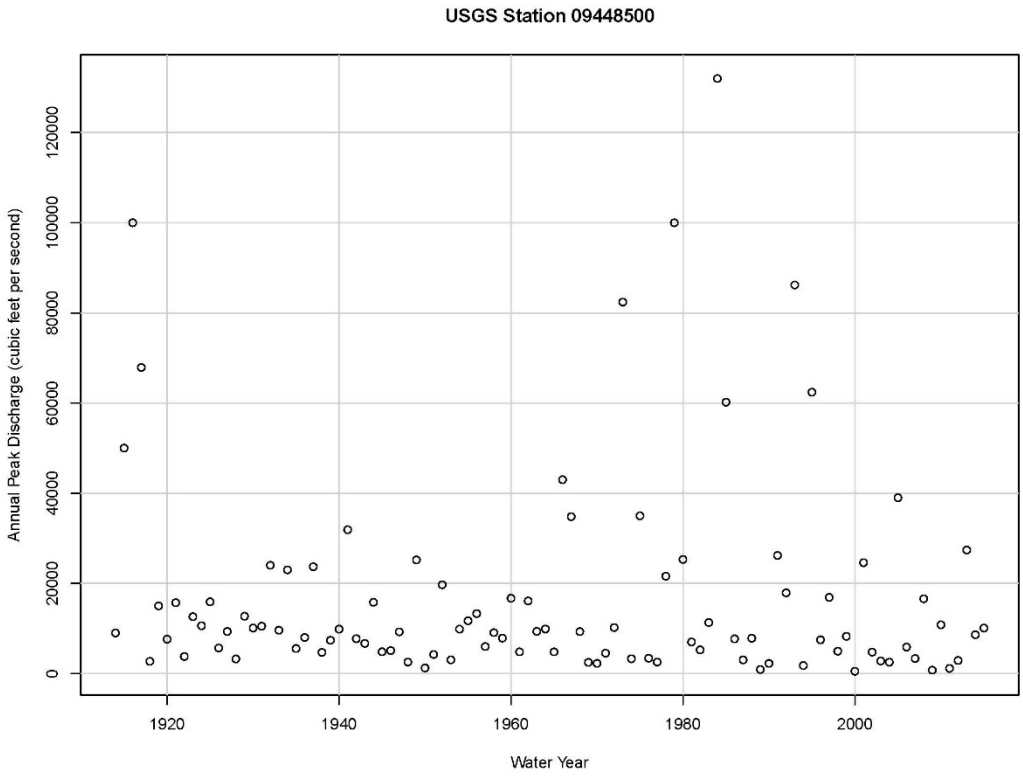

Figure 5.1: The annual peak discharge streamflow record for Gila River at the head of Safford Valley, Arizona (station 09448500).

REFERENCES

Alila, Y., and Mtiraoui, A., (2002), Implications of heterogeneous flood-frequency distributions on traditional stream-discharge prediction techniques, Hydrological Processes, 16(5), 1065-1084.

Archfield, S.A., Hirsch, R.M., Viglione, A., Bloschl, G., (2016), Fragmented patterns of flood change across the United States, Geophysical Research Letters, 43(19), 10,232-210,239, doi: 10.1002/2016GL070590.

Aryal, Y.N., Villarini, G., Zhang, W., and Vecchi, G.A., (2018), Long term changes in flooding and heavy rainfall associated with North Atlantic tropical cyclones: Roles of the North Atlantic Oscillation and El Niño-Southern Oscillation, Journal of Hydrology, 559, 698-710.

Asquith, W.H., Kiang, J.E., and Cohn, T.A., (2017), Application of at-site peak-streamflow frequency analyses for very low annual exceedance probabilities, U.S. Geological Survey Scientific Investigations Report, 2017-5038, 106 p.

Barlow, M., (2011), Influence of hurricane-related activity on North American extreme precipitation, Geophysical Research Letters, 38(4).

Barros, A. P., Duan, Y., Brun, J., and Medina, Jr., M.A., (2014), Flood nonstationarity in the Southeast and Mid-Atlantic Regions of the United States, Journal of Hydrologic Engineering, 19(10), 05014014.

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, doi:10.1002/2016WR019064.

Barth, N.A., Villarini, G., White, K., (2018a), Contribution of eastern North Pacific tropical cyclones and their remnants on flooding in the western United States, International Journal of Climatology;1–6.

Barth, N.A., Villarini, G., and White, K., (2018b), Accounting for mixed populations in flood frequency analysis: A Bulletin 17C perspective, Journal of Hydrologic Engineering, (in press).

Berghuijs, W.R., M. Sivapalan, C.J. Hutton, and R.A. Woods, (2016), Dominant flood generating mechanisms across the United States, Geophysical Research Letters, 43, 4382–4390.

Bernhardt, D. (2008), Glacier National Park flooding November 2006, NWS Western Reg. Tech. Attachment 08-23, 15 pp. [Available at https:// www.weather.gov/media/wrh/online_publications/talite/talite0823.pdf.]

Brands, S., Gutierrez, J. M., and San-Marın, D., (2016), Twentieth-century atmospheric river activity along the west coasts of Europe and North America: Algorithm formulation, reanalysis

uncertainty and links to atmospheric circulation patterns, Climate Dynamics, 1–25, doi: 10.1007/s00382-016-3095-6.

Climate Prediction Center (CPC) Unified Gauge-Based Analysis of Daily Precipitation (CPC UD United Precipitation data: https://www.esrl.noaa.gov/psd/data/gridded/data.unified.daily.conus.html), (accessed on 01 September 2017).

Cohn, T. A., England, J. F., Berenbrock, C. E., Mason, R. R., Stedinger, J. R., and Lamontagne, J. R., (2013), A generalized Grubbs-Beck test statistic for detecting multiple potentially influential low outliers in flood series, Water Resources Research, 49, 5047–5058, doi:10.1002/wrcr.20392.

Coles, S. (2001), An introduction to statistical modeling of extreme values, Springer.
Compo, G. P., et al., (2011), The twentieth century reanalysis project, Q. J. R. Meteorological Society, 137 (654), 1-28.

Copper, R. M., (2005), Estimates of peak discharges for rural, unregulated streams in western Oregon, U.S. Geological Survey Scientific Investigations Report, 2005-5116, 134 p.

Corbosiero, K., Dickinson, M., and Bosart, L., (2009), The contribution of eastern North Pacific tropical cyclones to the rainfall climatology of the southwest United States. Monthly Weather Review, 137, 2415–2435, doi: 10.1175/2009MWR2768.1.

CPC Global Unified Precipitation data provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at https://www.esrl.noaa.gov/psd/

Czajkowski, J., Villarini, G., Michel-Kerjan, E., and Smith, J.A., (2013), Determining tropical cyclone inland flooding loss on a large scale through a new flood peak ratio-based methodology, Environmental Research Letters, 8(4), 044056.

Czajkowski, J., Villarini, G., Montgomery, M., Michel-Kerjan, E., and Goska, R., (2017), Assessing current and future freshwater flood risk from North Atlantic tropical cyclones via insurance claims, Scientific Reports, 7, 1-10, 2017.

Dettinger, M. D., (2013), Atmospheric Rivers as Drought Busters on the U.S. West Coast, Journal of Hydrometeorology, 14(6), 1721–1732.

Dettinger, M. D., Ralph, F. M., Das, T., Neiman, P. J., and Cayan, D. R., (2011), Atmospheric rivers, floods and the water resources of California, Water, 3(2), 445–478.

England, J.F., Jr., Cohn, T.A., Faber, B.A., Stedinger, J.R., Thomas, W.O., Jr., Veilleux, A.G., Kiang, J.E., and Mason, R.R., Jr., (2018), Guidelines for determining flood flow frequency—Bulletin 17C : U.S. Geological Survey Techniques and Methods, book4, chap.B5, 148p., https://doi.org/10.3133/tm4B5.

Farfán, L.M., and Zehnder, J.A., (2001), An analysis of the landfall of Hurricane Nora (1997), Monthly Weather Review, 129(8), 2073-2088.

Fors, J.R., (1977), Tropical Cyclone Kathleen, NOAA Technical Memorandum NWS WR-114, 38.

Gotvald, A.J., Barth, N.A., Veilleux, A.G., and Parrett, C., (2012), Methods for determining magnitude and frequency of floods in California, based on data through water year 2006, U.S. Geological Survey Scientific Investigations Report 2012-5113, 38.

Guan, B., Molotch, N. P., Waliser, D. E., Fetzer, E. J., and Neiman, P. J., (2013), The 2010/2011 snow season in California's Sierra Nevada: Role of atmospheric rivers and modes of large-scale variability, Water Resources Research, 49, 6731–6743, doi:10.1002/wrcr.20537.

Hammond, S. E., and Harmon, J. G. , (1998), Publications document floods of January 1997 in California and Nevada, Rep. 093-98. [Available at https://pubs.er.usgs.gov/publication/fs09398.]
Helsel, D. R., and Hirsch, R. M., (1992), Statistical Methods in Water Resources, 326 pp., Elsevier, New York.

Hirsch, R. M., (2011), A perspective on nonstationarity and water management, Journal of the American Water Resources Association, 436-446, DOI: 10.1111/j.1752-1688.2011.00539.x.

Hirsch, R. M., (2017), Timothy A. Cohn (1957–2017), Eos, 98, Published on 23 June 2017, doi.org/10.1029/2017EO076523.

Hirsch, R. M., and Stedinger, J. R., (1987), Plotting positions for historical floods and their precision, Water Resources Research, 23(4), 715-727.

Hirschboeck, K.K., (1987), Hydroclimatically defined mixed distributions in partial duration flood series, in Singh, V.P., ed., Hydrologic Frequency Modeling, D. Reidel Publishing Company.

Hirschboeck, K. K., (1988), Flood hydroclimatology, in flood gomorphology, edited by R. Victor, et al., John Wiley and Sons, Inc., New York, New York, 49 p.

Hirschboeck, K.K., (1991), Climate and floods, in National Water Summary 1988-1989 Hydrologic Events and Floods and Droughts: U.S. Geological Survey Water Supply Paper 2375, 67-88.

Hirschboeck, K.K.,(2009), Flood flows of the San Pedro River. Pages 300-312 in Stromberg, Juliet C. and Tellman, Barbara (eds.) Ecology and Conservation of the San Pedro River. University of Arizona Press: Tucson, 656 pp.

Hjalmarson, H.W., (1989), Flood of October 1983 and history of flooding along the San Francisco River, Clifton, Arizona, U.S. Geological Survey Water Resources Investigations Report 85-4225-B, 48.

Hosking, J.R.M., 1990, L-moments—Analysis and estimation of distributions using linear combinations or order statistics, Journal of Royal Statistical Society, series B, v. 52, no. 1, p. 105–124.

House, P.K. and Hirschboeck, K.K., (1995), Hydroclimatological and paleohydrological context of extreme winter flooding in Arizona, 1993. In: Larson, R.A. and Slosson, J.E. (Eds.) Storm Induced Geological Hazards: Case Histories from the 1992 1993 Winter Storm in Southern California and Arizona.Reviews in Engineering Geology, Vol. 11. Boulder, CO: Geological Societyof America, pp. 1–24.

Interagency Advisory Committee on Water Data, (1982), Guidelines for Determining Flood Flow Frequency, Hydrology Subcommittee Bulletin 17B,U.S. Department of the Interior Geological Survey, Office of Water Data Coordination, Reston, Virginia, 28 p.

Jackson, D. L., Hughes, M., and Wick, G. A., (2016), Evaluation of landfalling atmospheric rivers along the U.S. West Coast in reanalysis data sets, Journal of Geophysical Research Atmospheres, 121, 2705–2718, doi:10.1002/2015JD024412.

Khouakhi, A., Villarini, G. and Vecchi, G., (2017), Contribution of Tropical Cyclones to Rainfall at the Global Scale. Journal of Climate, 30, 359–372, doi: 10.1175/JCLI-D-16-0298.1.

Klemeš, V., (1974), Some problems in pure and applied stochastic hydrology. Proceedings of the Symposium on Statistical Hydrology Miscellaneous Publication, No. 1275 Washington, DC: U.S. Department of Agriculture, 2–15.

Lamontagne, J. R., Stedinger, J. R., Yu, X., Whealton, C. A., and Xu, Z., (2016), Robust flood frequency analysis: Performance of EMA with multiple Grubbs-Beck outlier tests, Water Resources Research, 52, 3068–3084, doi:10.1002/2015WR018093.

Landsea, C. and Franklin, J., (2013), Atlantic hurricane database uncertainty and presentation of a new database Format. Monthly Weather Review, 141, 3576–3592, doi: 10.1175/MWR-D-12-00254.1.

Lang, M., Ouarda, T. B. M. J., and Bobée, B., (1999), Towards operational guidelines for over-threshold modeling, Journal of Hydrology, 225, 103–117.

Lavers, D. A., and Villarini, G., (2013), Atmospheric rivers and flooding over the central United States, Journal of Climate, 26(20), 7829–7836.

Lavers, D. A., and Villarini, G., (2015), The contribution of atmospheric rivers to precipitation in Europe and the United States, Journal of Hydrology, 522, 382–390.

MacDonald, L. H., and Hoffman, J. A., (1995), Causes of peak flows in northwestern Montana and northeastern Idaho, Water Resources Bulletin, 31(1), 79–95.

Mastin, M.C., Konrad, C.P., Veilleux, A.G., and Tecca, A.E., (2016), Magnitude, frequency, and trends of floods at gaged and ungaged sites in Washington, based on data through water year 2014, U.S. Geological Survey Scientific Investigations Report 2016-5118, 80.

National Oceanic and Atmospheric Administration (NOAA), (2010): A history of significant weather events in southern California: Organized by weather type, (http://www.wrh.noaa.gov/sgx/document/weatherhistory.pdf) (accessed on 06 February 2017).

National Oceanic and Atmospheric Administration (NOAA) Weather Prediction Center (WPC) (http://www.wpc.ncep.noaa.gov/tropical/rain/tcpointofentry.html) (accessed on 01 September 2017).

Neiman, P. J., Ralph, F. M., Wick, G. A., Lundquist, J. D., and Dettinger, M. D., (2008), Meteorological characteristics and overland precipitation impacts of atmospheric rivers affecting the west coast of North America based on eight years of SSM/I satellite observations, Journal of Hydrometeorology, 9(1), 22–47.

Neiman, P. J., Schick, L. J., Ralph, F. M., Hughes, M., and Wick, G. A., (2011), Flooding in western Washington: The connection to atmospheric rivers, Journal of Hydrometeorology, 12(6), 1337–1358.

Neiman, P. J., Ralph, F. M., Moore, B. J., Hughes, M. , Mahoney, K. M., Cordeira, J. M., and Dettinger, M. D., (2013), The landfall and inland penetration of a flood-producing atmospheric river in Arizona. Part I: Observed synoptic-scale, orographic, and hydrometeorological characteristics, Journal of Hydrometeorology, 14(2), 460–484.

Paretti, N.V., Kennedy, J.R., Turney, L.A., and Veilleux, A.G., (2014), Methods for estimating magnitude and frequency of floods in Arizona, developed with unregulated and rural peak-flow data through water year 2010, U.S. Geological Survey Scientific Investigations Report 2014-5211, 61.

Parrett, C., and Johnson, D. R., (2003), Methods for estimating flood frequency in Montana based on data through water year 1998, U.S. Geological Survey Water-Resources Investigations Report, 03-4308, 109 p.

Parrett, C., Veilleux, A., Stedinger, J. R., Barth, N. A., Knifong, D. K., and Ferris, J. C., (2010), Regional skew for California, and flood frequency for selected sites in the Sacramento-San Joaquin River Basin, based on data through water year 2006, U.S. Geological Survey Scientific Investigations Report, 2010-5260, 94 p.

Prat, O.P., and Nelson, B.R., (2013), Mapping the world's tropical cyclone rainfall contribution over land using the TRMM Multi-satellite Precipitation Analysis, Water Resources Research, 49(11), 7236-7254.

Rahman, A. S., Haddad, K., and Rahman, A., (2014), Identification of outliers in flood frequency analysis: Comparison of original and multiple Grubbs-Beck Test, World Academy of Science,

Engineering and Technology: International Journal of Environmental, Chemical, Ecological, Geological and Geophysical Engineering, 8(12), 840–848.

Ralph, F. M., and Dettinger, M. D., (2011), Storms, floods, and the science of atmospheric rivers, Eos Transactions, AGU, 92(32), 265–266.

Ralph, F. M., and Dettinger, M. D., (2012), Historical and national perspectives on extreme west coast precipitation associated with atmospheric rivers during December 2010, Bulletin of the American Meteorological Society, 93(6), 783–790.

Ralph, F. M., Neiman, P. J., Wick, G. A., Gutman, S. I., Dettinger, M. D. , Cayan, D. R. and White, A. B., (2006), Flooding on California's Russian River: Role of atmospheric rivers, Geophysical Research Letters, 33, L13801, doi:10.1029/2006GL026689.

Rappaport, E.N., (2014), Fatalities in the United States from Atlantic tropical cyclones: New data and interpretation, Bulletin of the American Meteorological Society, 95, 341–346, doi:10.1175/BAMS-D-12-00074.1.

Ritchie, E., Wood, K., Gutzler, D., and White, S., (2011), The influence of Eastern Pacific tropical cyclone remnants on the southwestern United States. Monthly Weather Review, 139, 192–210, doi: 10.1175/2010MWR3389.1.

Rivera, E. R., Dominguez, F., and Castro, C. L., (2014), Atmospheric rivers and cool season extreme precipitation events in the Verde River Basin of Arizona, Journal of Hydrometeorology, 15(2), 813–829.

Roeske, R.H., Garrett, J.M., Eychaner, J.H., (1989), Floods of October 1983 in southeastern Arizona, U.S, Geological Survey, Water Resources Investigations Report 85-4225-C, 82.

Rutz, J. J., and Steenburgh, W. J., (2012), Quantifying the role of atmospheric rivers in the interior western United States, Atmospheric Science Letters,13(4), 257–261.

Rutz, J. J., Steenburgh, W. J., and Ralph, F. M., (2014), Climatological characteristics of atmospheric rivers and their inland penetration over the western United States, Monthly Weather Review, 142(2), 905–921.

Rutz, J. J., Steenburgh, W. J., and Ralph, F. M., (2015), The inland penetration of atmospheric rivers over western North America: A Lagrangian analysis, Monthly Weather Review, 143(5), 1924–1944.

Sando, S.K., McCarthy, P.M., Dutton, D.M., (2015), Peak-flow frequency analyses and results based on data through water year 2011 for selected streamflow-gaging stations in or near Montana, U.S. Geological Survey Scientific Investigations Report, 2015-5019-C, 38 p.
Smith, J.A., Villarini, G. and Baeck, M. L., (2011), Mixture distributions and the hydroclimatology of extreme rainfall and flooding in the eastern United States, Journal of Hydrometeorology, 12(2), 294-309.

Smith, W., (1986), The effects of eastern north Pacific tropical cyclones on the southwestern United States, National Oceanic and Atmospheric Administration Technical Memorandum NWS WS-197, 229.

Stedinger, J.R. and Griffis, V.W., (2008), Flood frequency analysis in the United States: Time to update, Journal of Hydrologic Engineering, 13(4), 199-204.

Thomas, B.E., Hjalmarson, H.W., and Waltemeyer, S.D., (1997), Methods for estimating magnitude and frequency of floods in the southwestern United States, U.S. Geological Survey Water-Supply Paper 2433, 205.

Villarini, G., (2016), On the seasonality of flooding across the continental United States, Advances in Water Resources, 87, 80-91.

Villarini, G., Goska, R., Smith, J.A., and Vecchi, G.A., (2014a), North Atlantic tropical cyclones and U.S. flooding, Bulletin of the American Meteorological Society, 95, 1381–1388, doi:10.1175/BAMS-D-13-00060.1.

Villarini, G., Lavers, D.A., Scoccimarro, E., Zhao, M., Wehner, M.F., Vecchi, G.A., Knutson, T.R., and Reed, K.A., (2014b), Sensitivity of tropical cyclone rainfall to idealized global-scale forcings. J. Climate, 27, 4622–4641, doi:10.1175/JCLI-D-13-00780.1.

Villarini, G. and Slater, L., (2017), Climatology of flooding in the United States, Oxford Research Encyclopedia of Natural Hazard Science, 37 p., doi:10.1093/acrefore/9780199389407.013.123.

Webb, R.H. and Betancourt, J.L., (1992), Climatic variability and flood frequency of the Santa Cruz River, Pima County, Arizona, U.S. Geological Survey Water-Supply Paper 2379, 47.

Wood, K.M., and Ritchie, E.A., (2013), An updated climatology of tropical cyclone impacts on the southwestern United States, Monthly Weather Review, 141(12), 4322-4336.

Zhu, Y., and Newell, R. E., (1998), A proposed algorithm for moisture fluxes from atmospheric rivers, Monthly Weather Review, 126(3), 725-735.

SUPPLEMENTARY MATERIAL

S.1 Figures

Figures S.1-S.43: Flood frequency plots for the single homogeneous population and AR/non-AR subpopulations based on the LP3 distribution using the EMA-MGB methodology and the GEV distribution using L moments. And Figures S.44-S.86: Graphical comparisons between the single and the weighted mixed population flood frequency curves based on the LP3 (EMA-MGB) methodology and GEV (L moments) approach. Confidence intervals for the single and weighted mixed population are based on the LP3 (EMA-MGB) probability distribution.

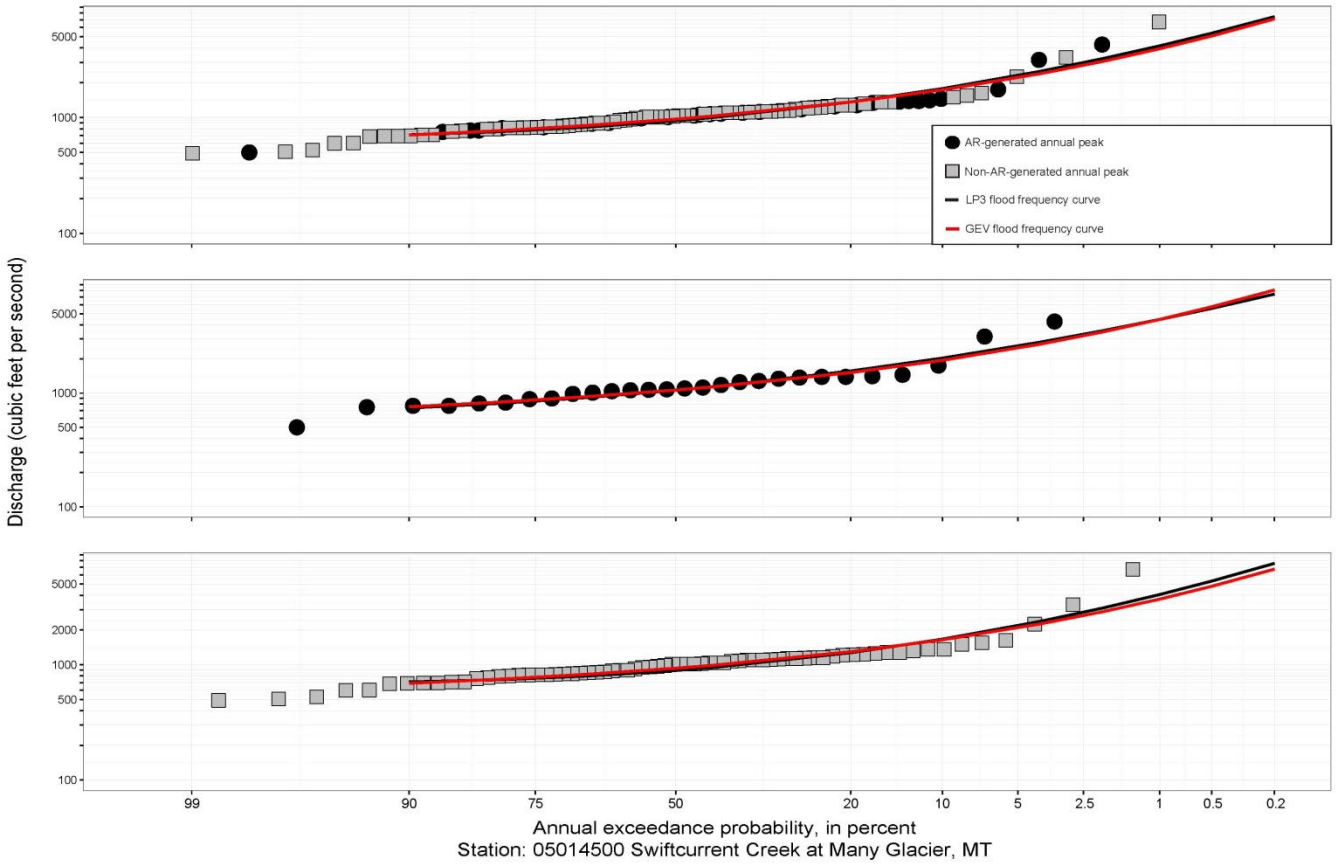

Figure S.1: Fitted frequency curves for (i) a single homogeneous population (top panel) comprised of annual peak flows generated from AR (black circles) /non-AR (grey squares) generated peaks, (ii) the subpopulation of AR-generated (middle panel), and (iii) the subpopulation of non-AR-generated peaks (bottom panel) for the log-Pearson Type III (LP3) distribution (black line) and the generalized extreme value (GEV) distribution (red line).

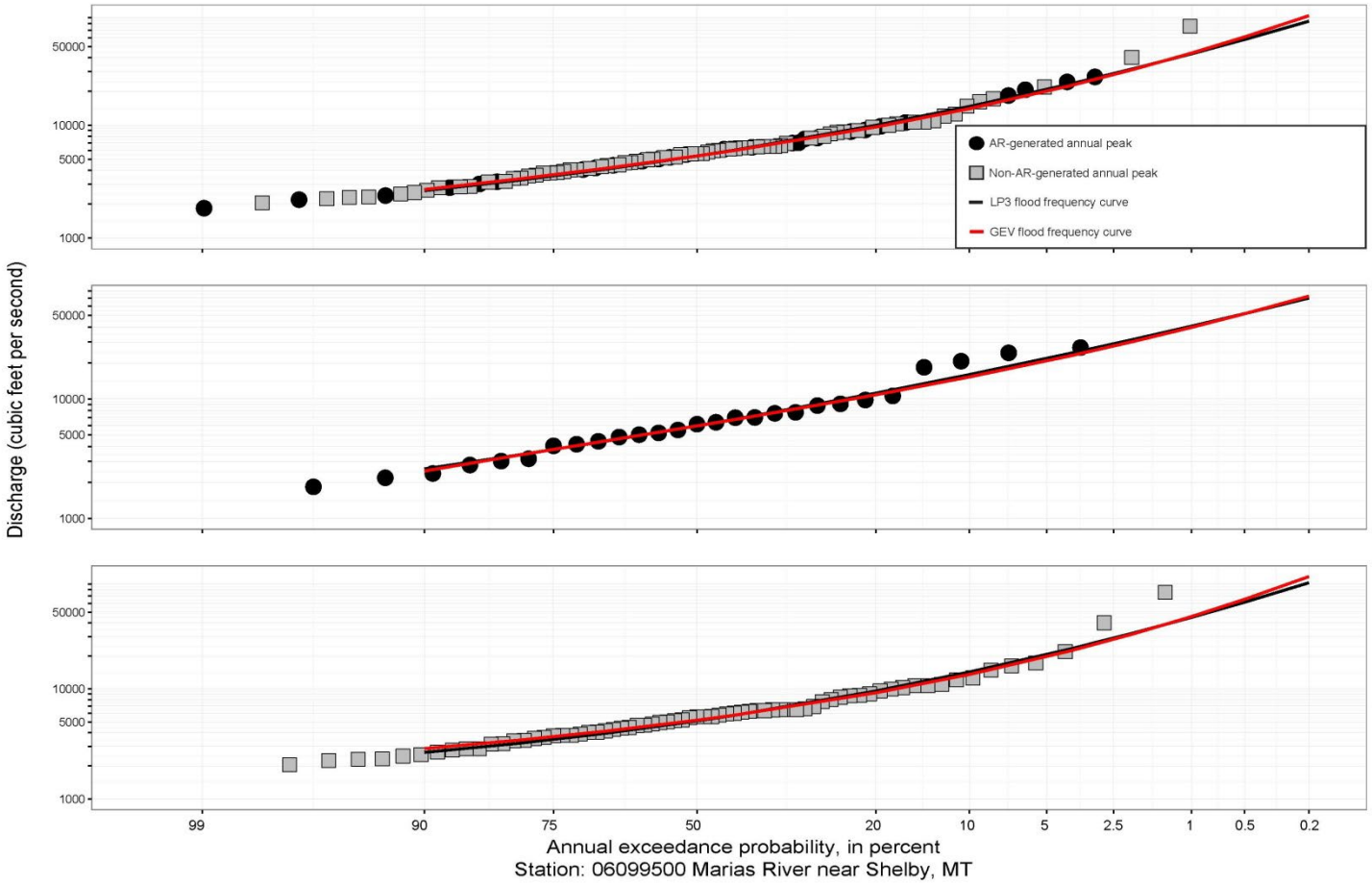

Figure S.2: Fitted frequency curves as described in figure caption S.1.

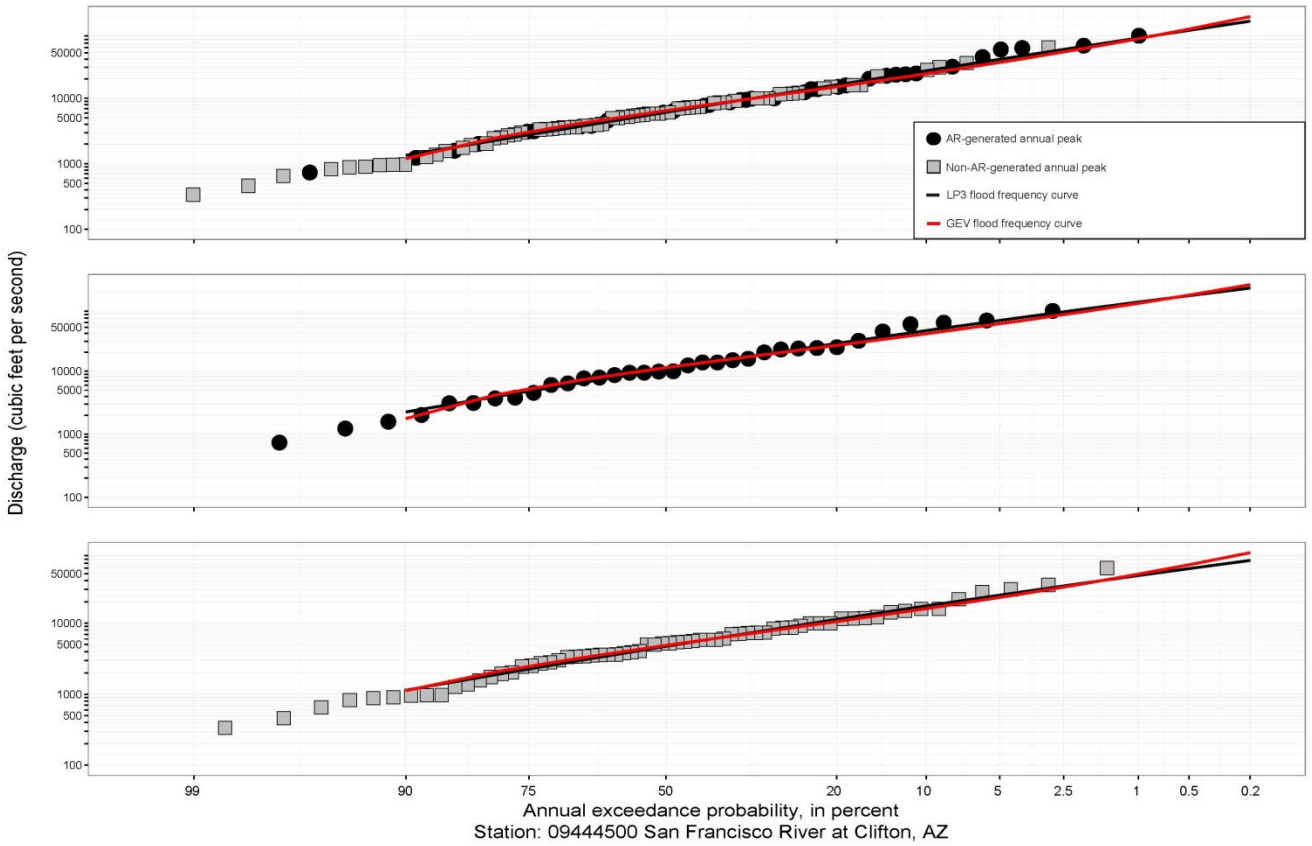

Figure S.3: Fitted frequency curves as described in figure caption S.1.

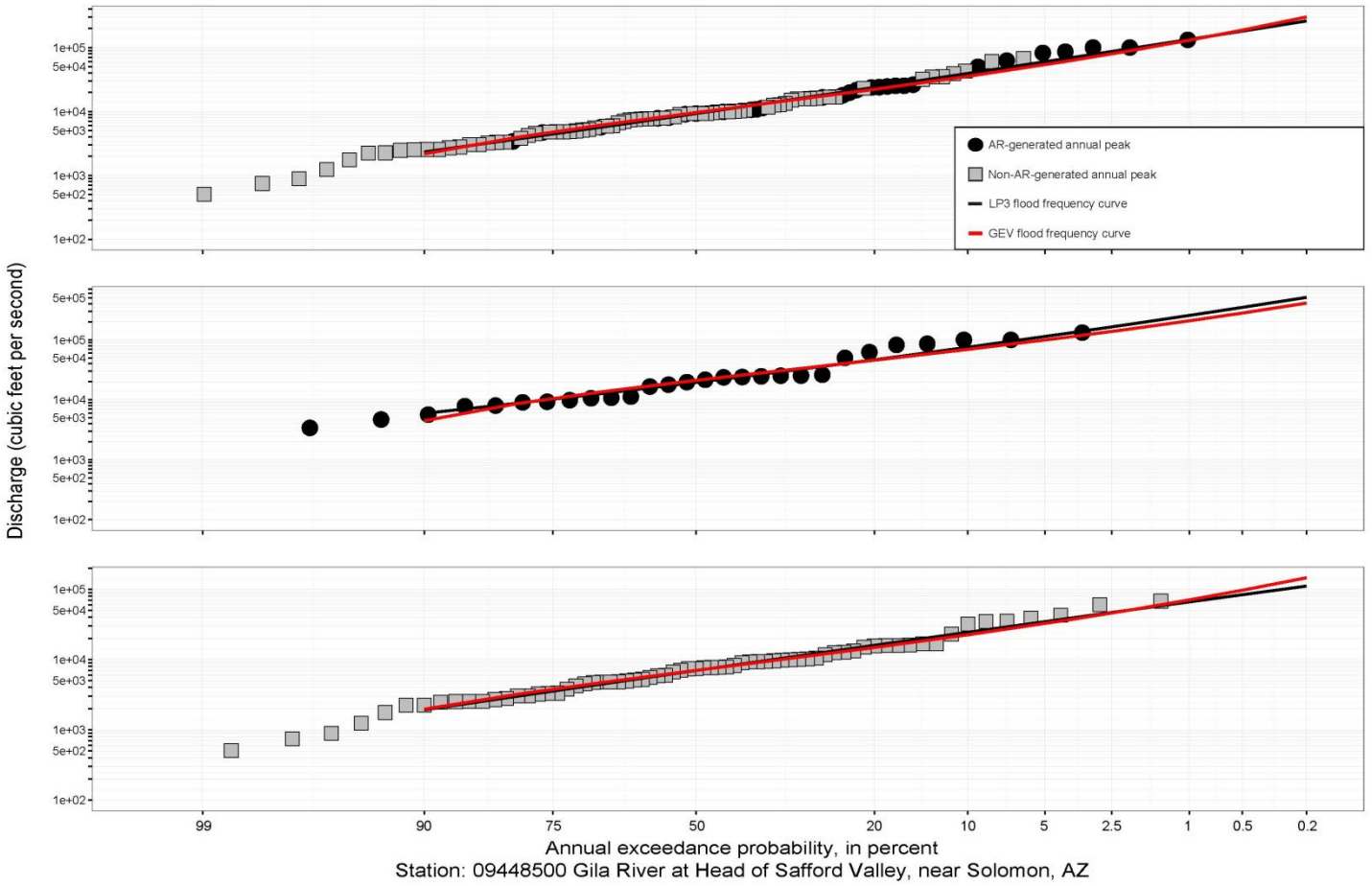

Figure S.4: Fitted frequency curves as described in figure caption S.1.

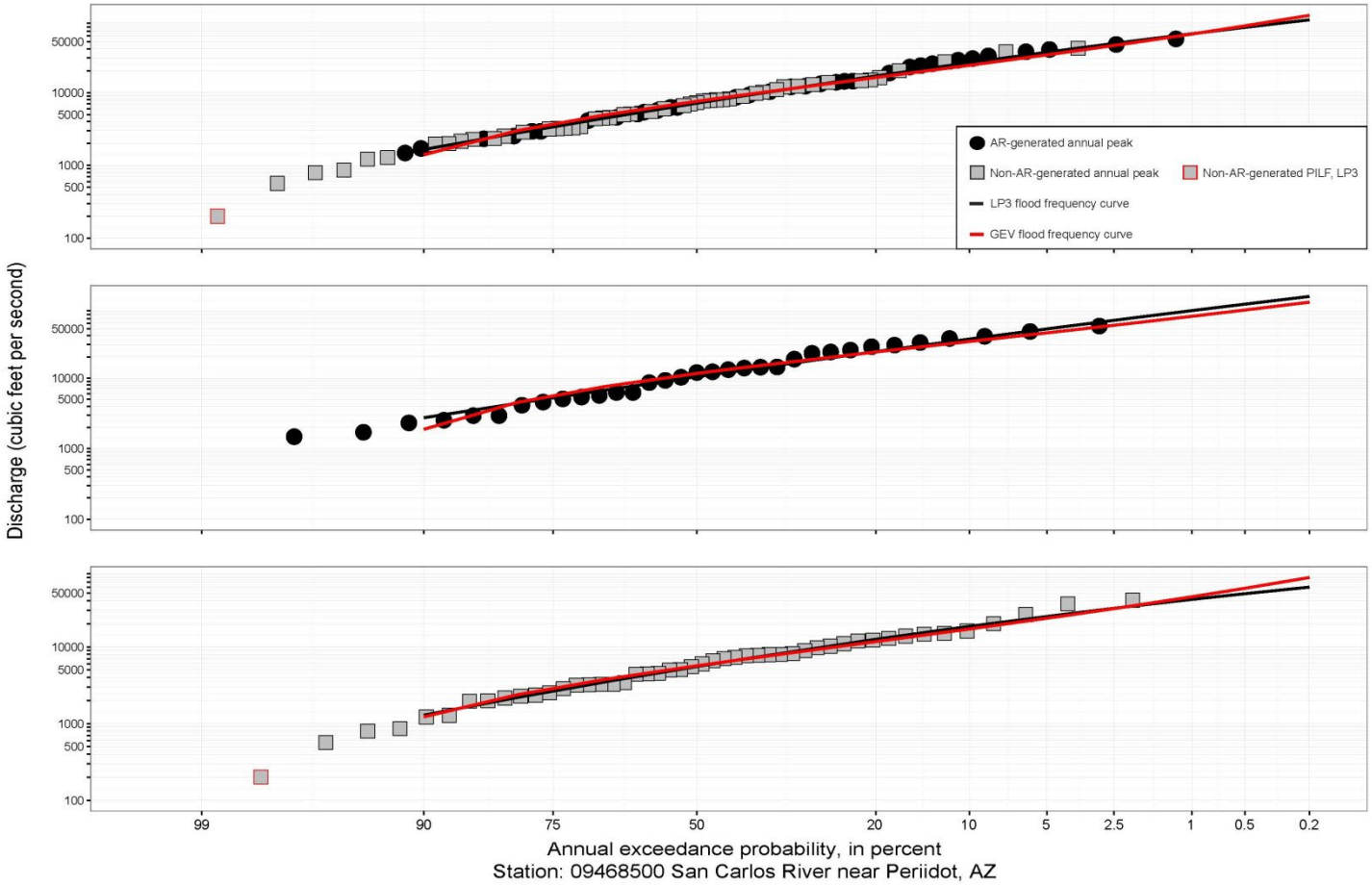

Figure S.5: Fitted frequency curves for (i) a single homogeneous population (top panel) comprised of annual peak flows generated from AR (black circles)/non-AR (grey squares) generated peaks,(ii) the subpopulation of AR-generated (middle panel), and (iii) the subpopulation of non-AR-generated peaks (bottom panel) for the log-Pearson Type III (LP3) distribution (black line) and the generalized extreme value (GEV) distribution (red line). Potentially influential low floods (PILFs) for the LP3 approach are highlighted with red.

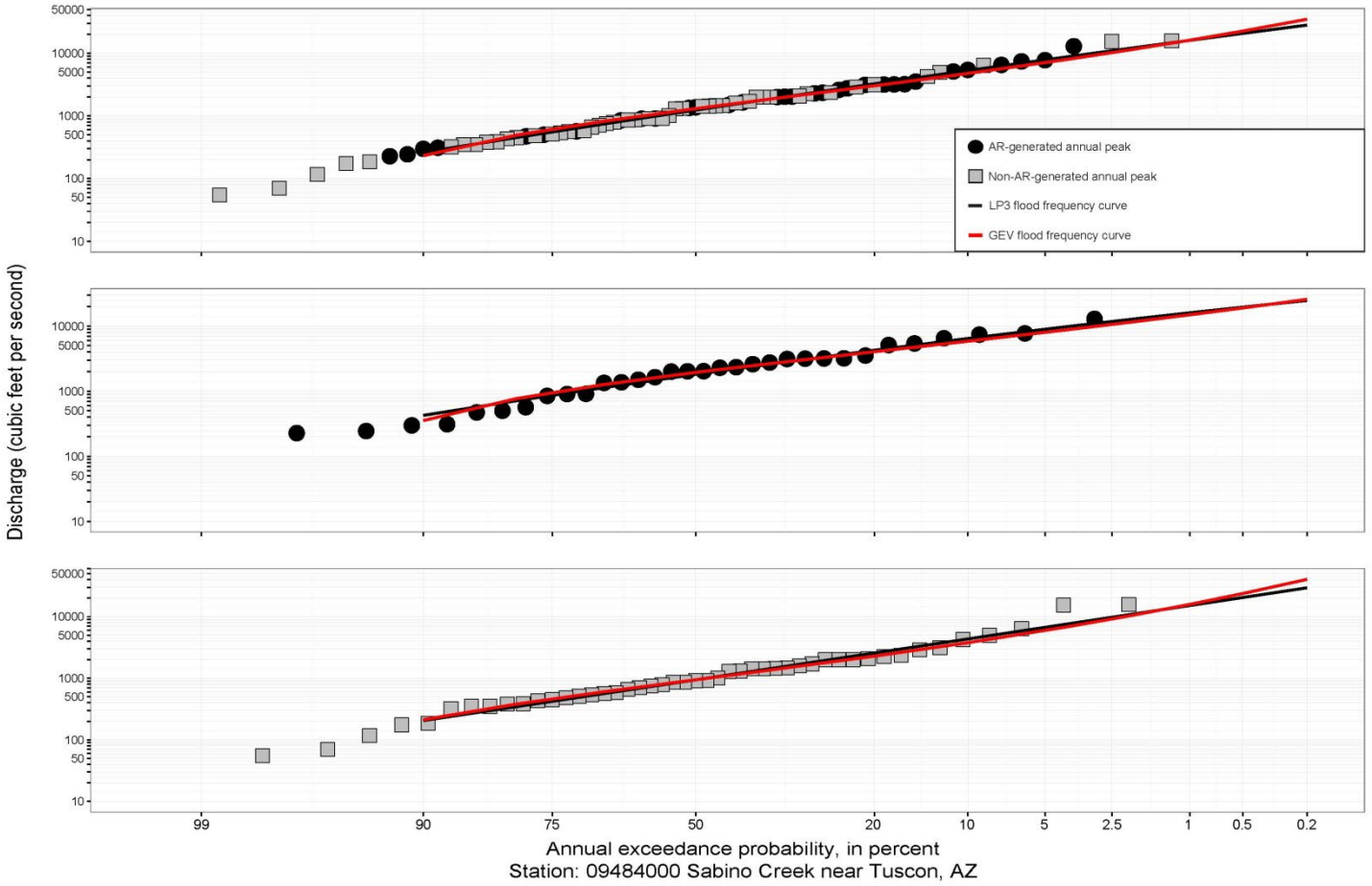

Figure S.6: Fitted frequency curves as described in figure caption S.1.

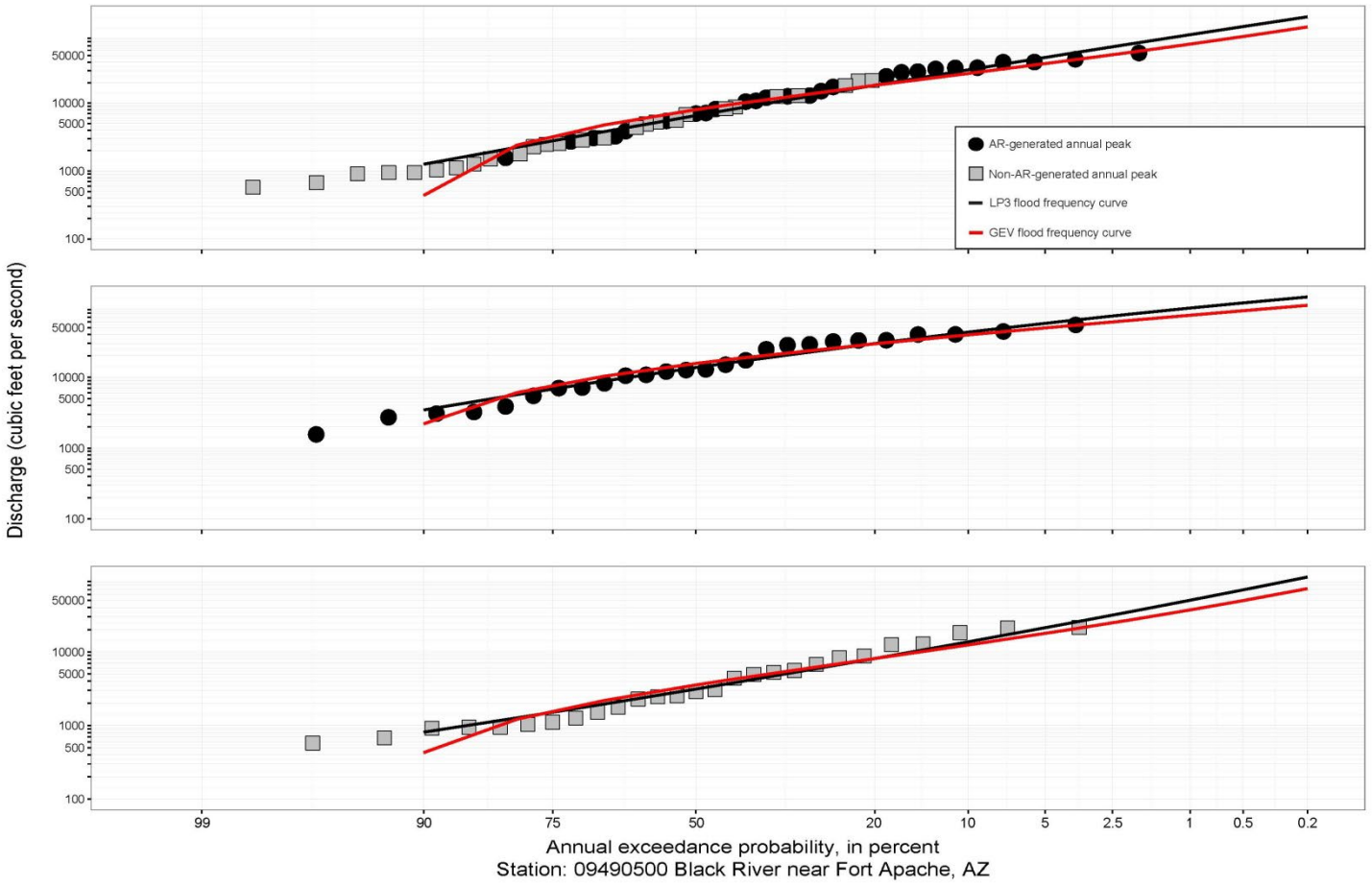

Figure S.7: Fitted frequency curves as described in figure caption S.1.

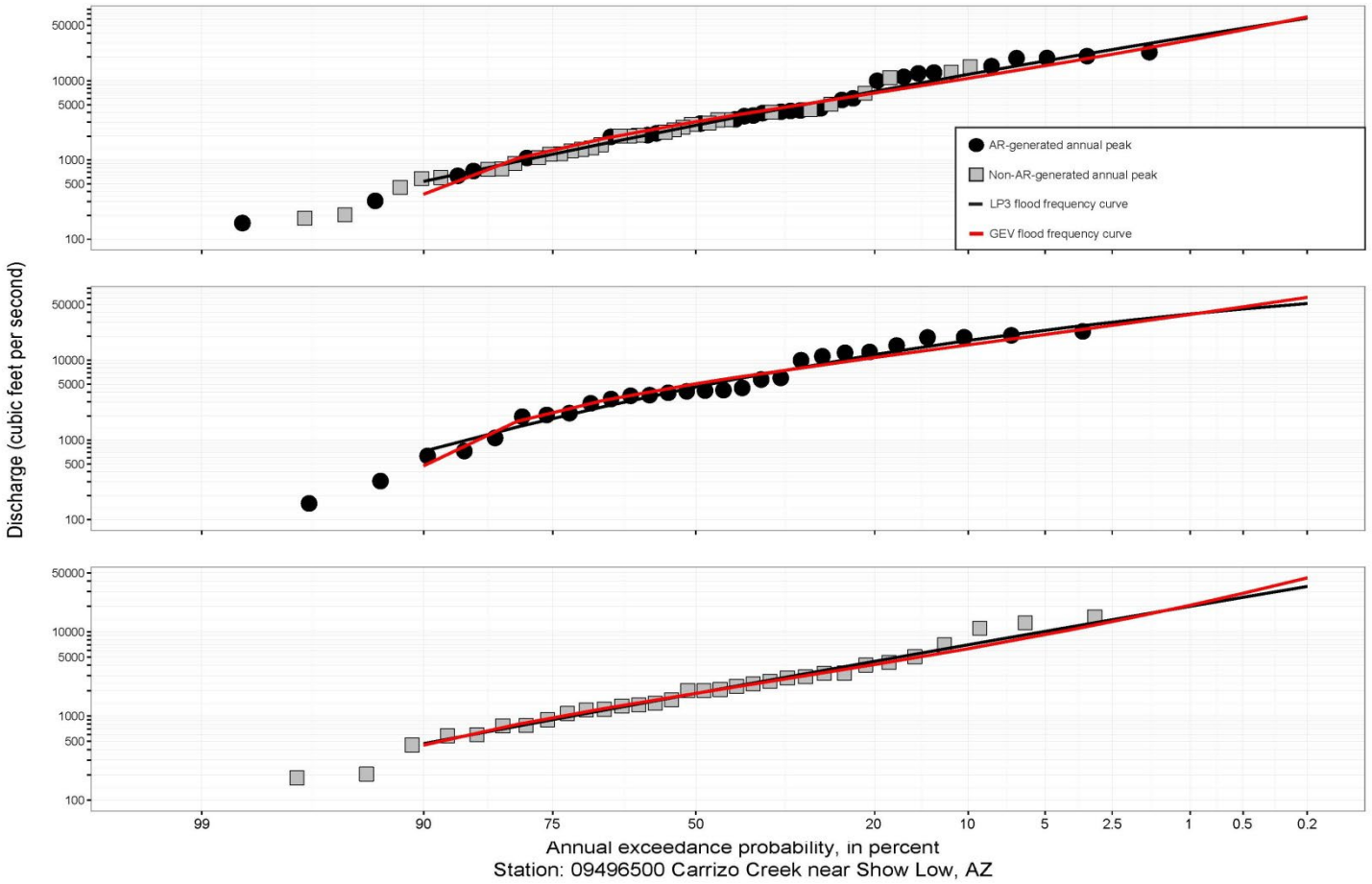

Figure S.8: Fitted frequency curves as described in figure caption S.1.

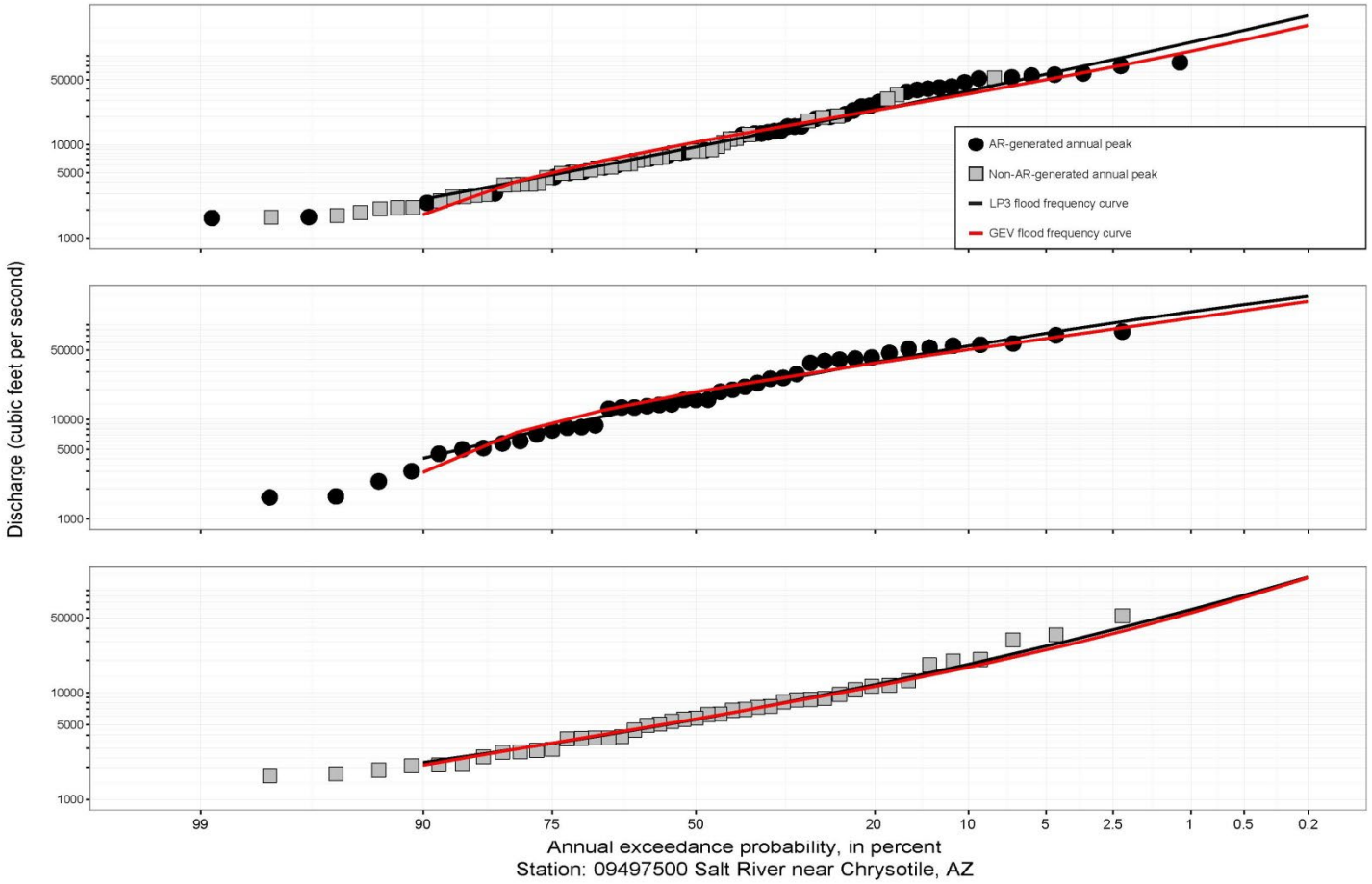

Figure S.9: Fitted frequency curves as described in figure caption S.1.

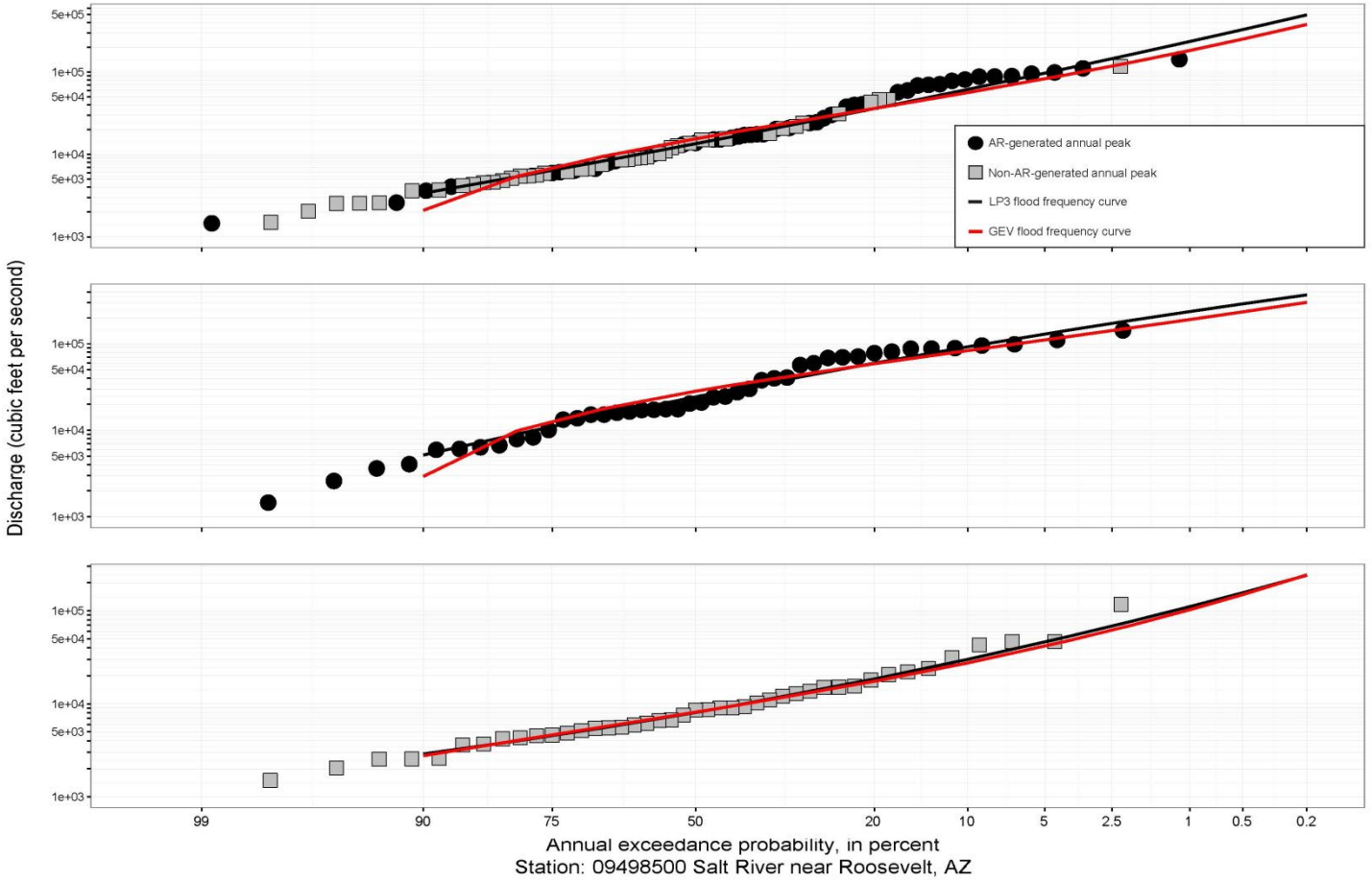

Figure S.10: Fitted frequency curves as described in figure caption S.1.

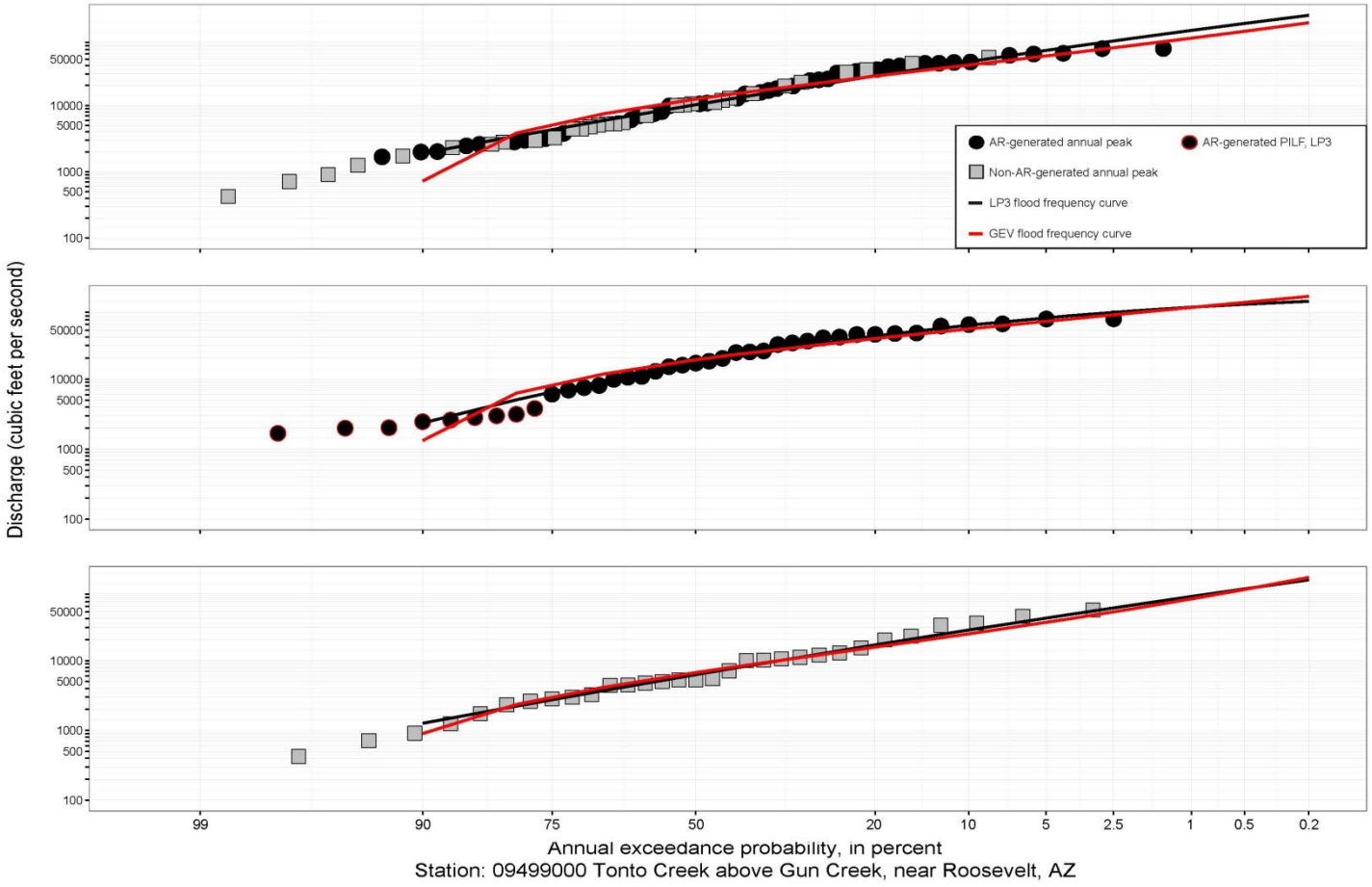

Figure S.11: Fitted frequency curves as described in figure caption S.5.

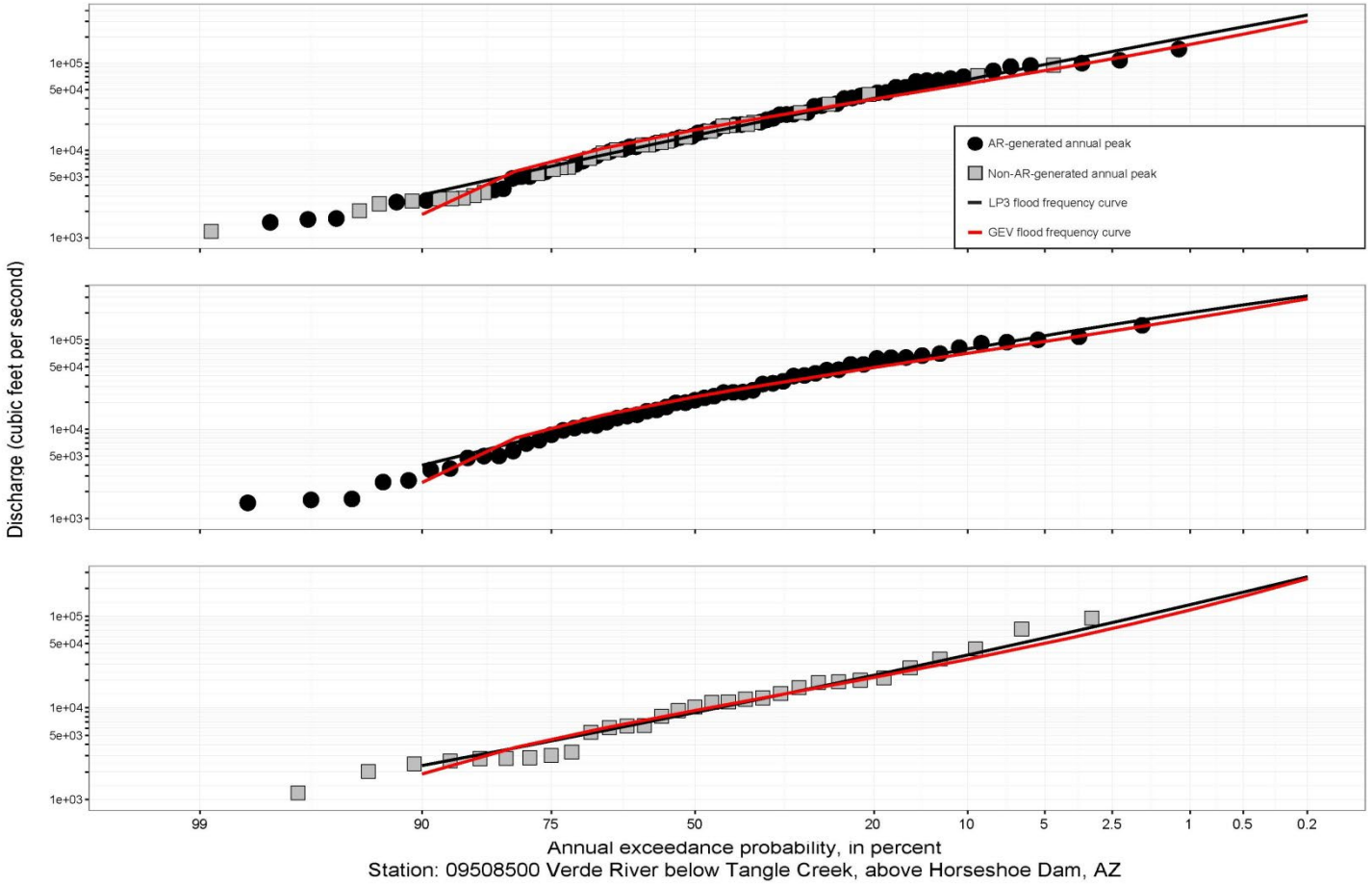

Figure S.12: Fitted frequency curves as described in figure caption S.1.

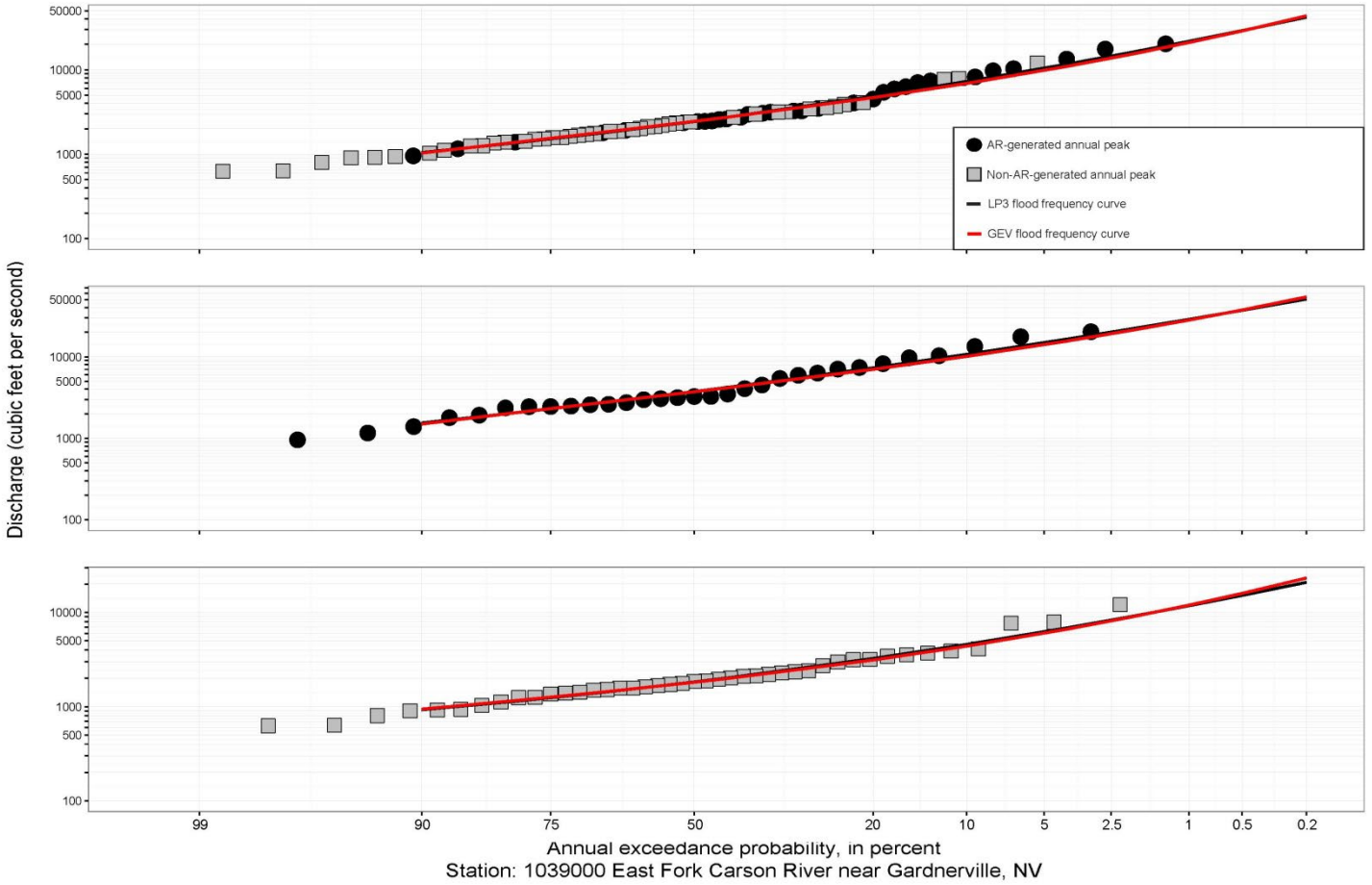

Figure S.13: Fitted frequency curves as described in figure caption S.1.

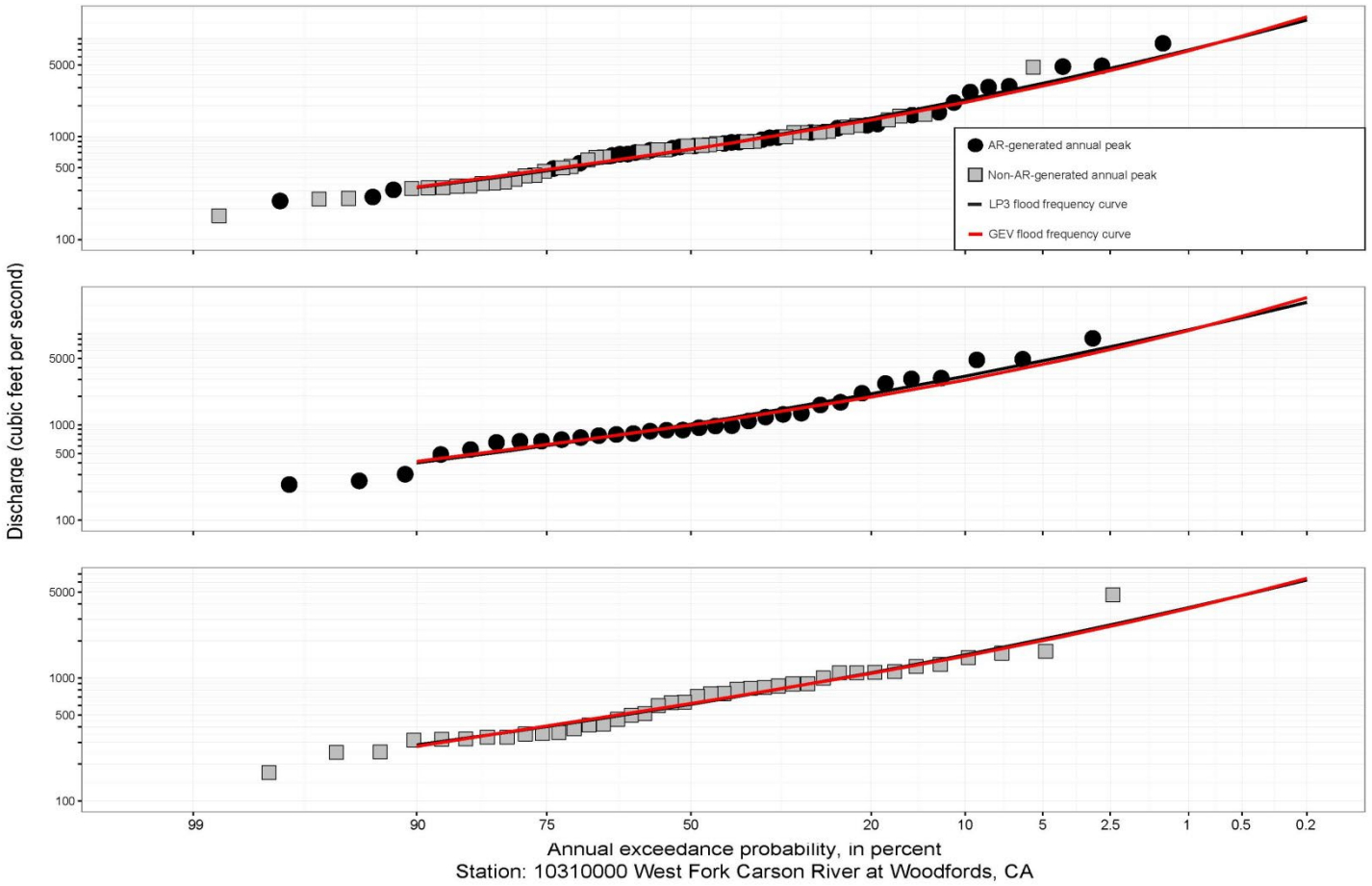

Figure S.14: Fitted frequency curves as described in figure caption S.1.

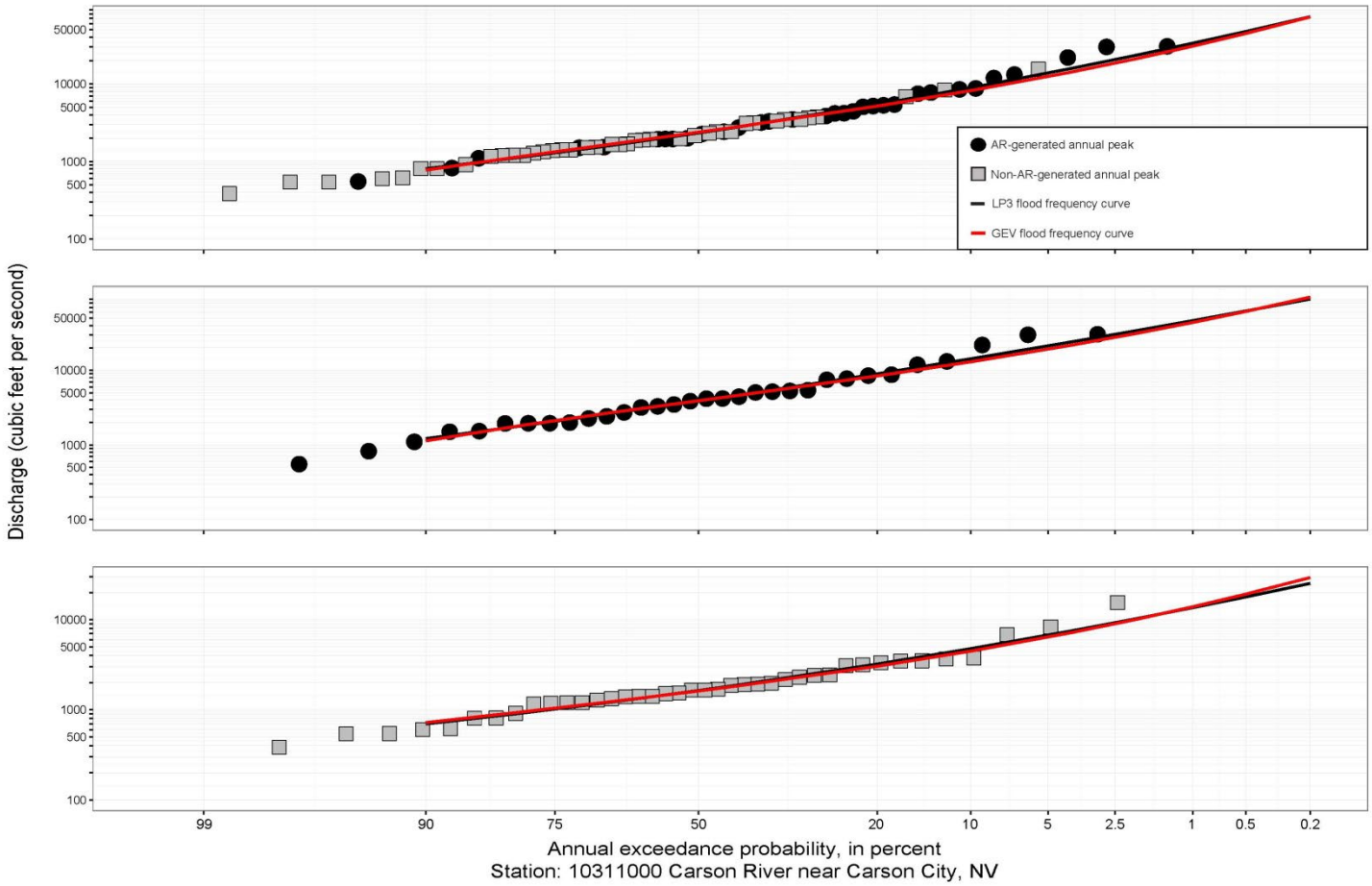

Figure S.15: Fitted frequency curves as described in figure caption S.1.

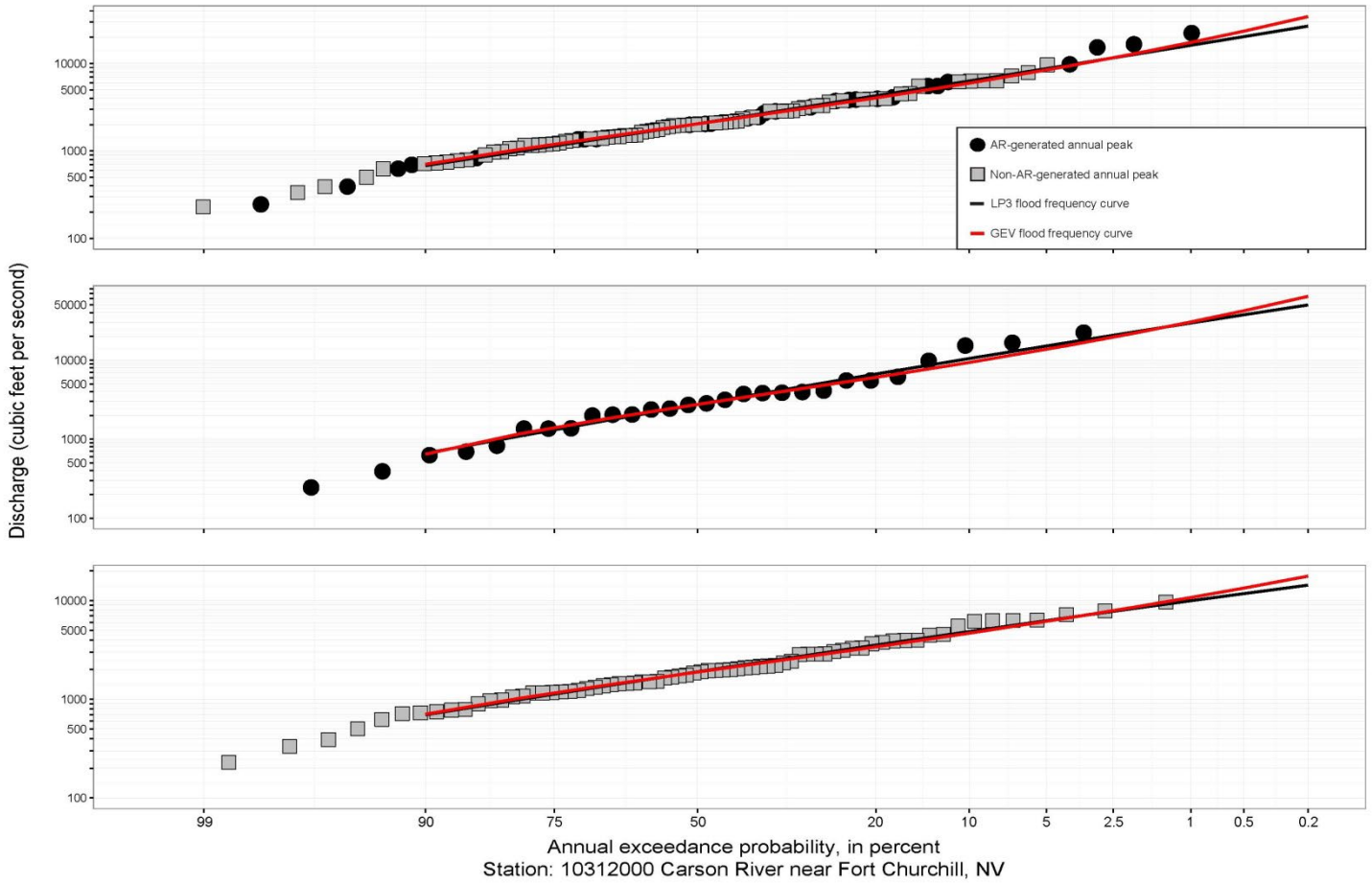

Figure S.16: Fitted frequency curves as described in figure caption S.1.

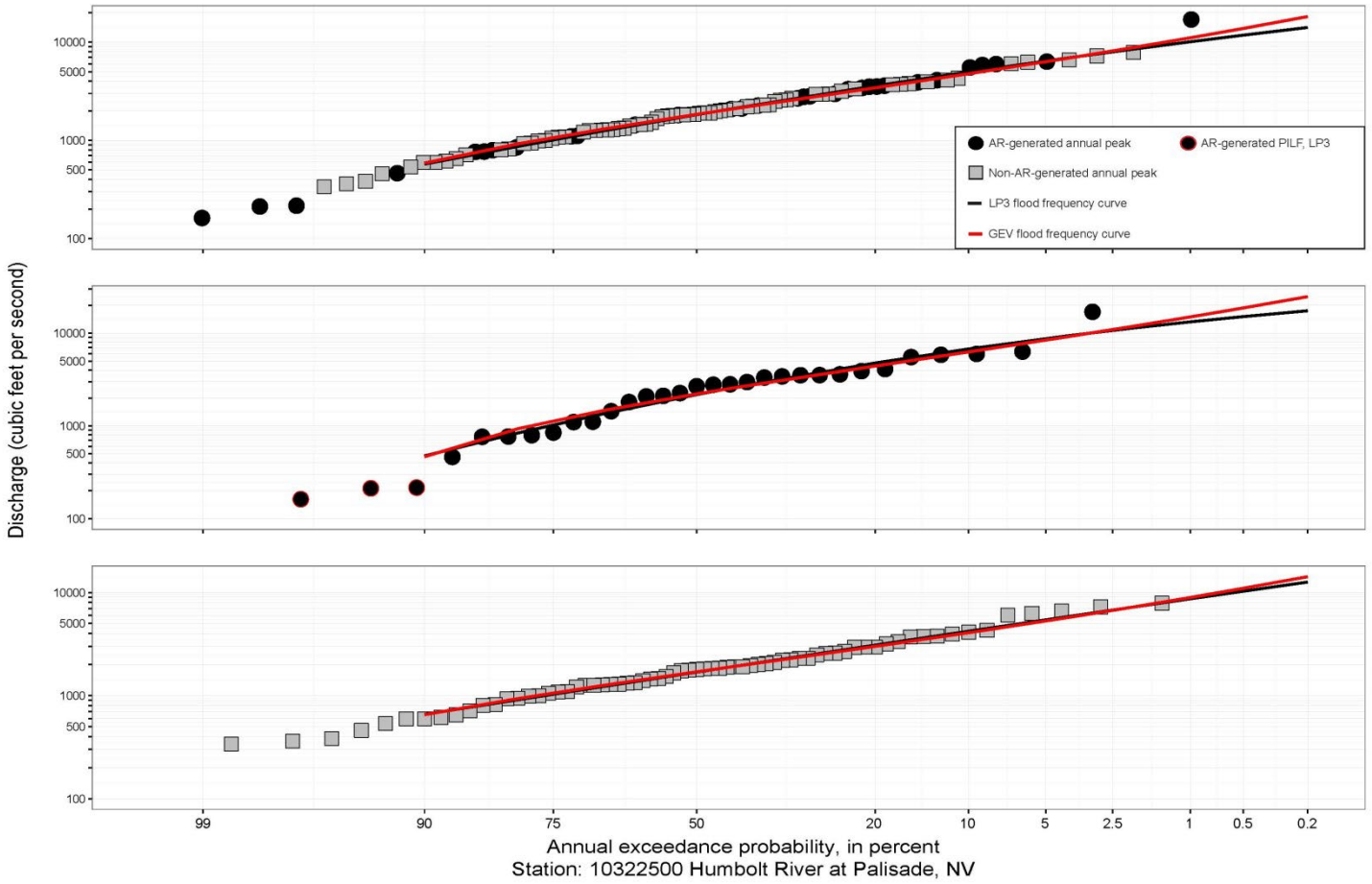

Figure S.17: Fitted frequency curves as described in figure caption S.5.

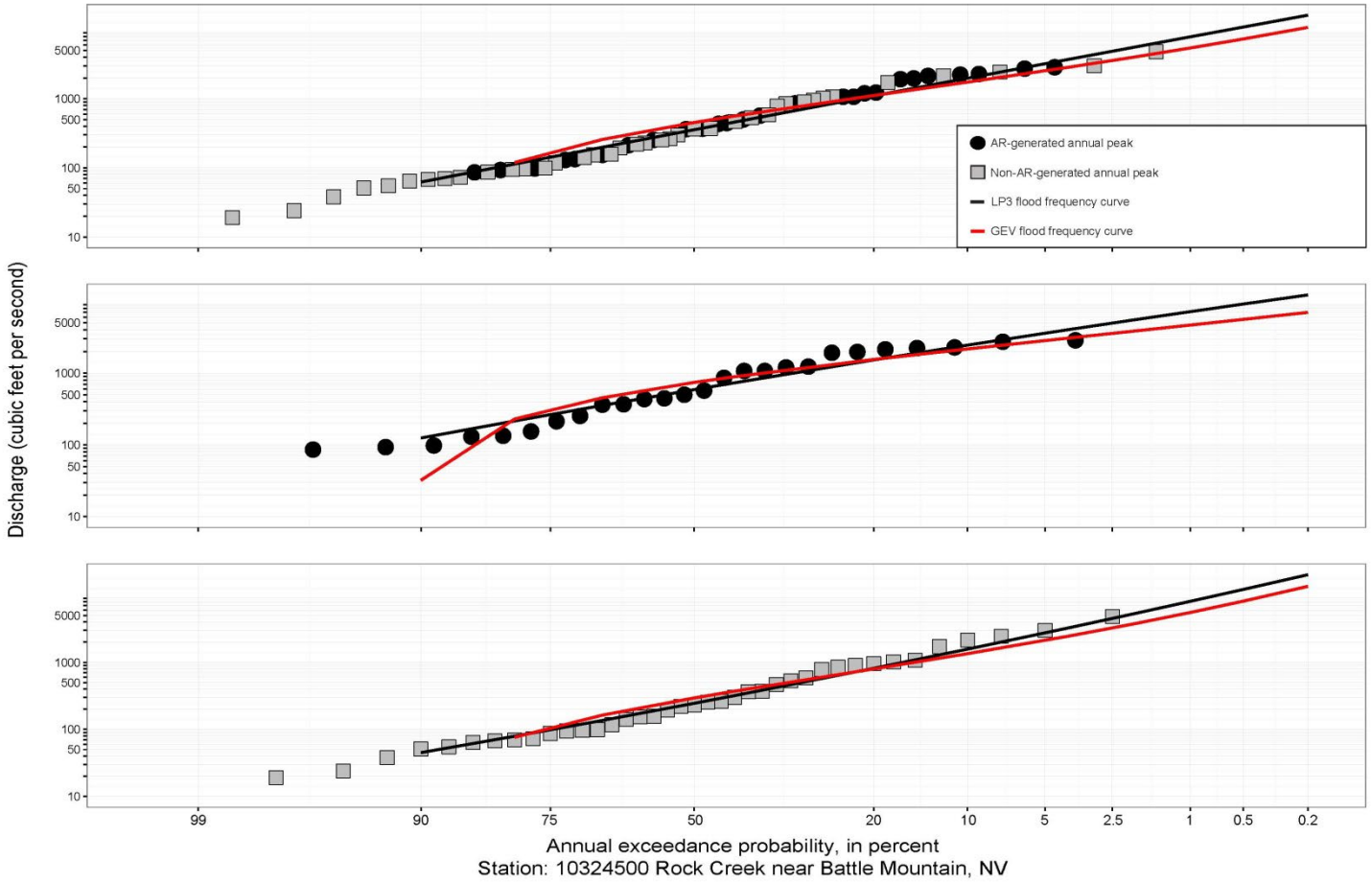

Figure S.18: Fitted frequency curves as described in figure caption S.1.

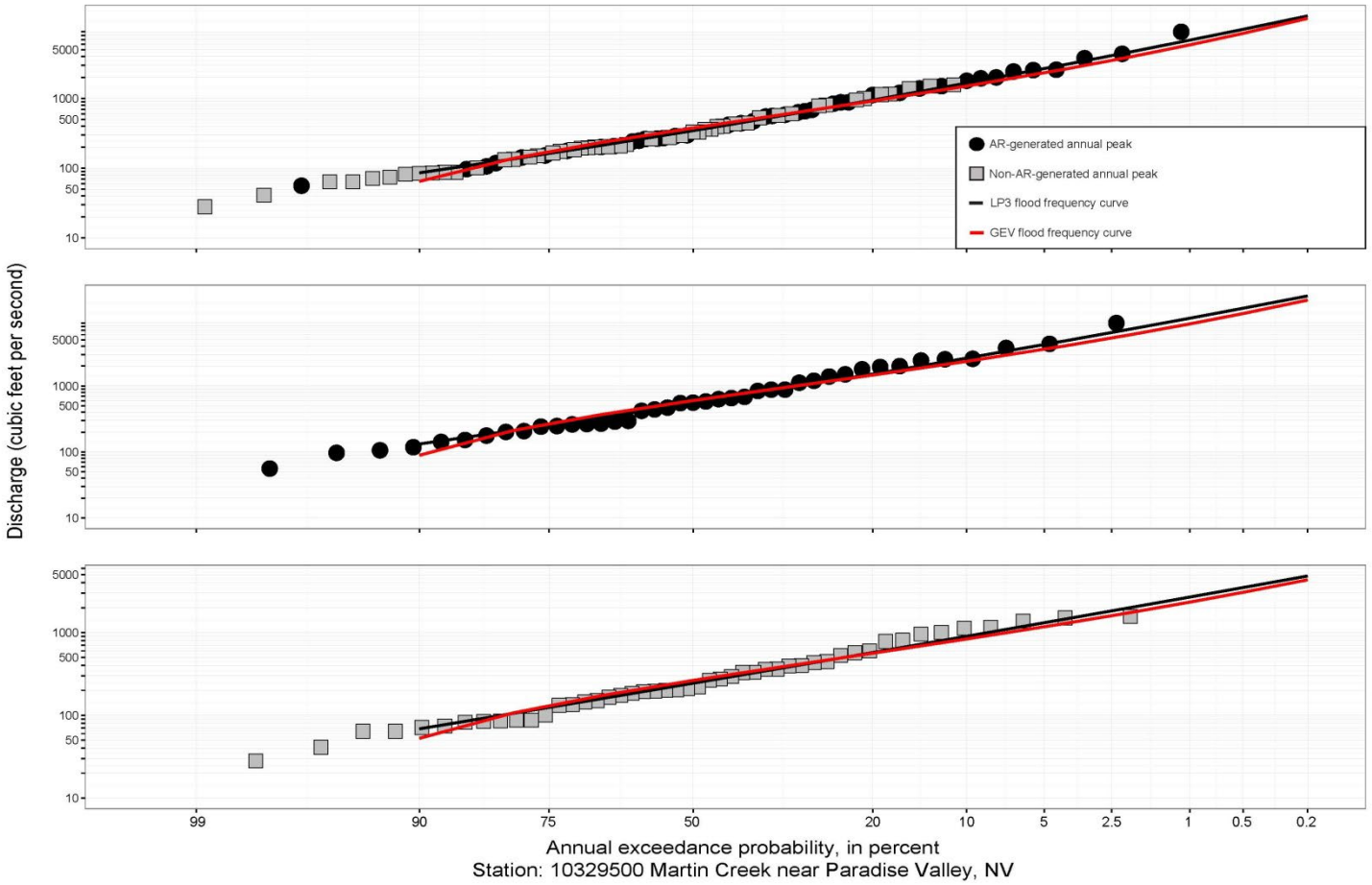

Figure S.19: Fitted frequency curves as described in figure caption S.1.

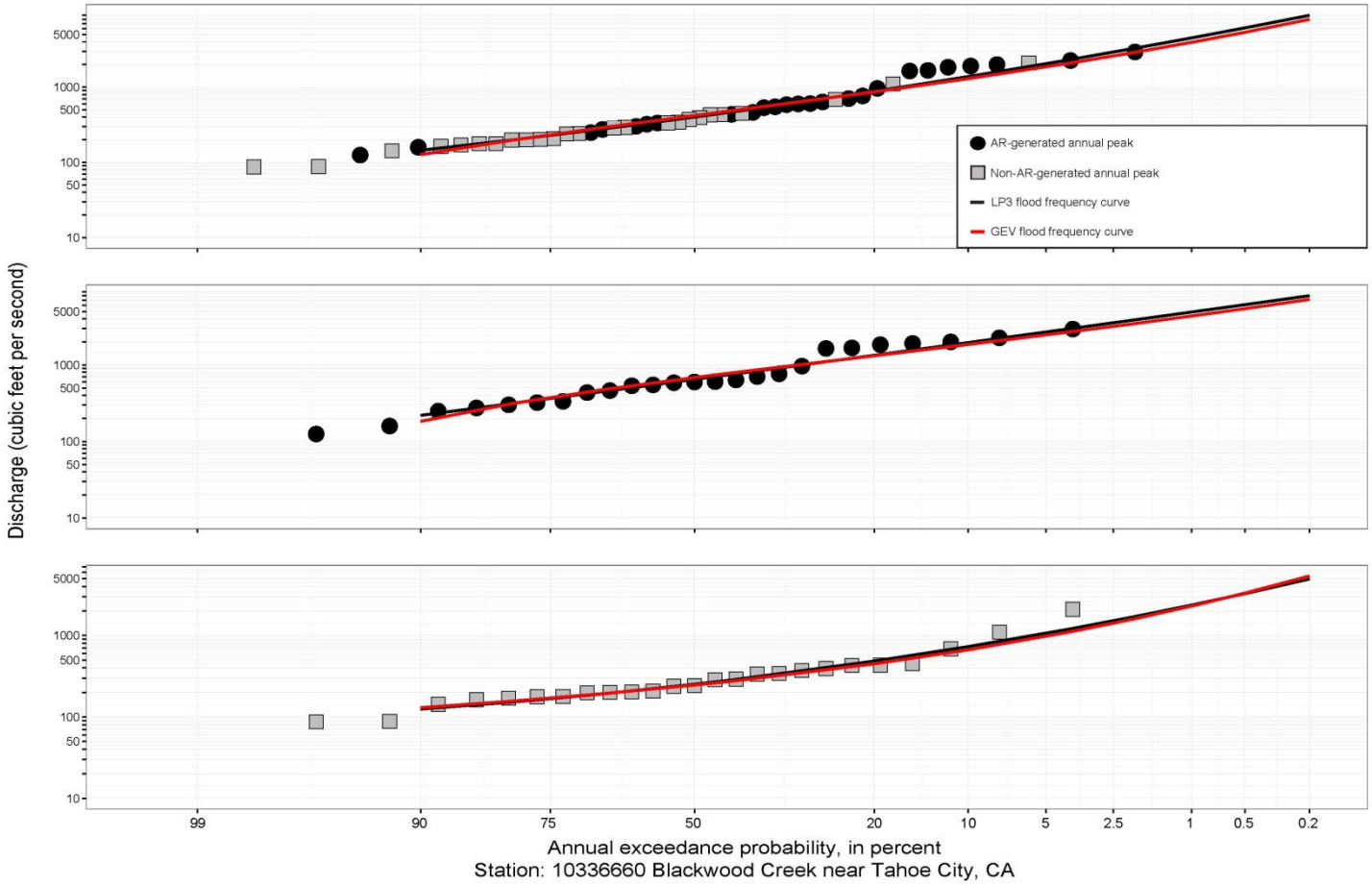

Figure S.20: Fitted frequency curves as described in figure caption S.1.

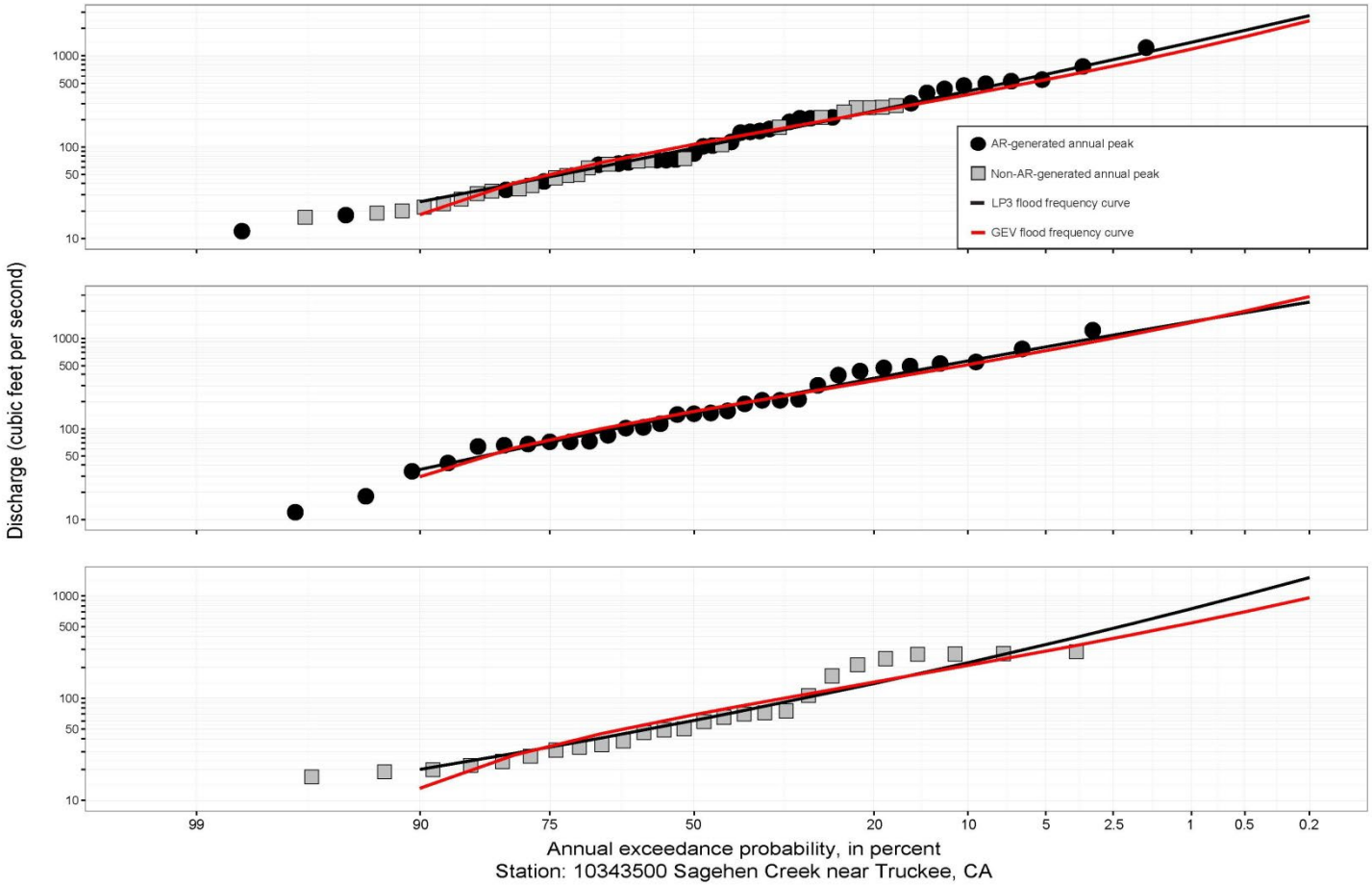

Figure S.21: Fitted frequency curves as described in figure caption S.1.

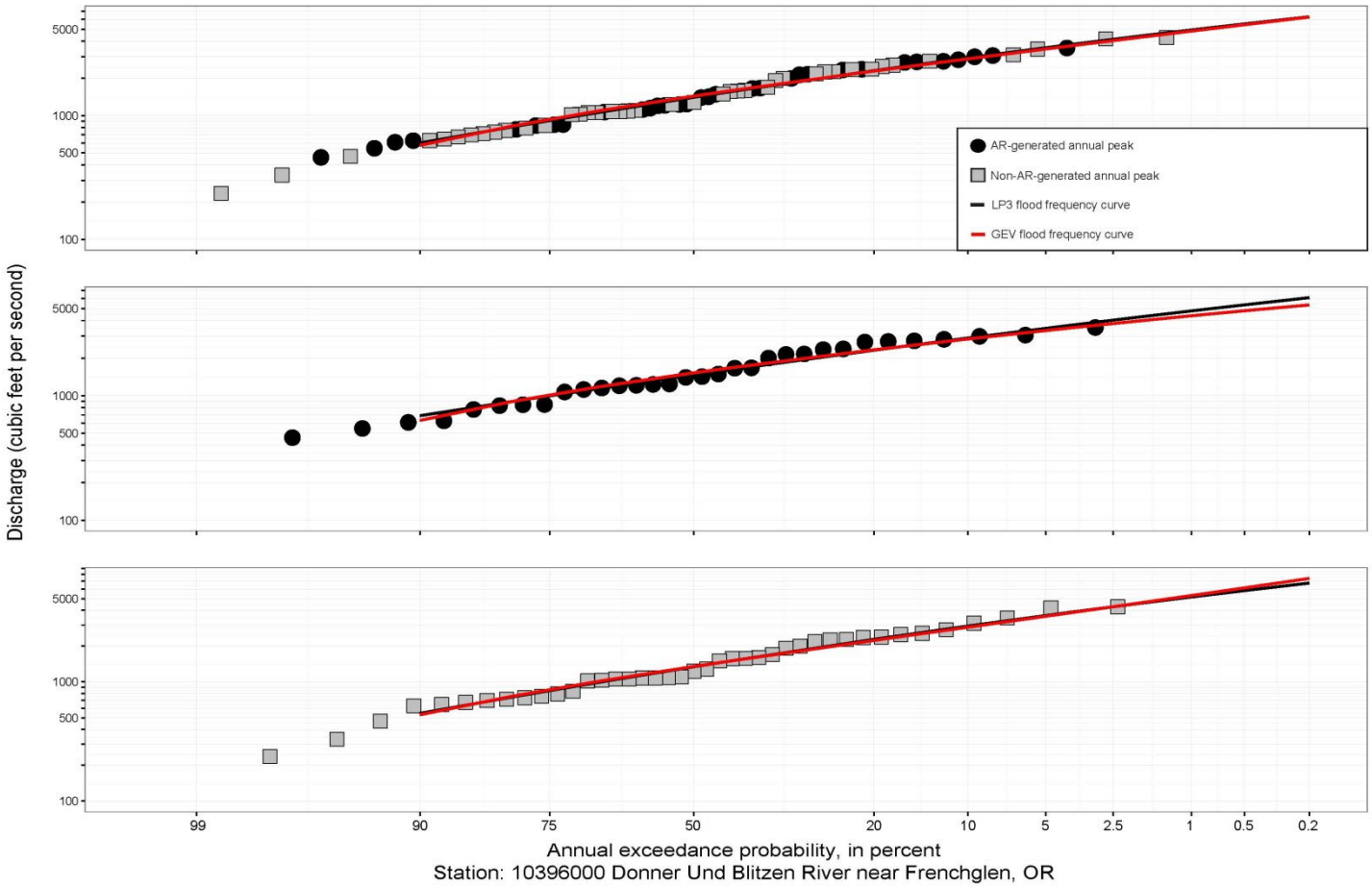

Figure S.22: Fitted frequency curves as described in figure caption S.1.

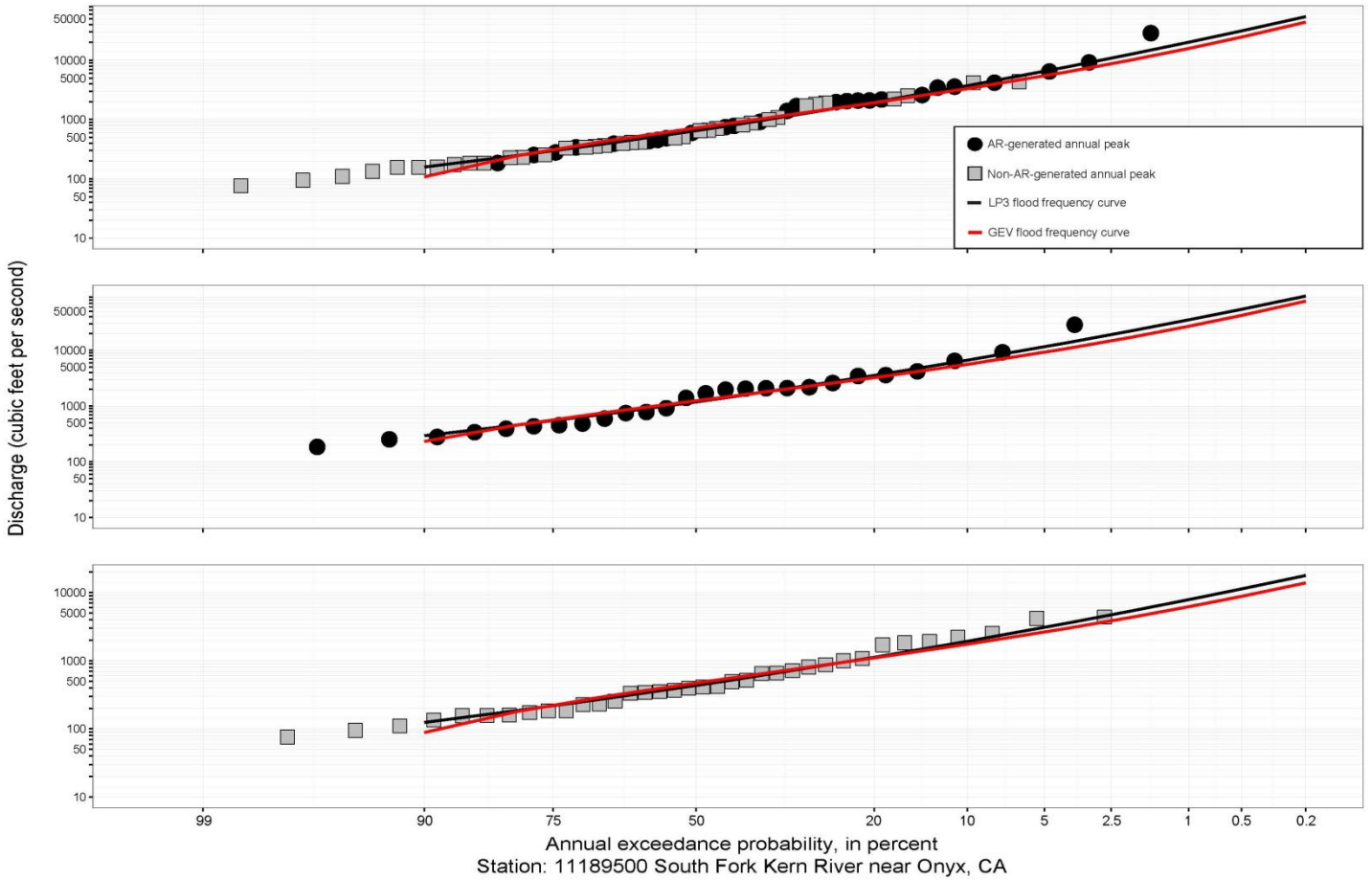

Figure S.23: Fitted frequency curves as described in figure caption S.1.

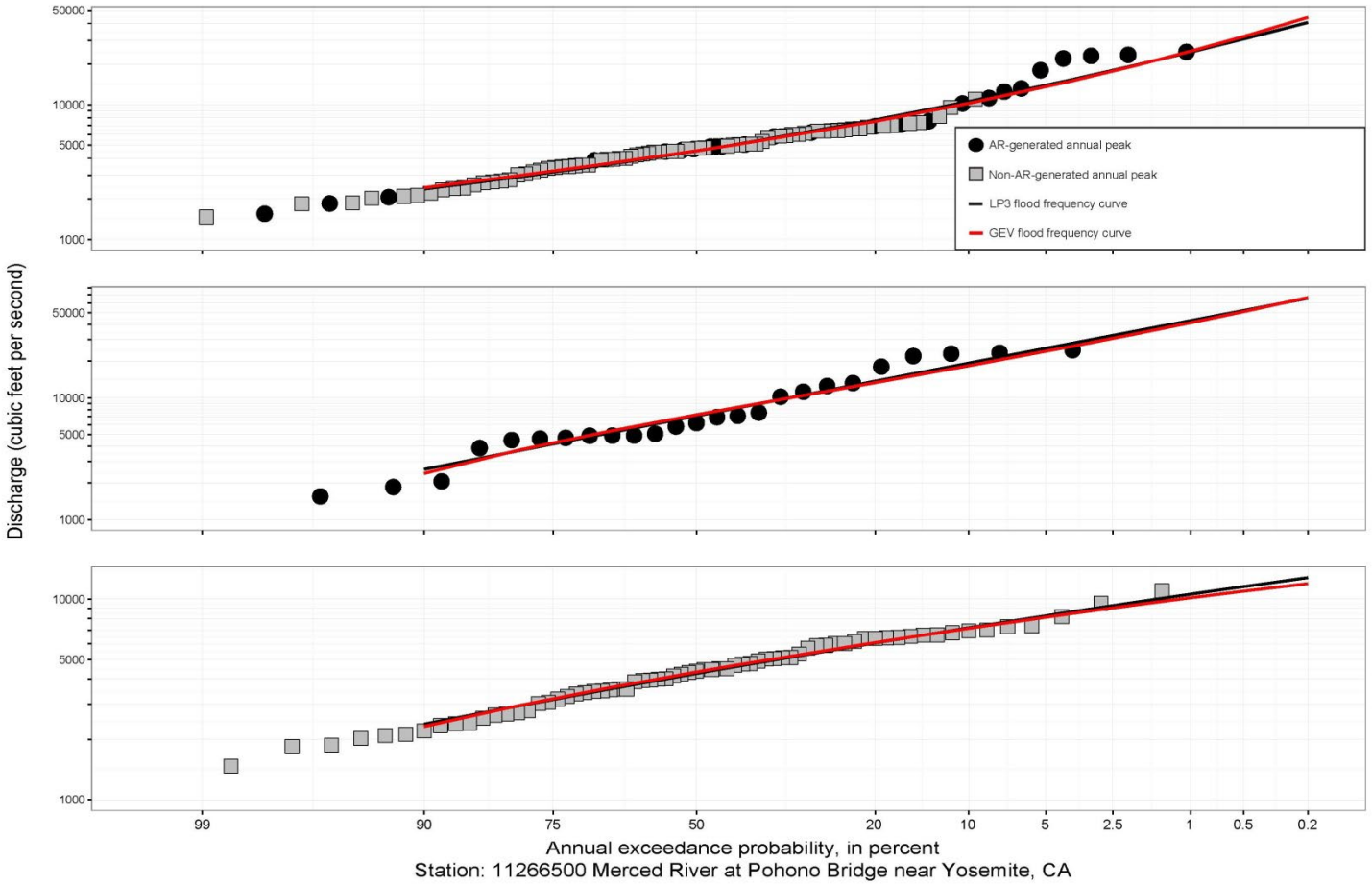

Figure S.24: Fitted frequency curves as described in figure caption S.1.

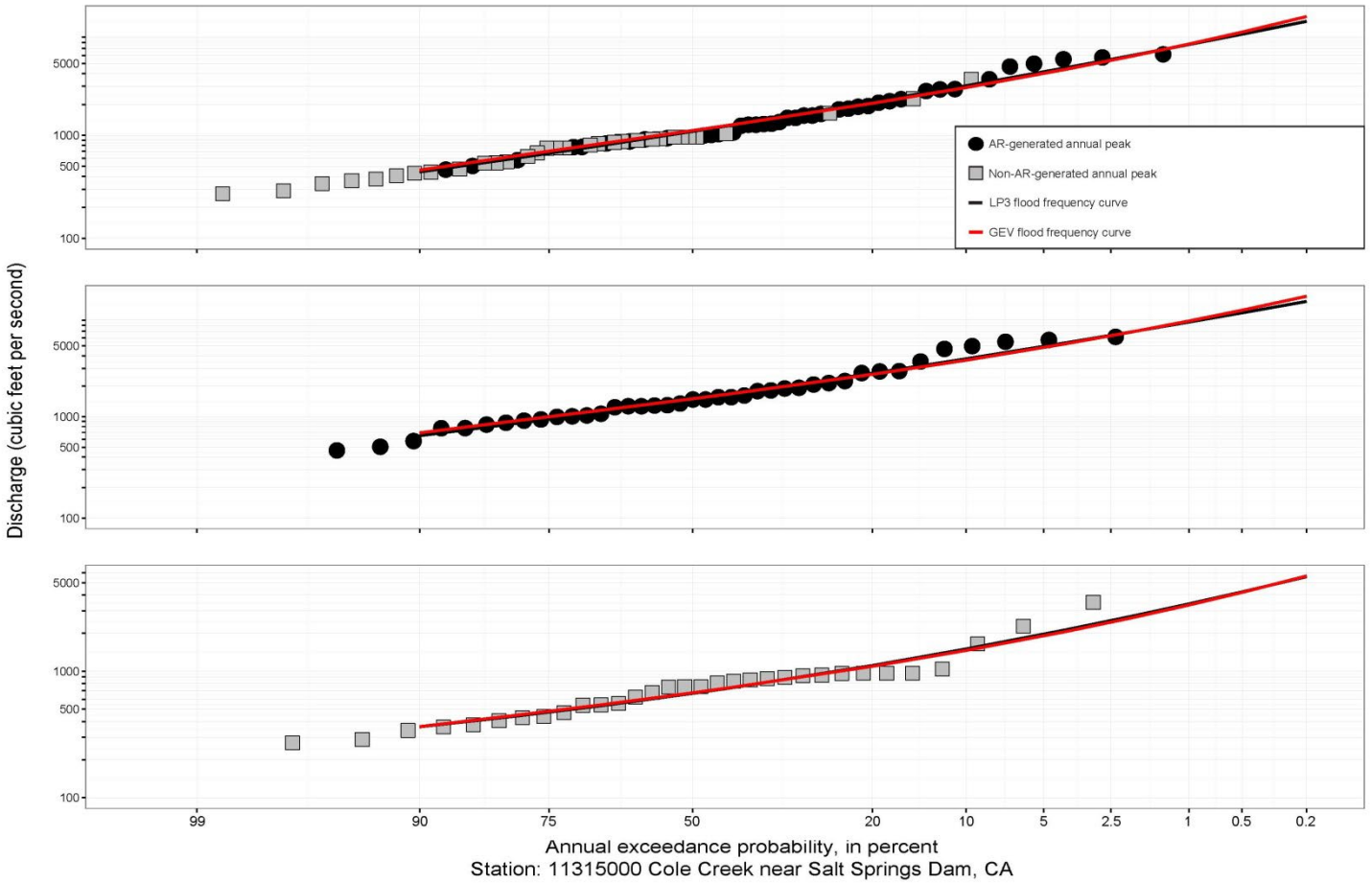

Figure S.25: Fitted frequency curves as described in figure caption S.1.

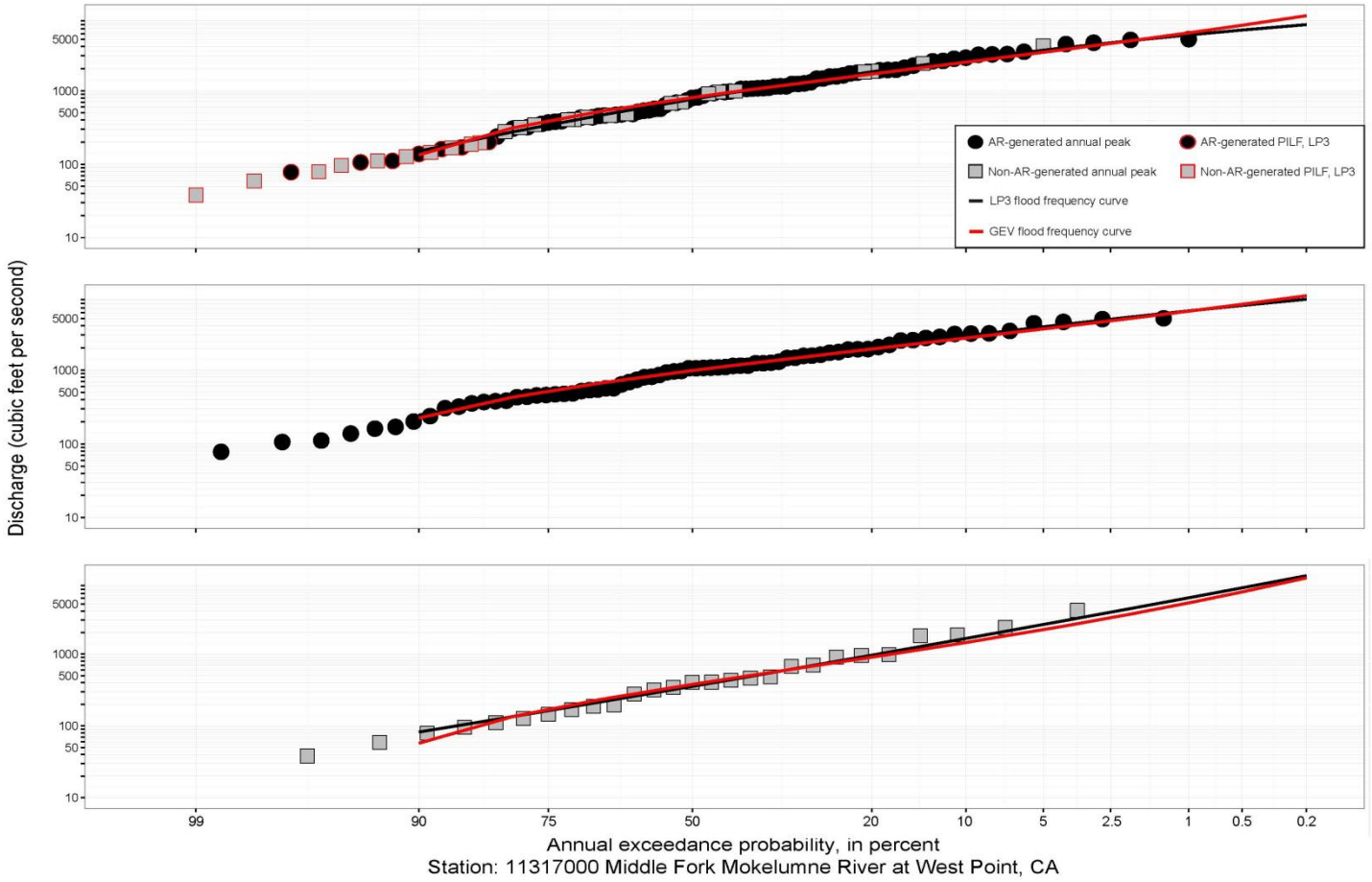

Figure S.26: Fitted frequency curves as described in figure caption S.5.

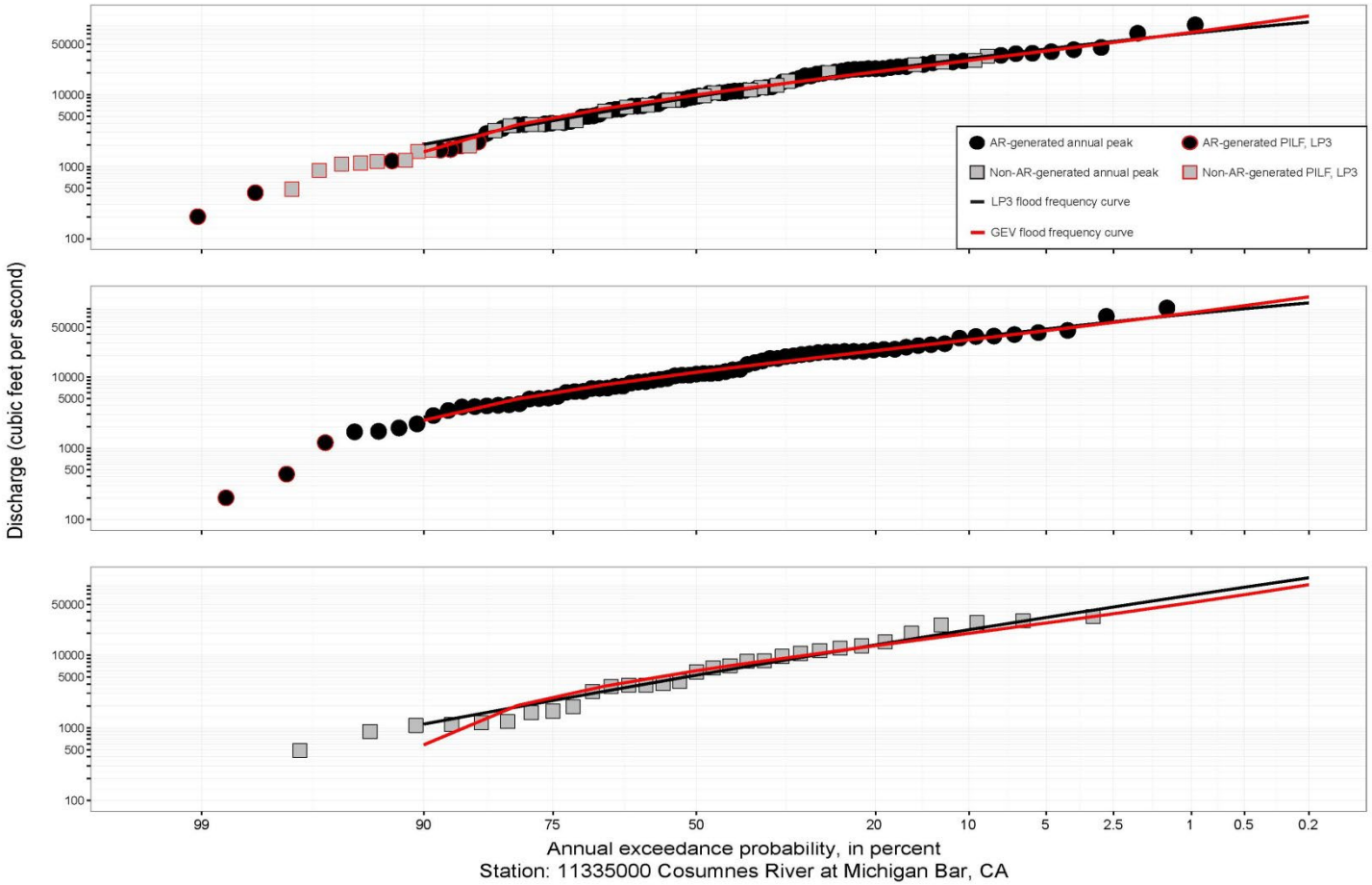

Figure S.27: Fitted frequency curves as described in figure caption S.5.

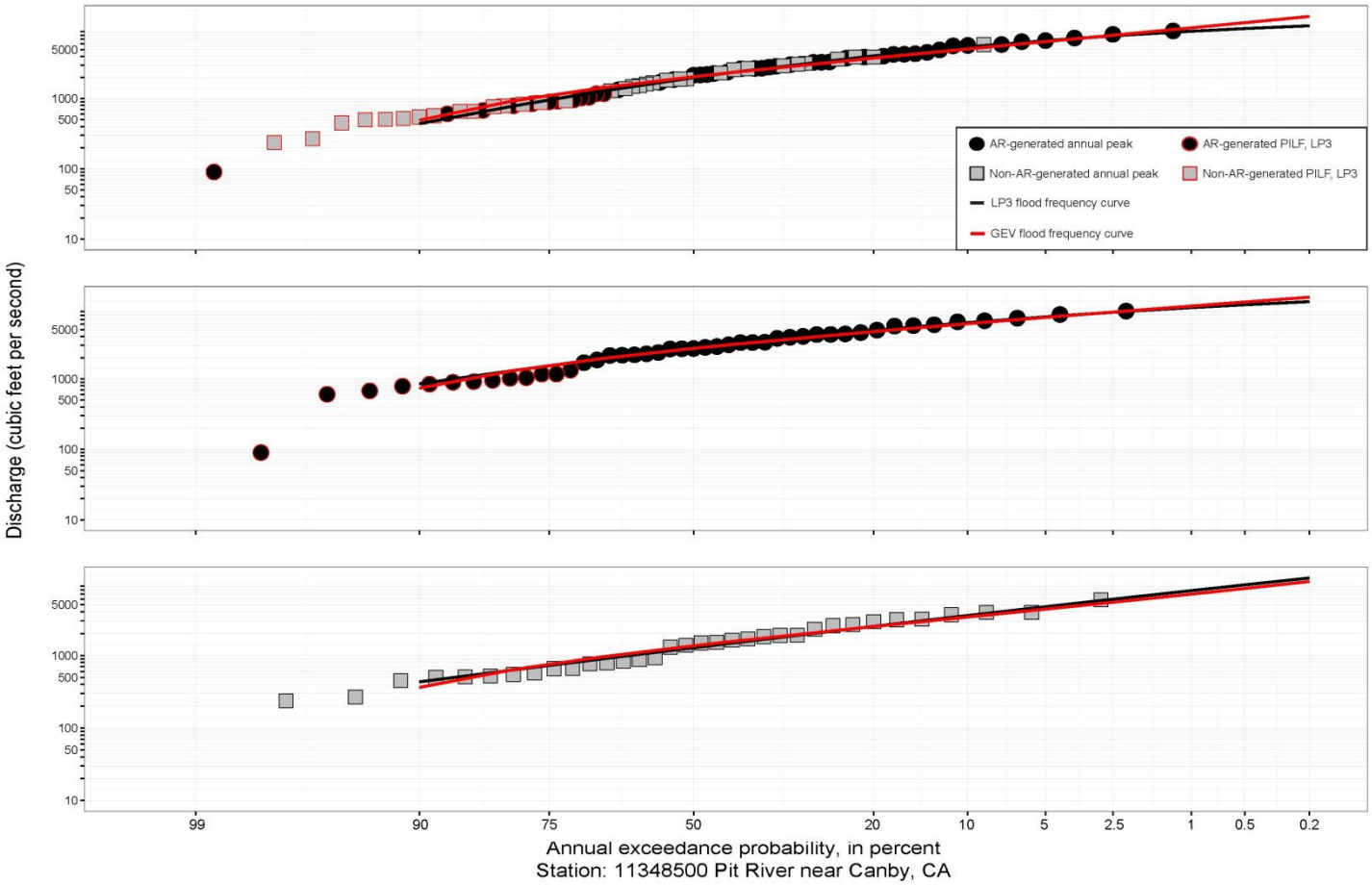

Figure S.28: Fitted frequency curves as described in figure caption S.5.

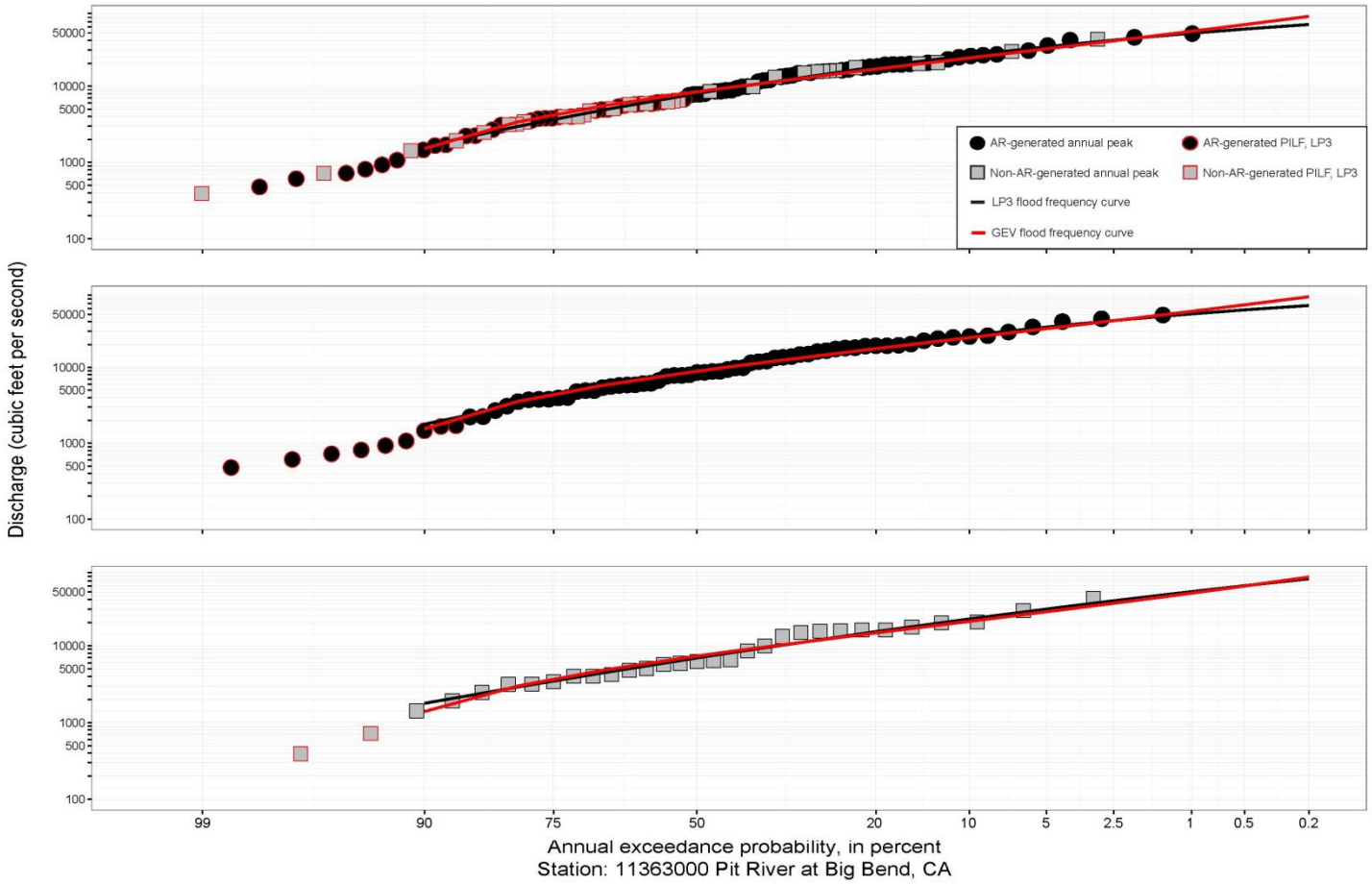

Figure S.29: Fitted frequency curves as described in figure caption S.5.

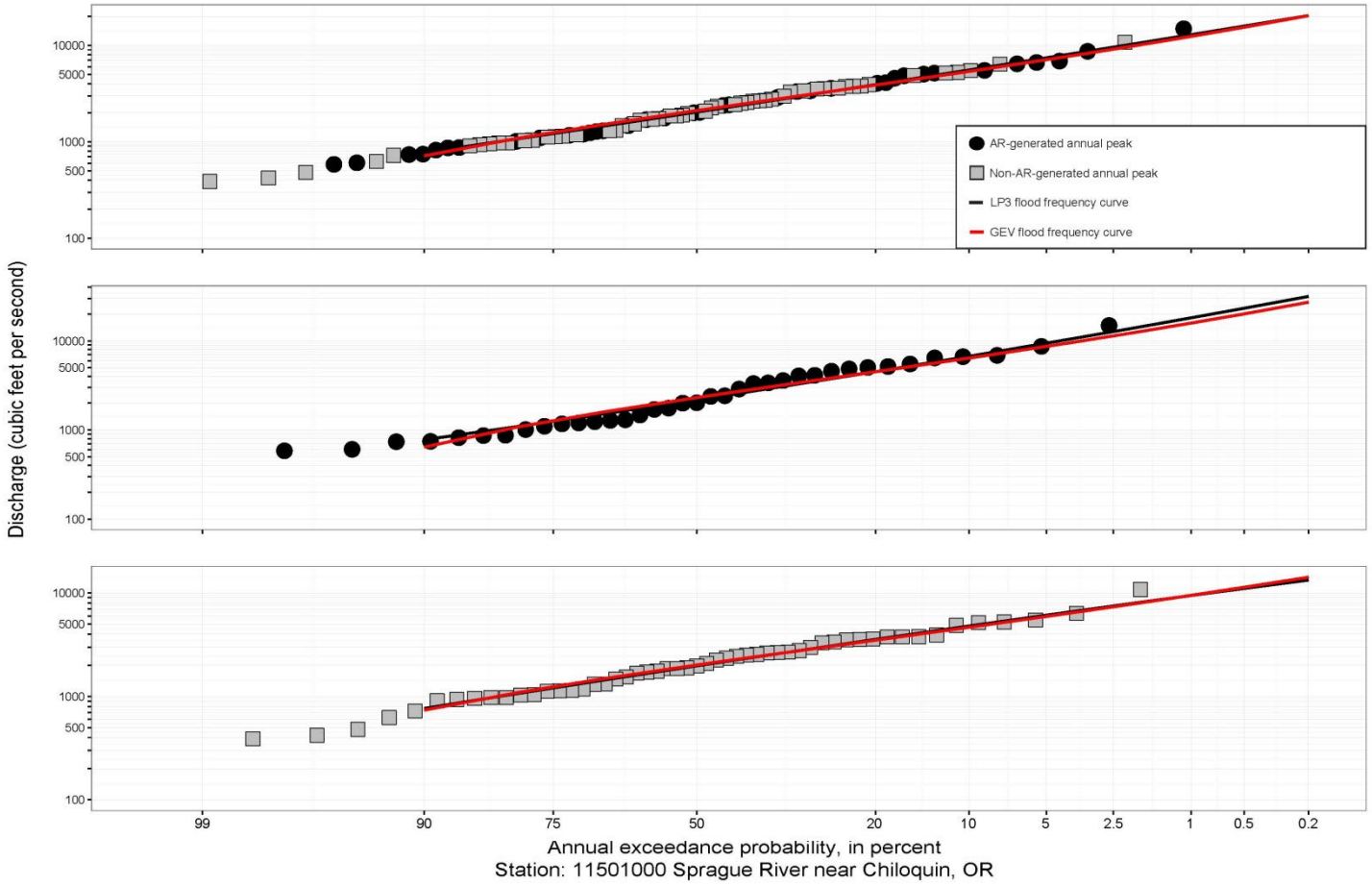

Figure S.30: Fitted frequency curves as described in figure caption S.1.

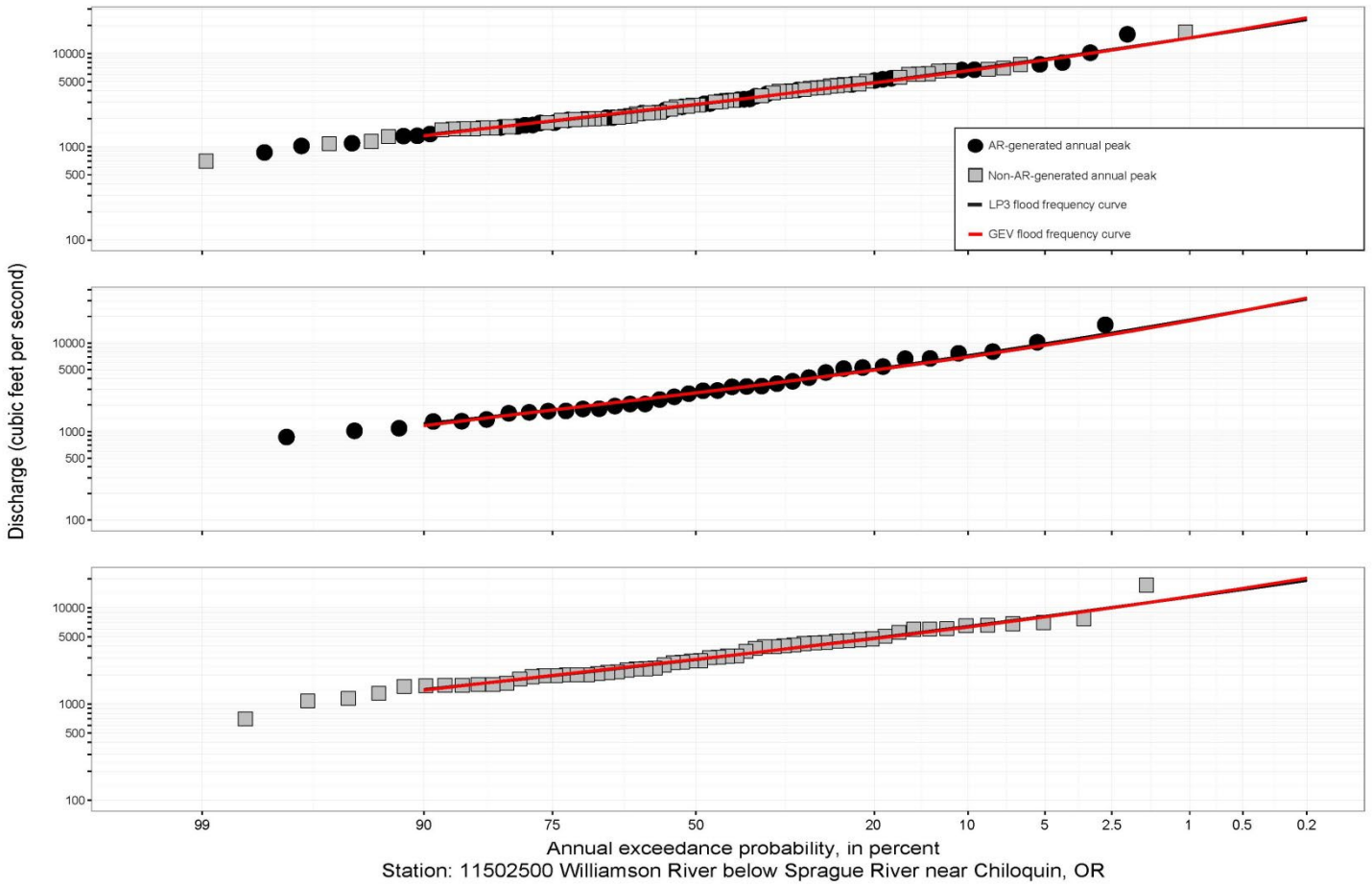

Figure S.31: Fitted frequency curves as described in figure caption S.1.

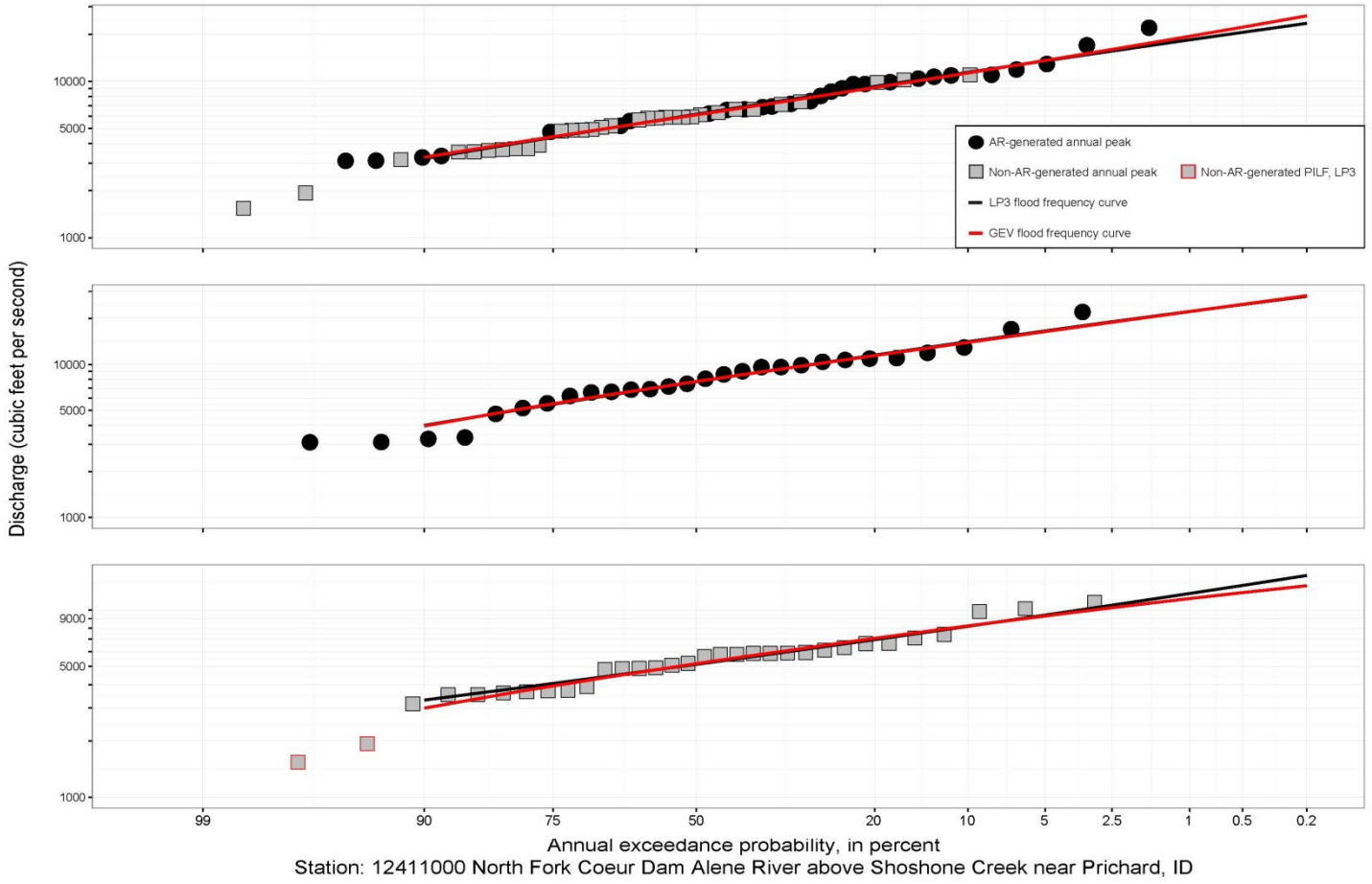

Figure S.32: Fitted frequency curves as described in figure caption S.5.

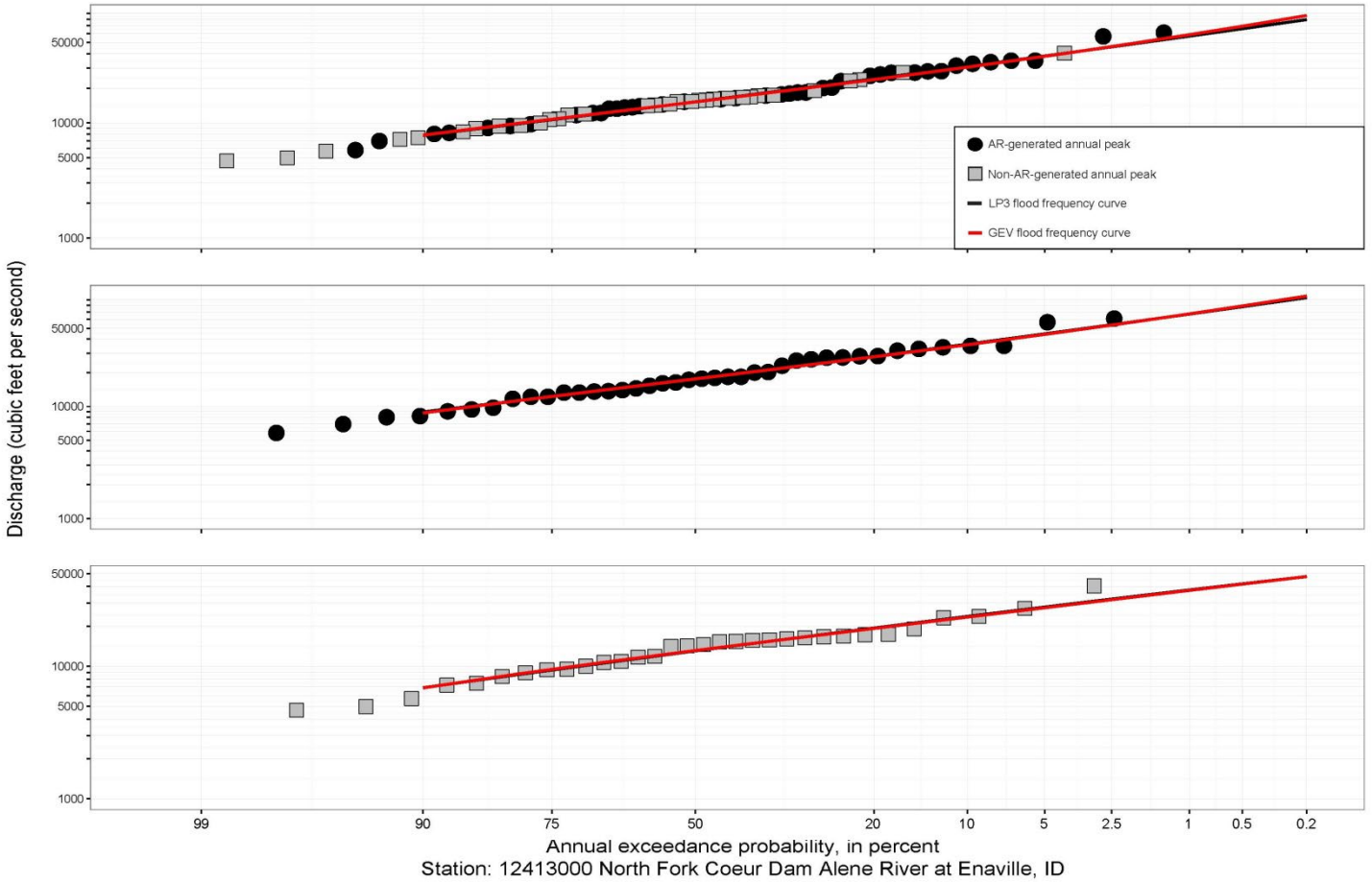

Figure S.33: Fitted frequency curves as described in figure caption S.1.

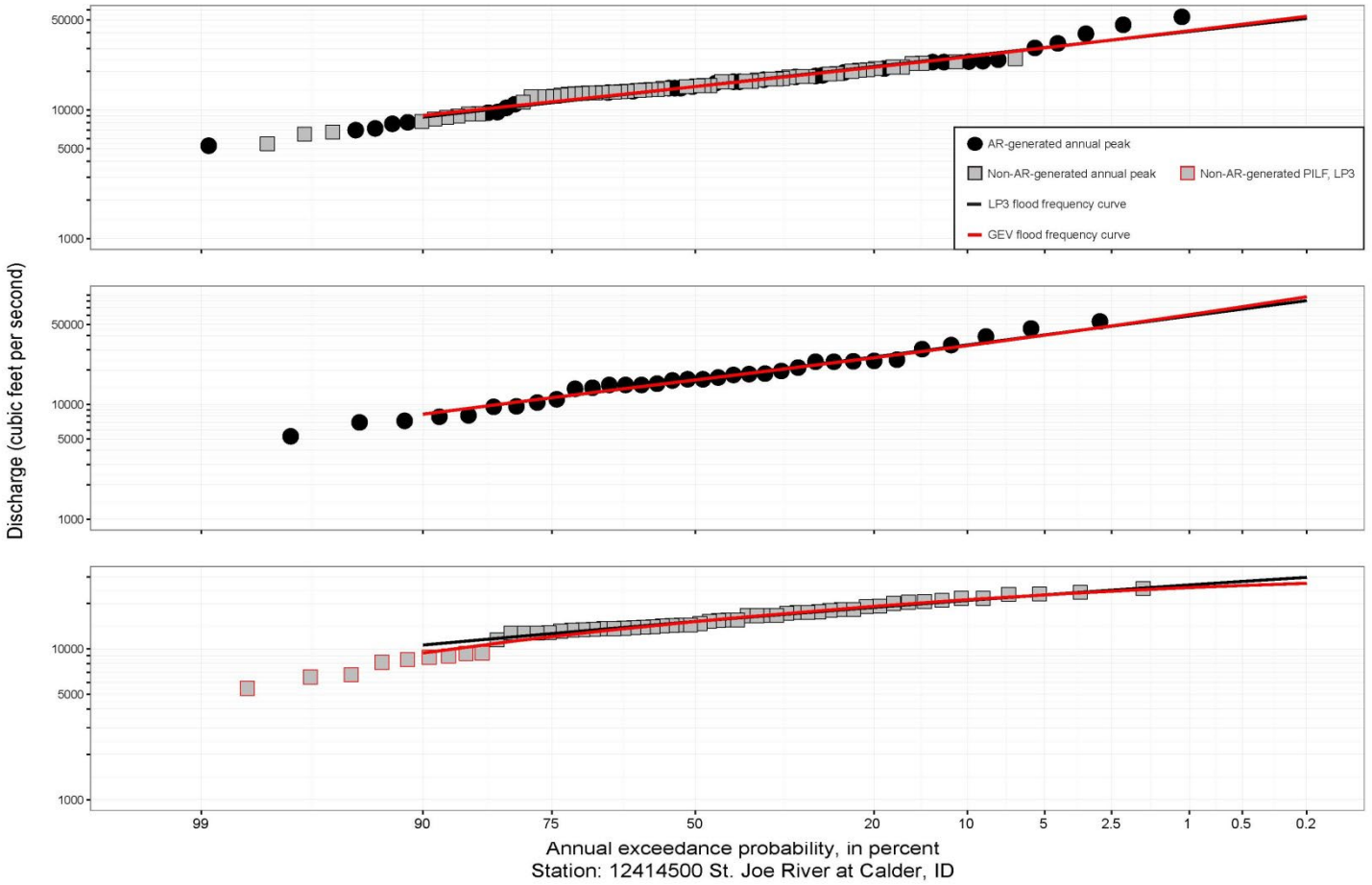

Figure S.34: Fitted frequency curves as described in figure caption S.5.

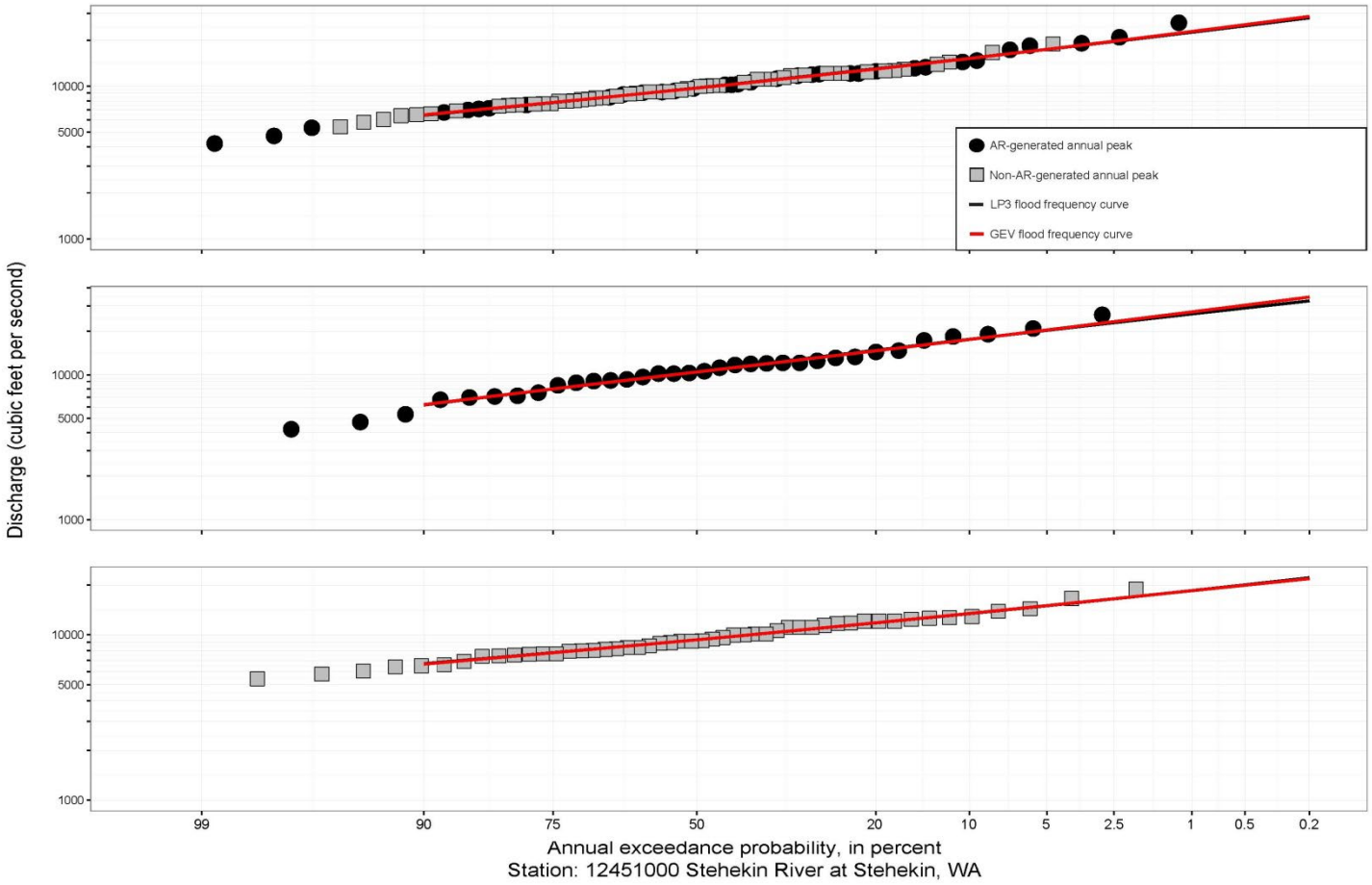

Figure S.35: Fitted frequency curves as described in figure caption S.1.

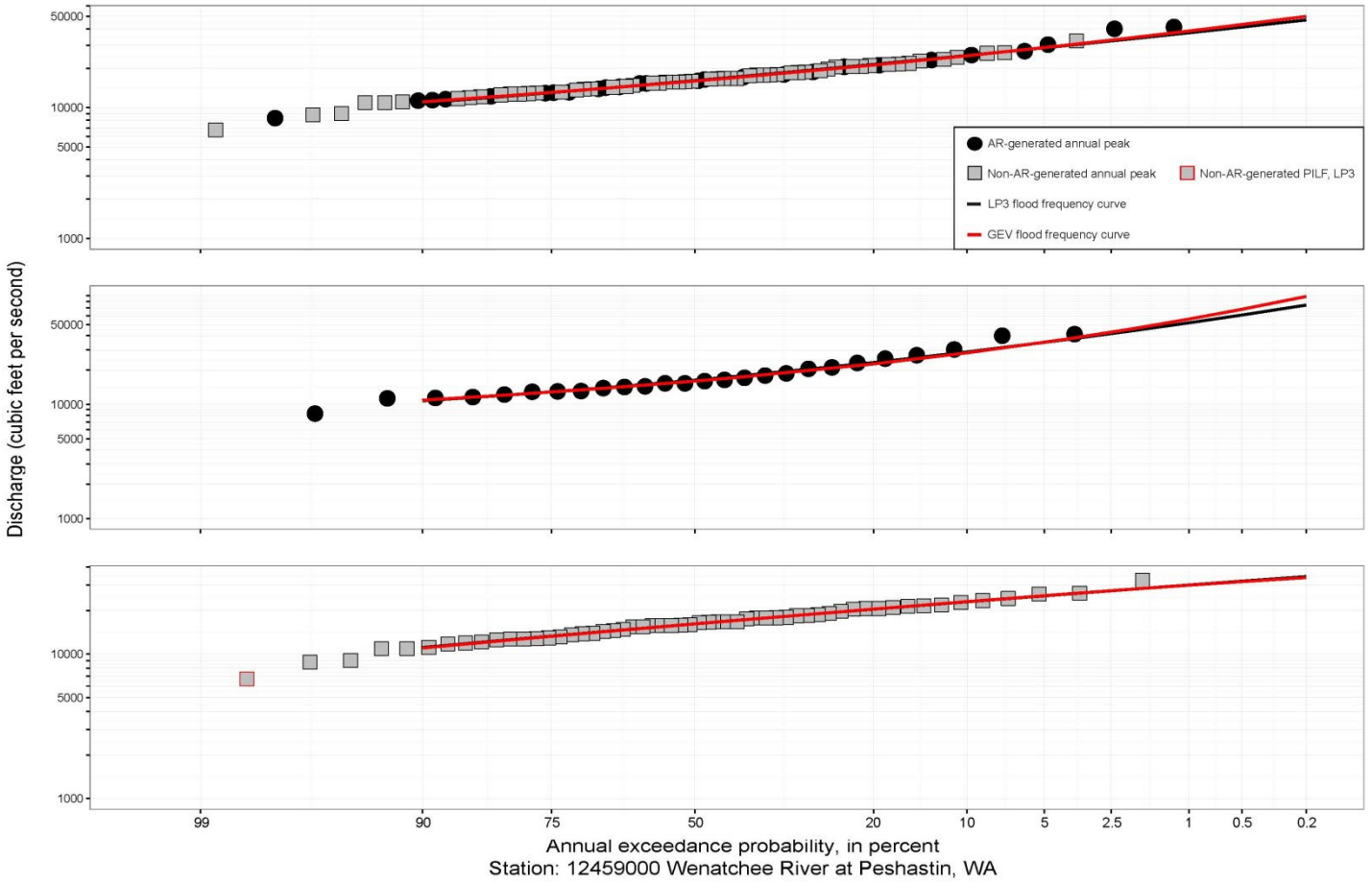

Figure S.36: Fitted frequency curves as described in figure caption S.5.

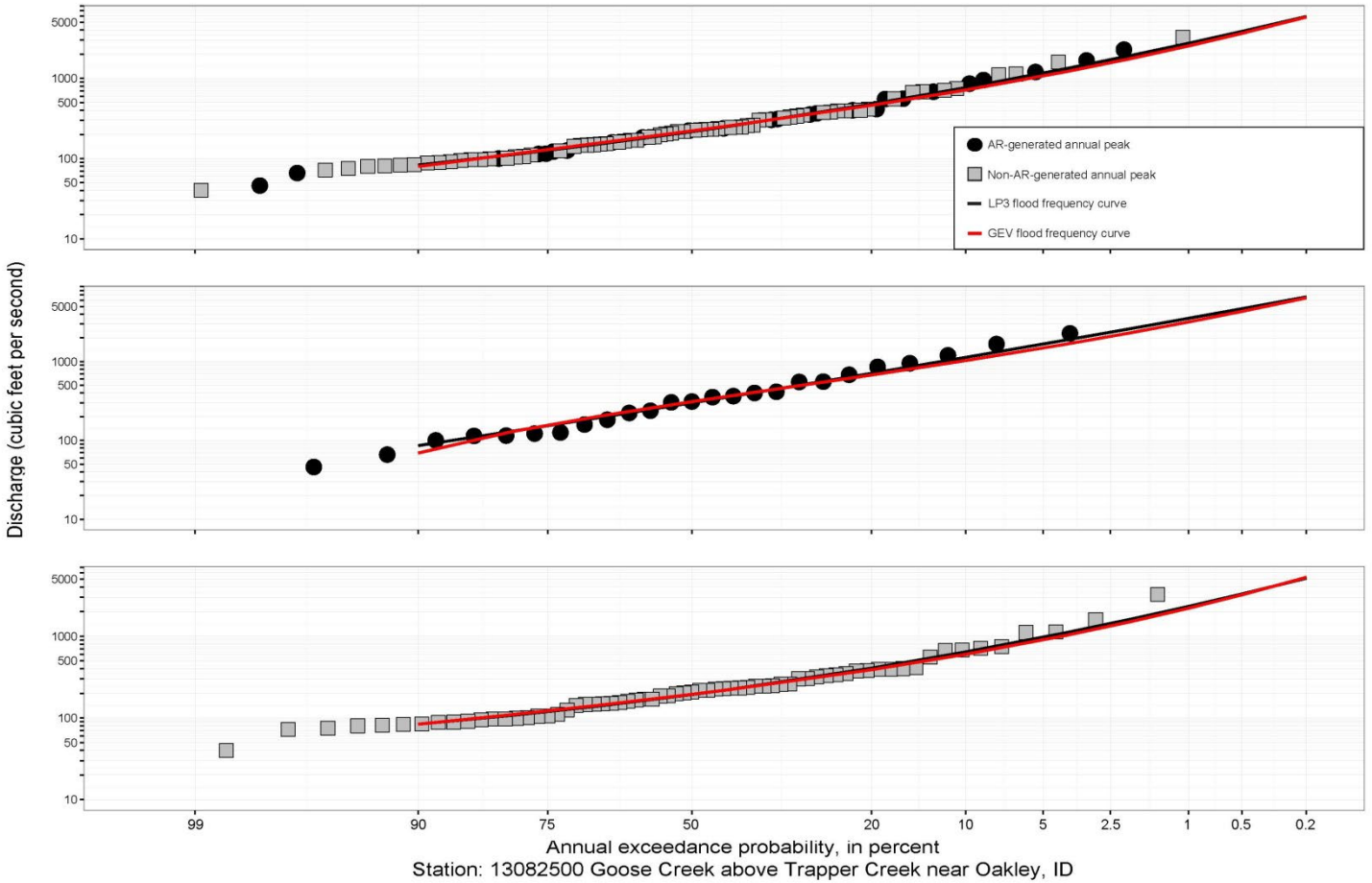

Figure S.37: Fitted frequency curves as described in figure caption S.1.

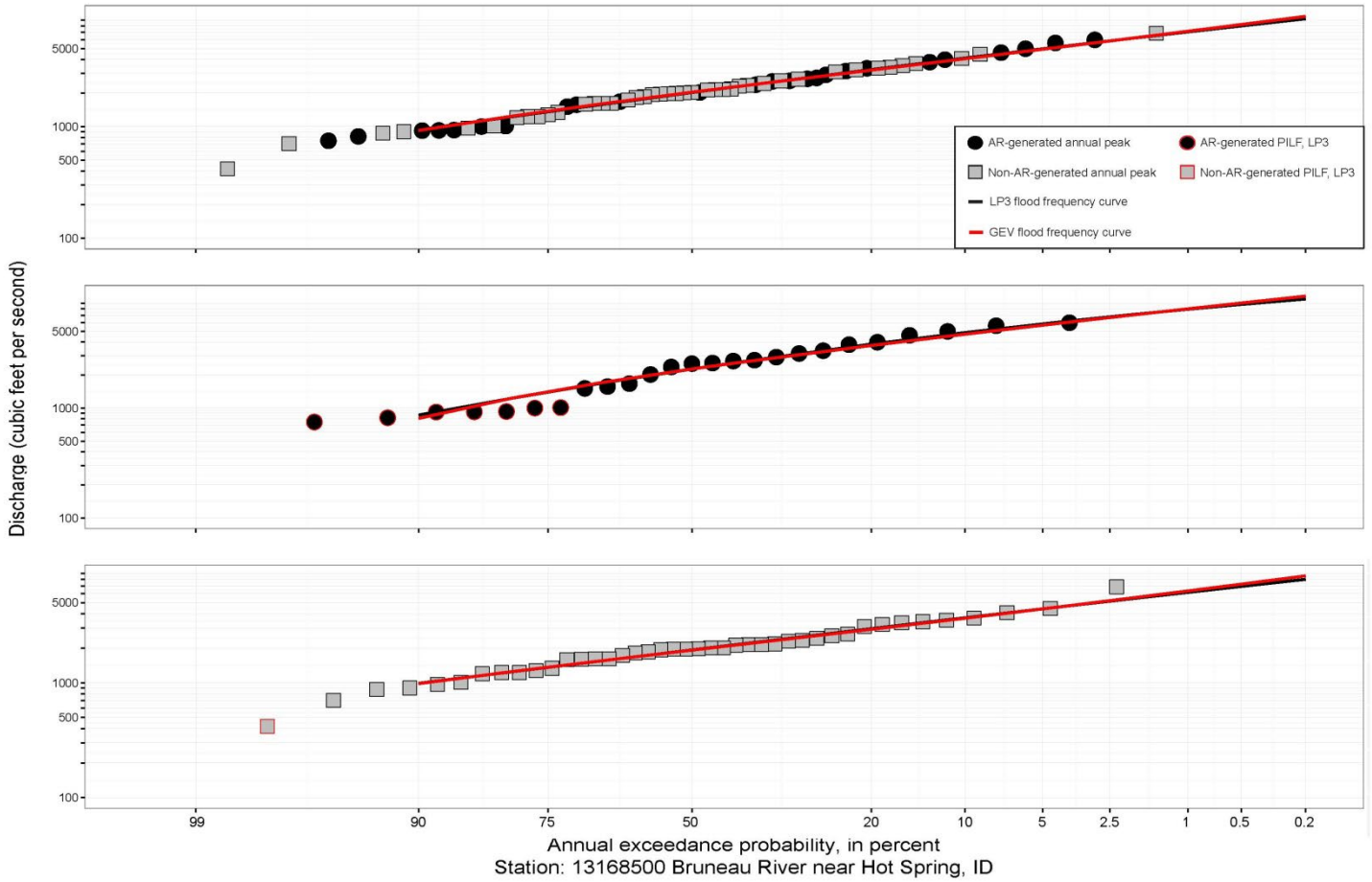

Figure S.38: Fitted frequency curves as described in figure caption S.5.

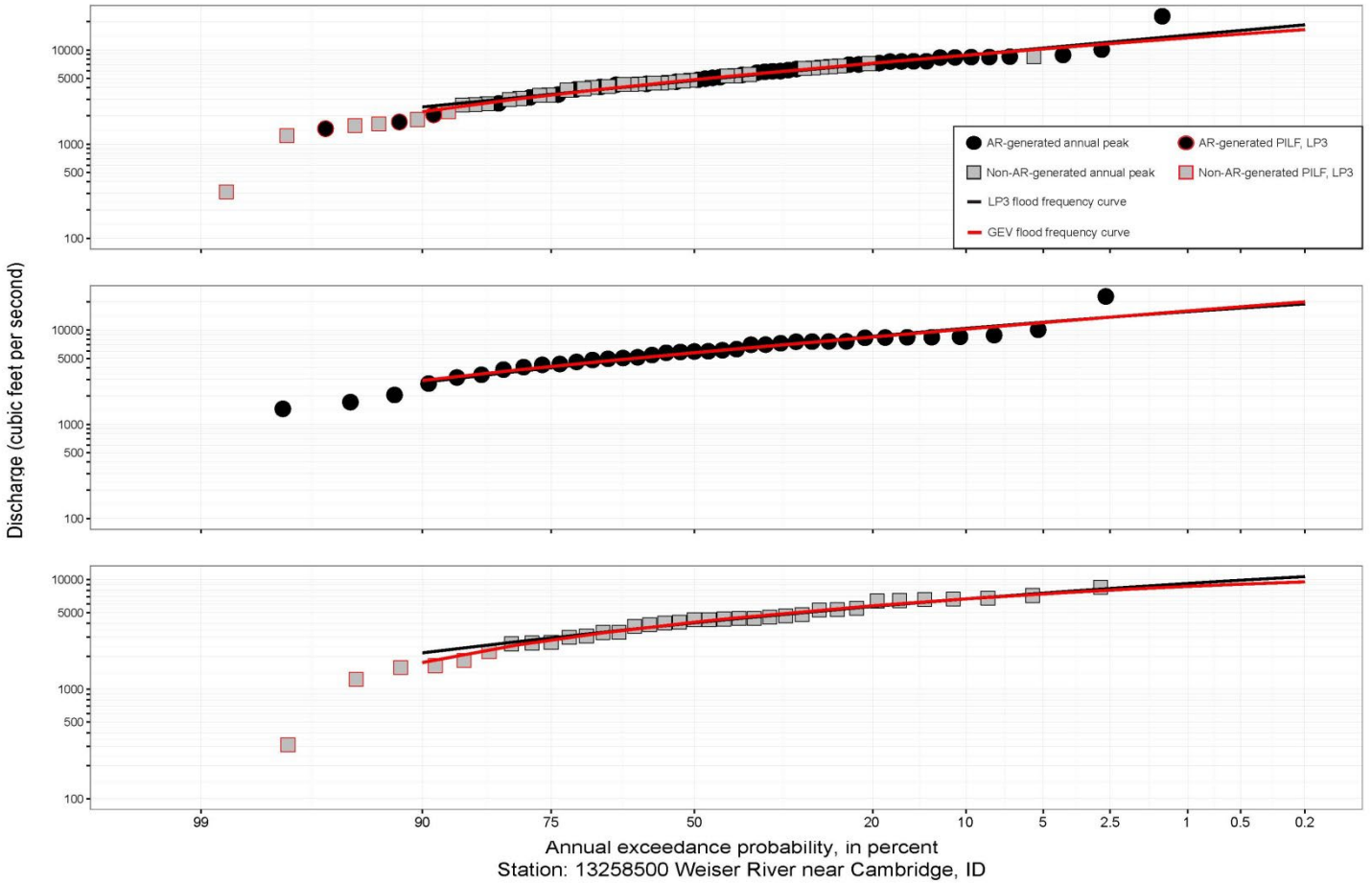

Figure S.39: Fitted frequency curves as described in figure caption S.5.

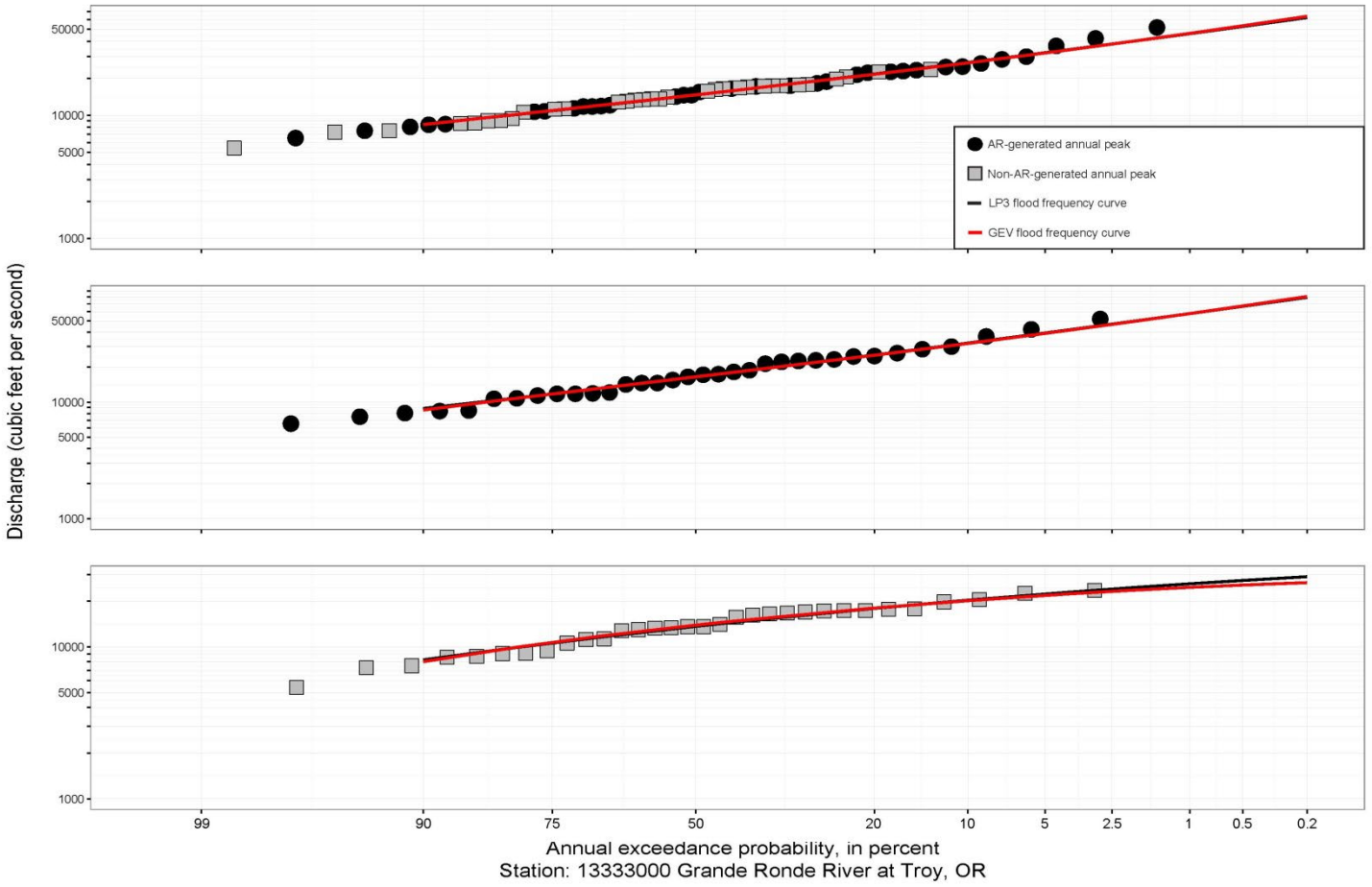

Figure S.40: Fitted frequency curves as described in figure caption S.1.

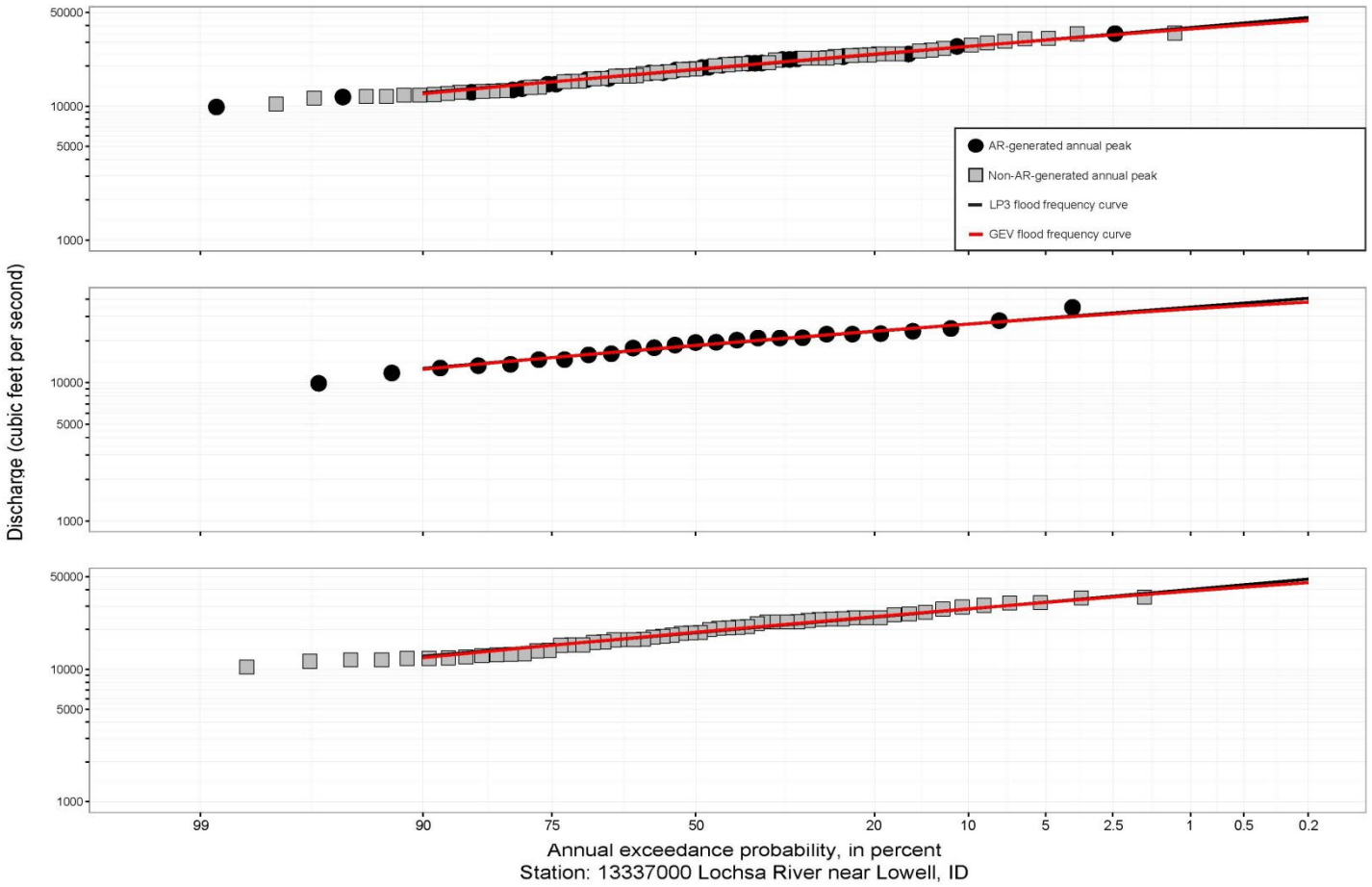

Figure S.41: Fitted frequency curves as described in figure caption S.1.

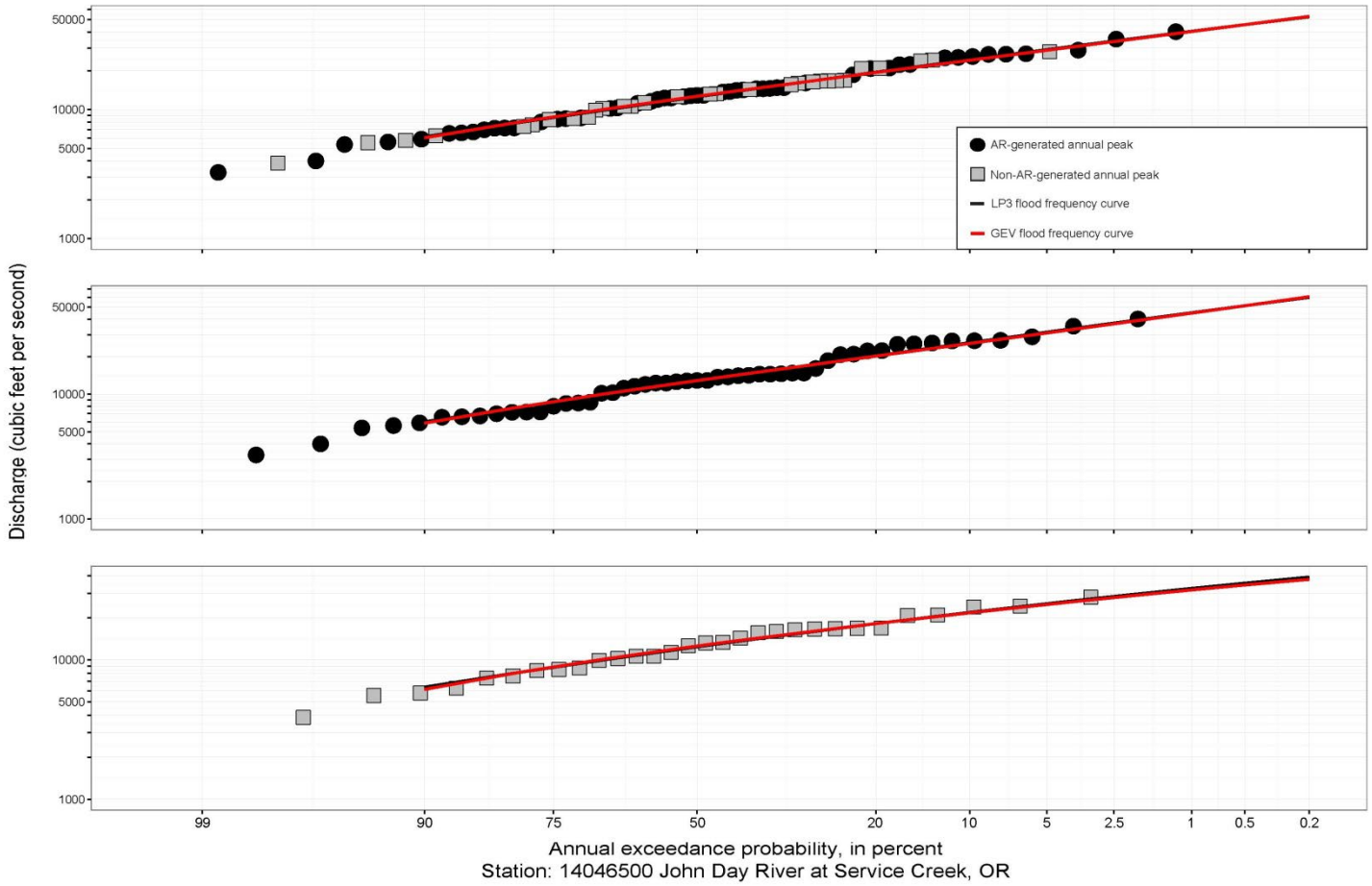

Figure S.42: Fitted frequency curves as described in figure caption S.1.

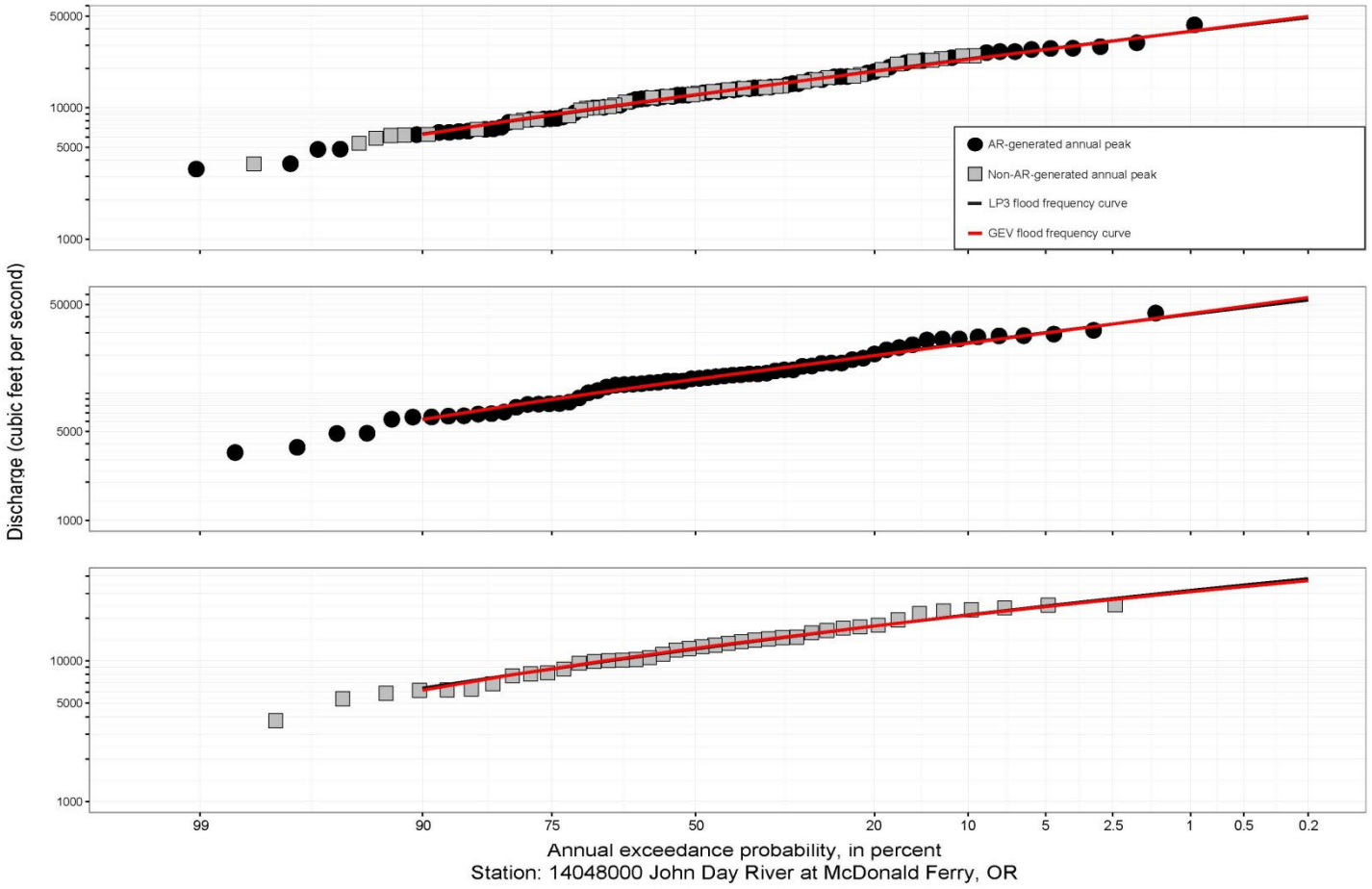

Figure S.43: Fitted frequency curves as described in figure caption S.1.

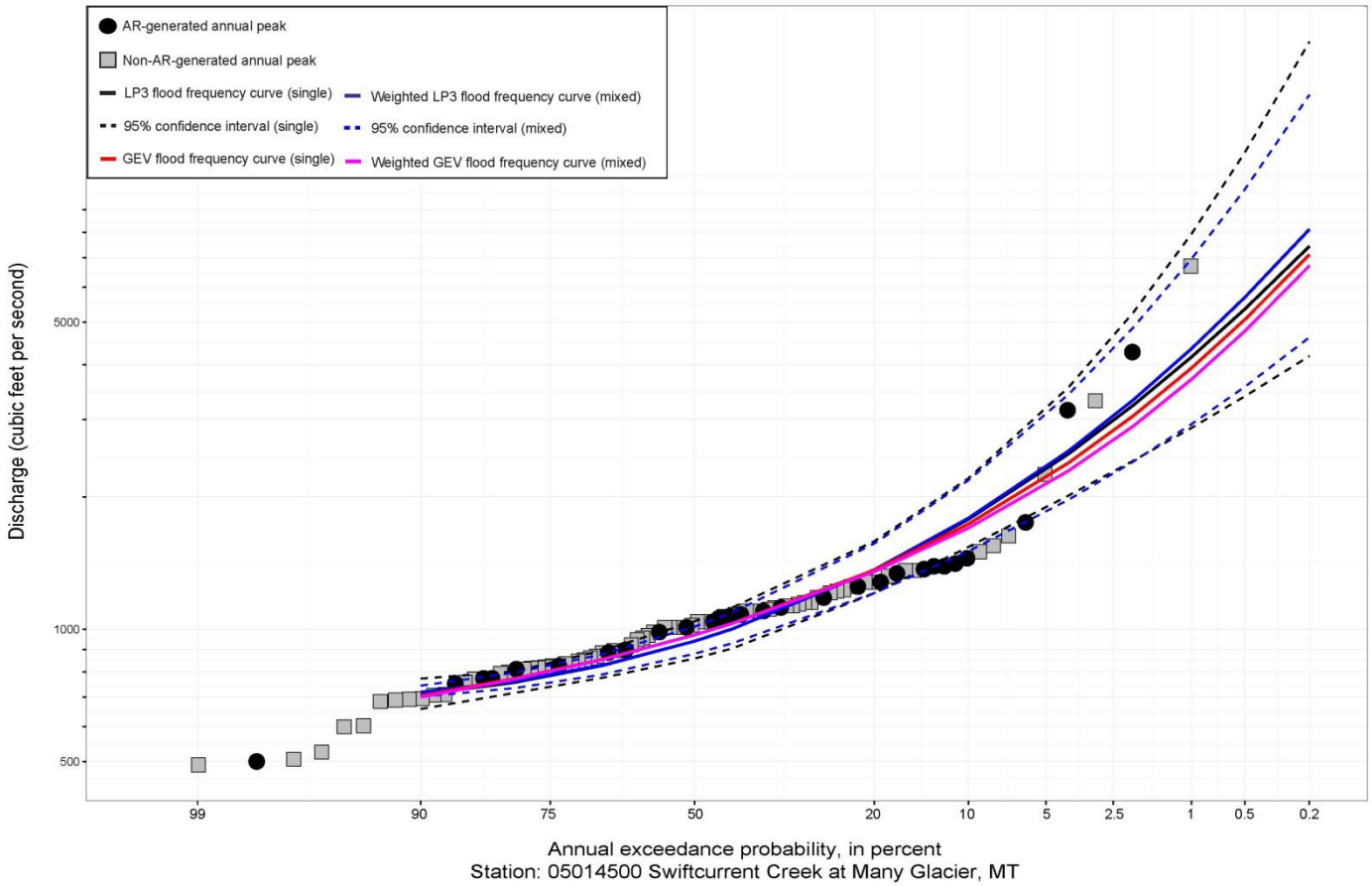

Figure S.44: Fitted frequency curves for (i) the log-Pearson Type III (LP3) single homogenous population (black line) comprised of AR (black circles)/ non-AR (grey squares) generated peaks with the corresponding 95% CIs (black dashed lines), (ii) the weighted LP3 mixed population approach (blue line) with the corresponding 95% CIs (dashed blue lines) that account for sampling and mixing uncertainties, (iii) the generalized extreme value (GEV) single homogenous population (red line), and (iv) the weighted GEV mixed population approach (magenta line).

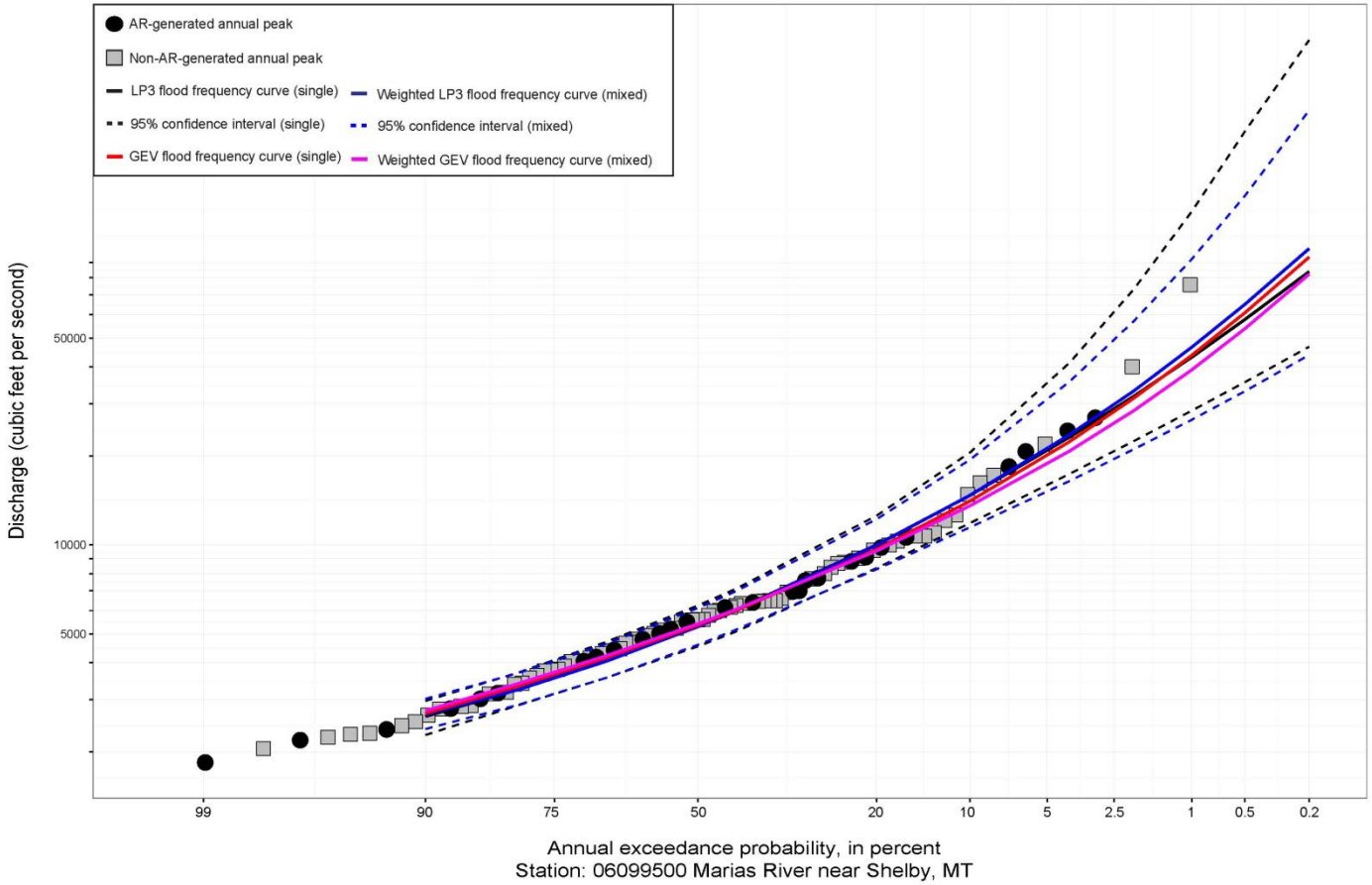

Figure S.45: Fitted frequency curves as described in figure caption S.44.

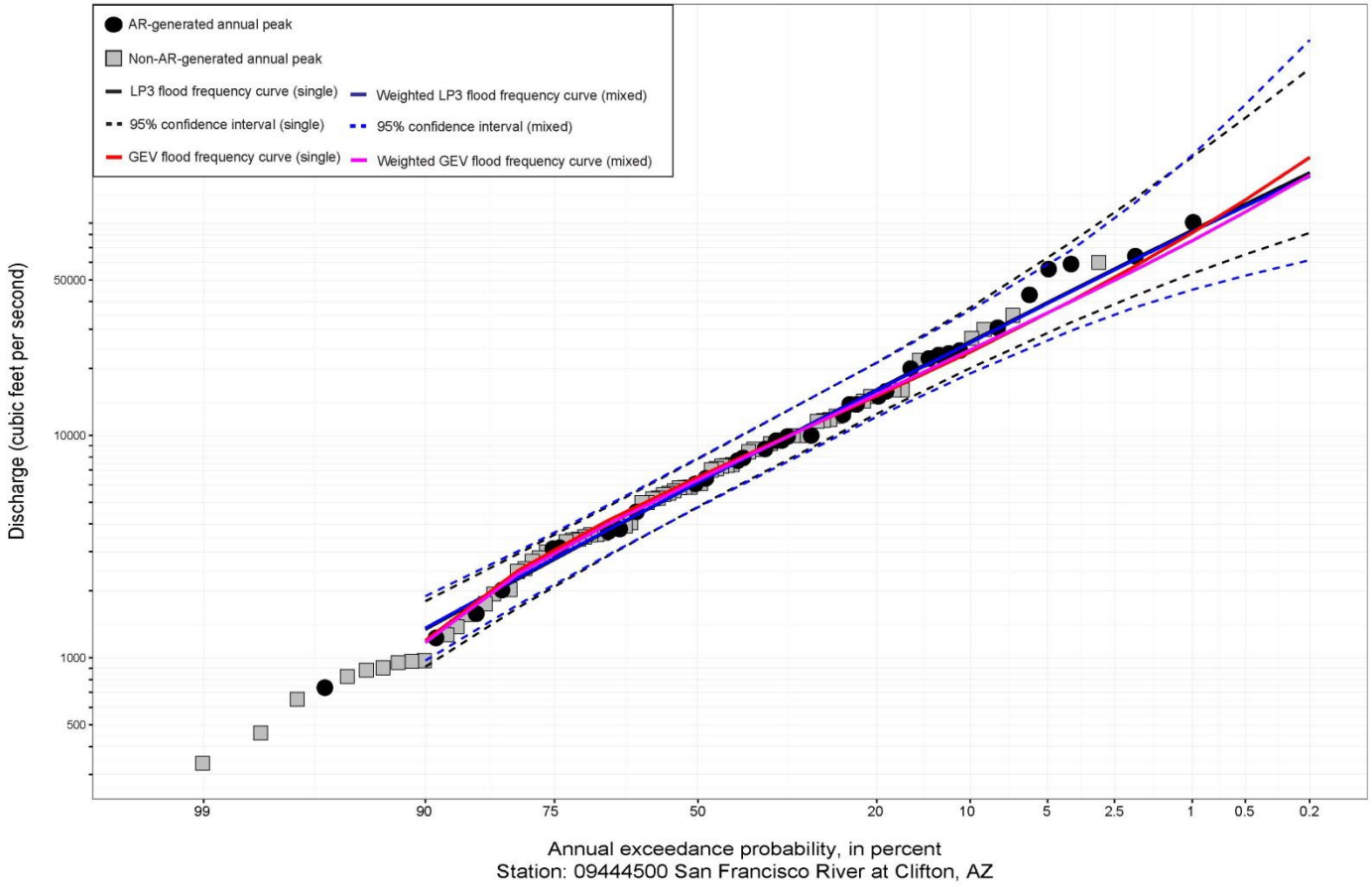

Figure S.46: Fitted frequency curves as described in figure caption S.44.

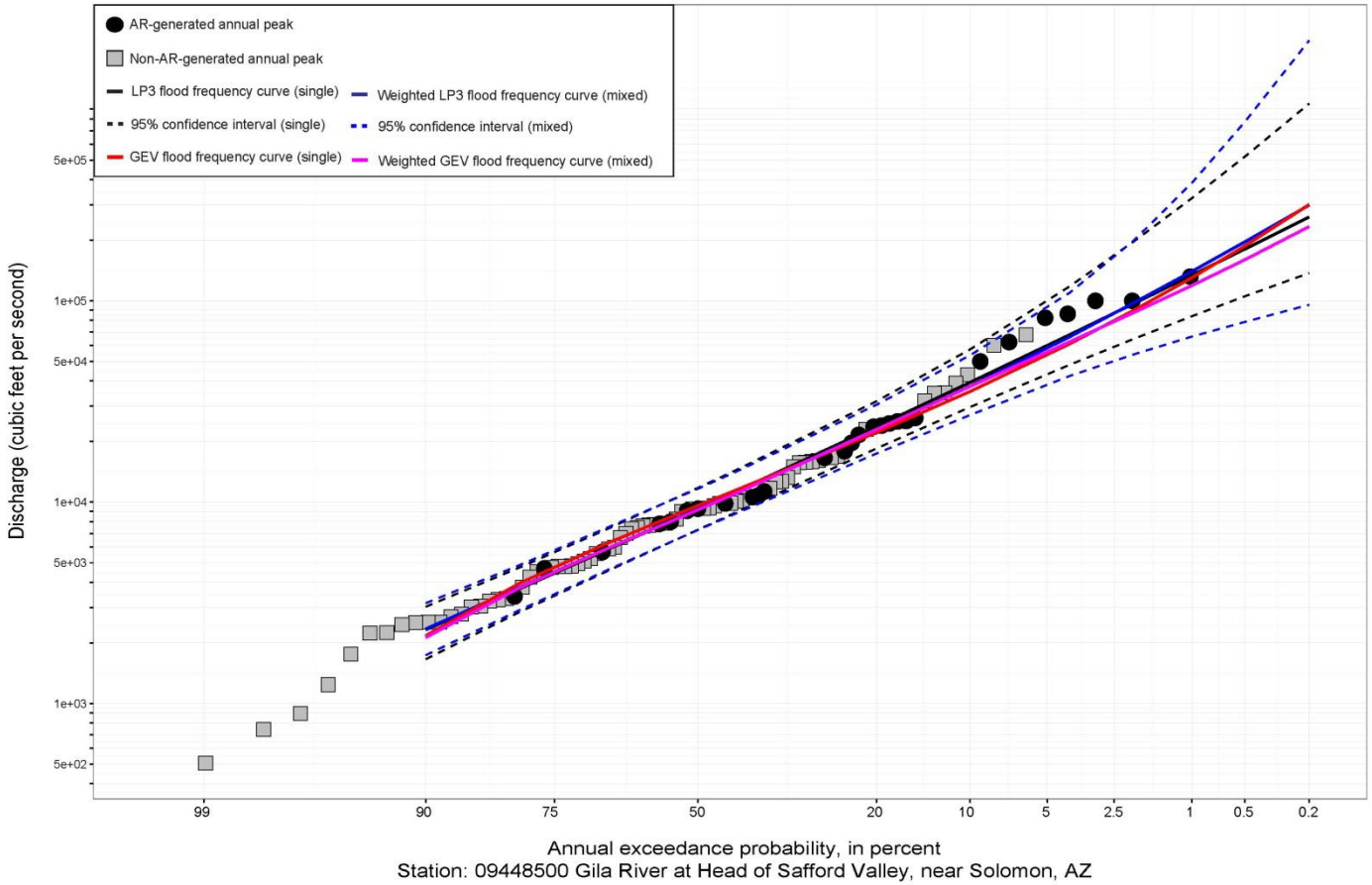

Figure S.47: Fitted frequency curves as described in figure caption S.44.

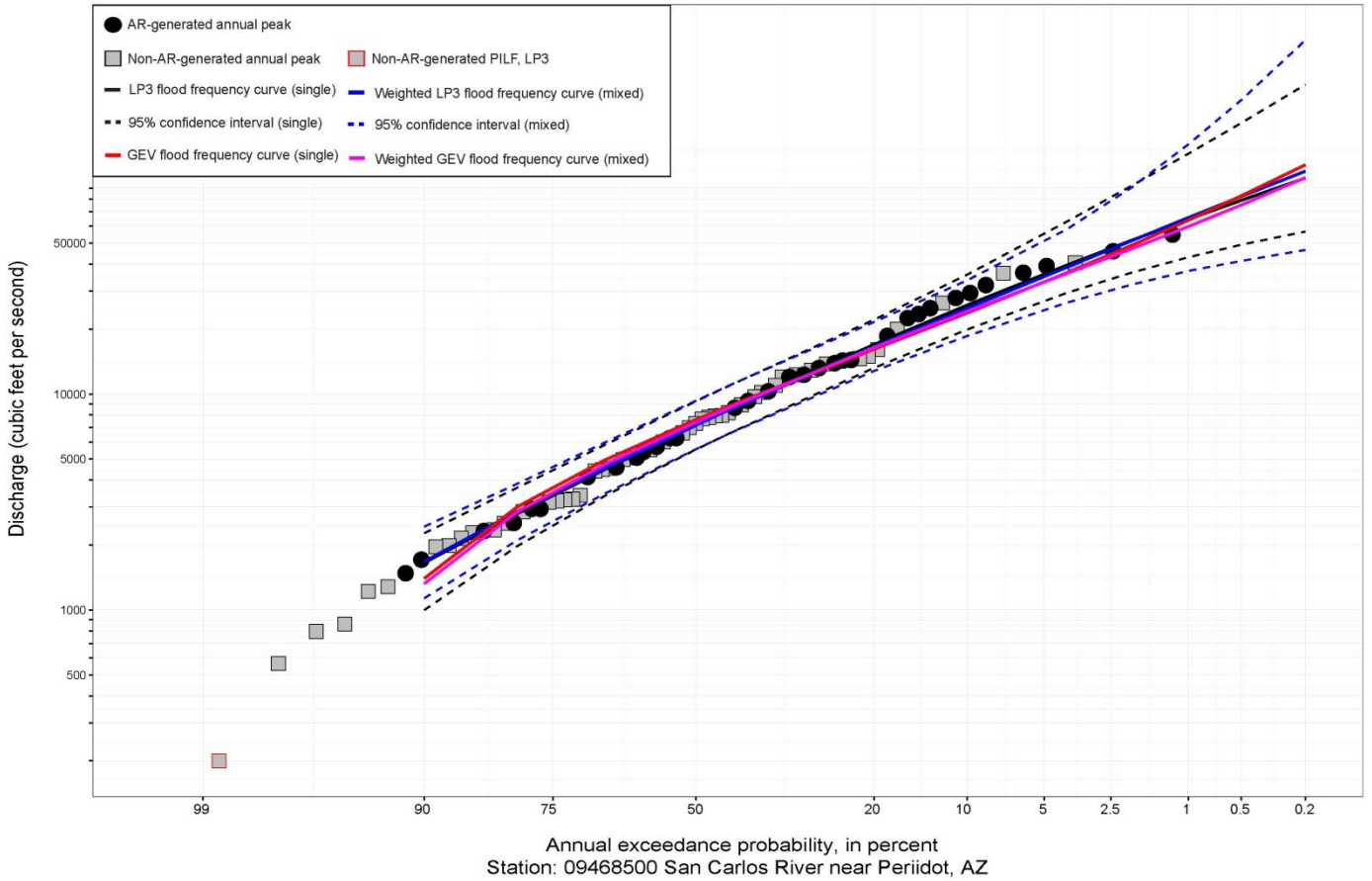

Figure S.48: Fitted frequency curves for (i) the log-Pearson Type III (LP3) single homogenous population (black line) comprised of AR (black circles)/ non-AR (grey squares) generated peaks with the corresponding 95% CIs (black dashed lines), (ii) the weighted LP3 mixed population approach (blue line) with the corresponding 95% CIs (dashed blue lines) that account for sampling and mixing uncertainties, (iii) the generalized extreme value (GEV) single homogenous population (red line), and (iv) the weighted GEV mixed population approach (magenta line). Potentially influential low floods (PILFs) identified among the LP3 single population are highlighted with red.

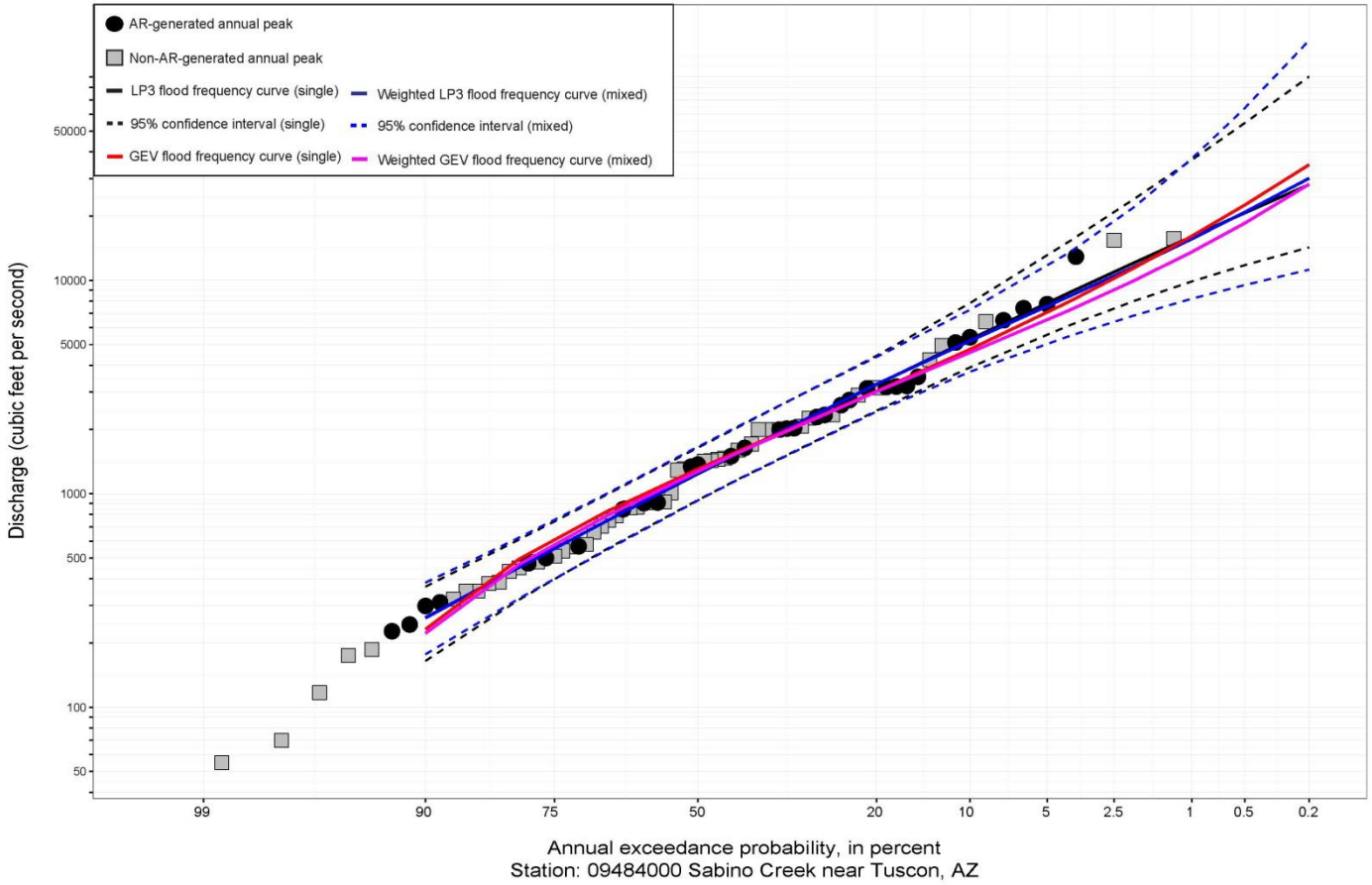

Figure S.49: Fitted frequency curves as described in figure caption S.44.

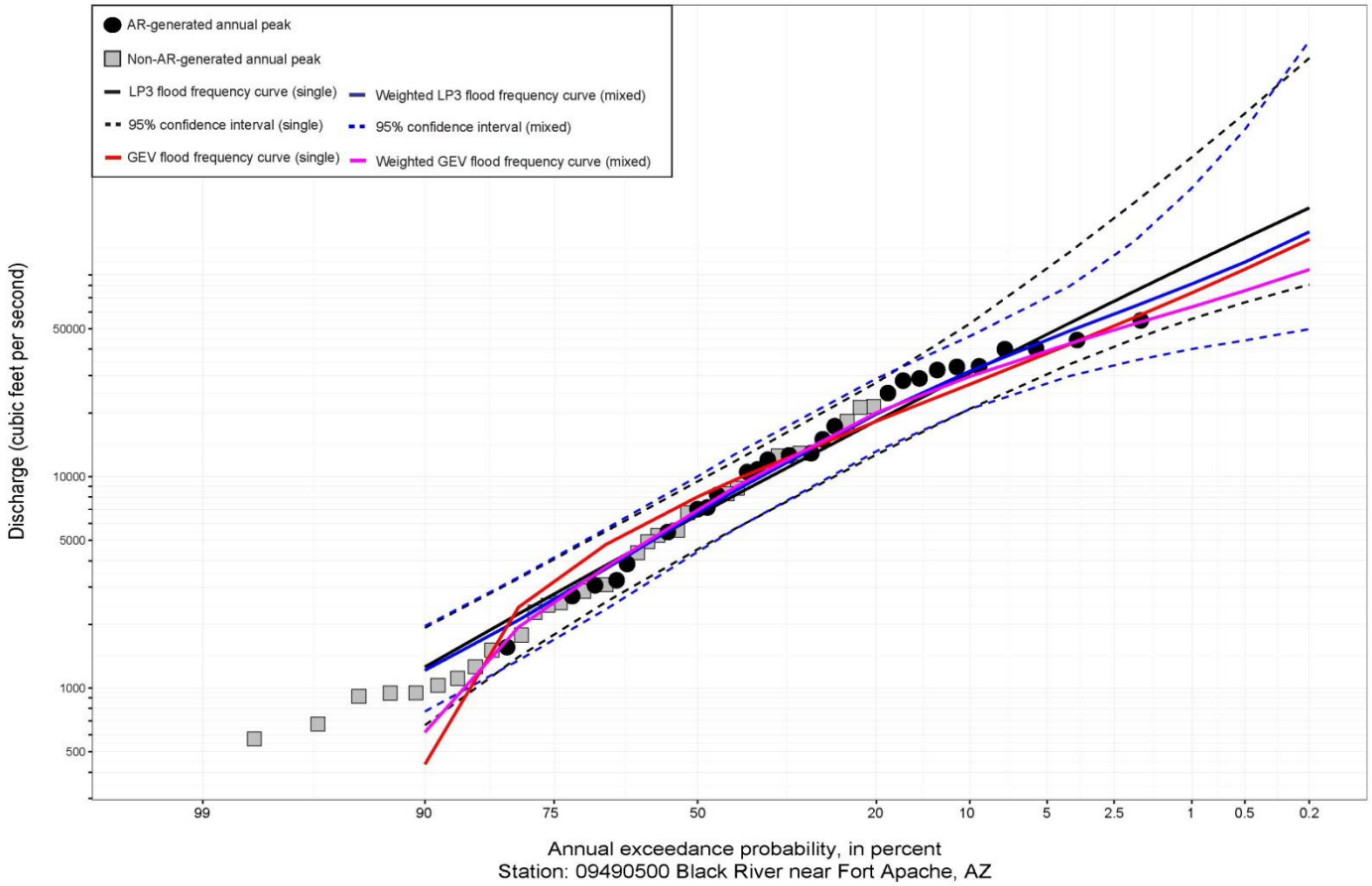

Figure S.50: Fitted frequency curves as described in figure caption S.44.

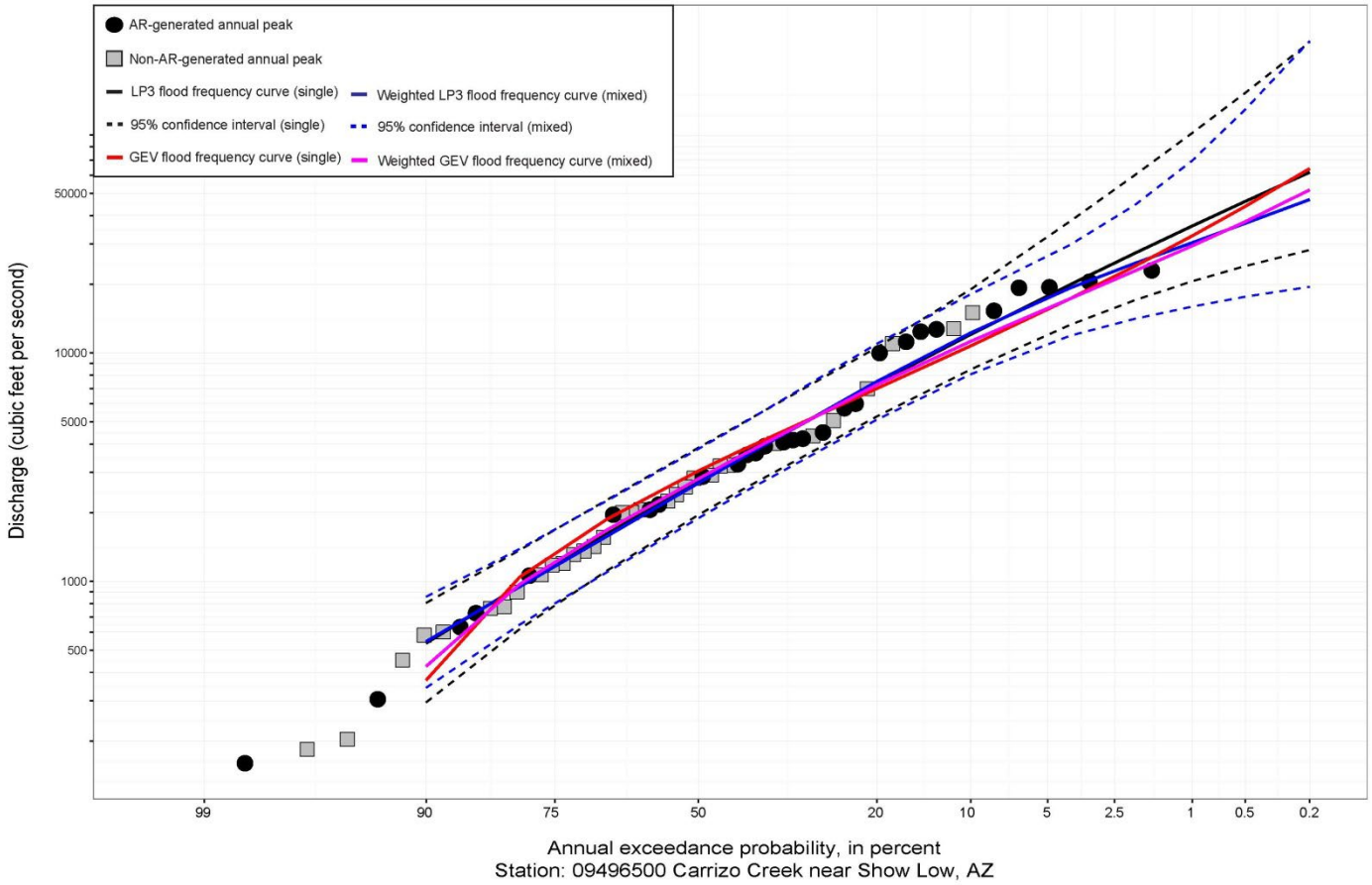

Figure S.51: Fitted frequency curves as described in figure caption S.44.

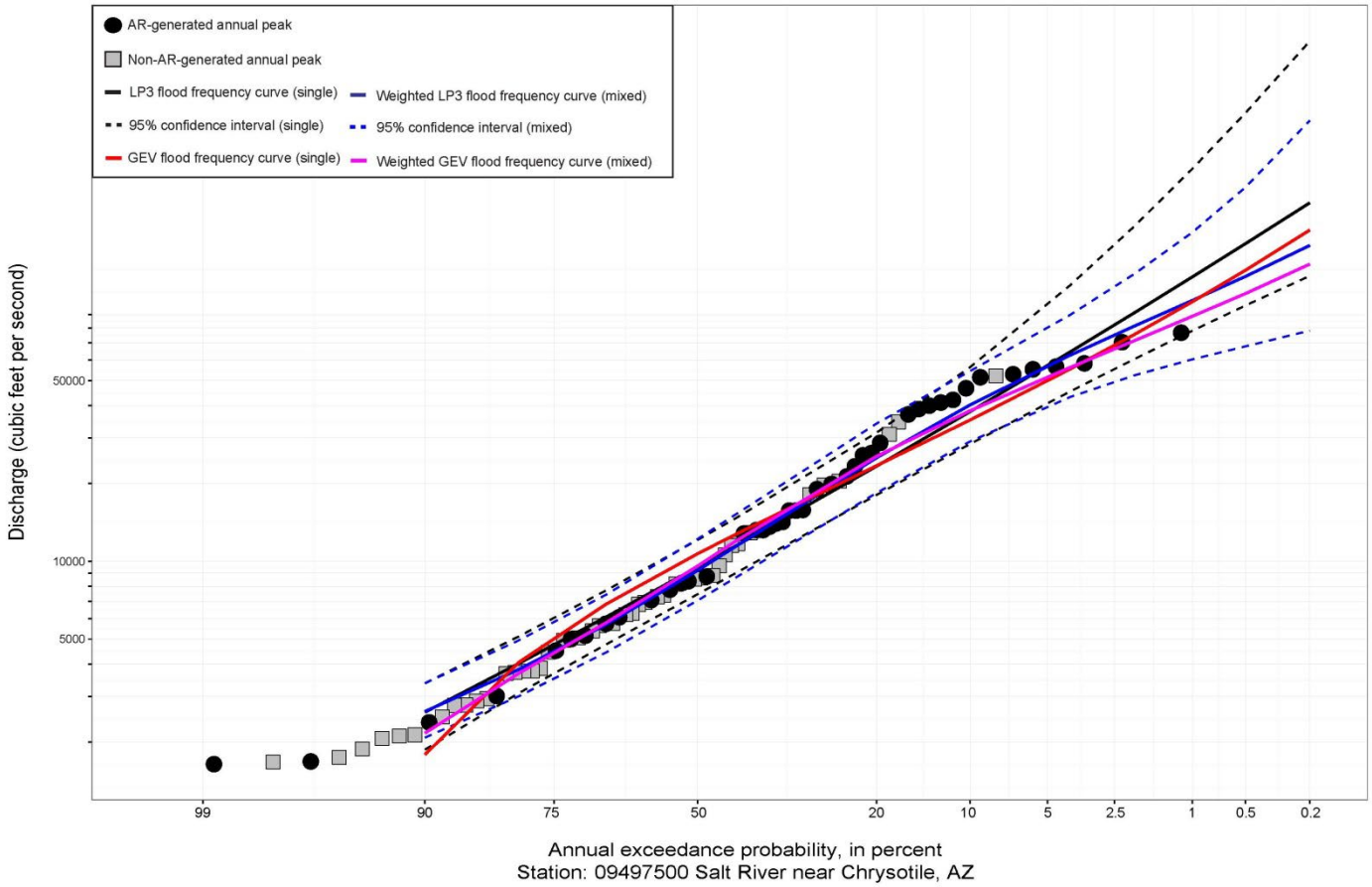

Figure S.52: Fitted frequency curves as described in figure caption S.44.

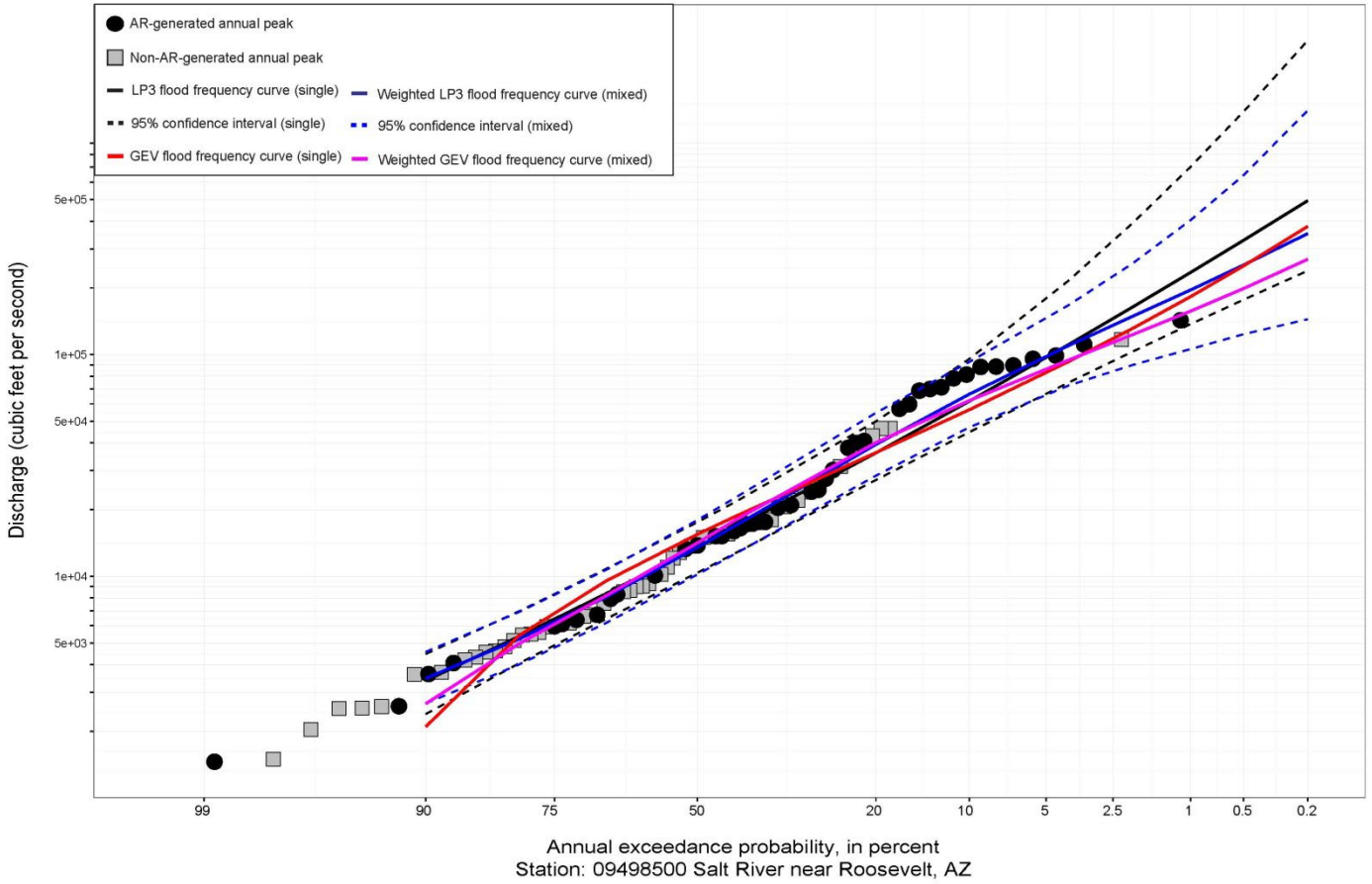

Figure S.53: Fitted frequency curves as described in figure caption S.44.

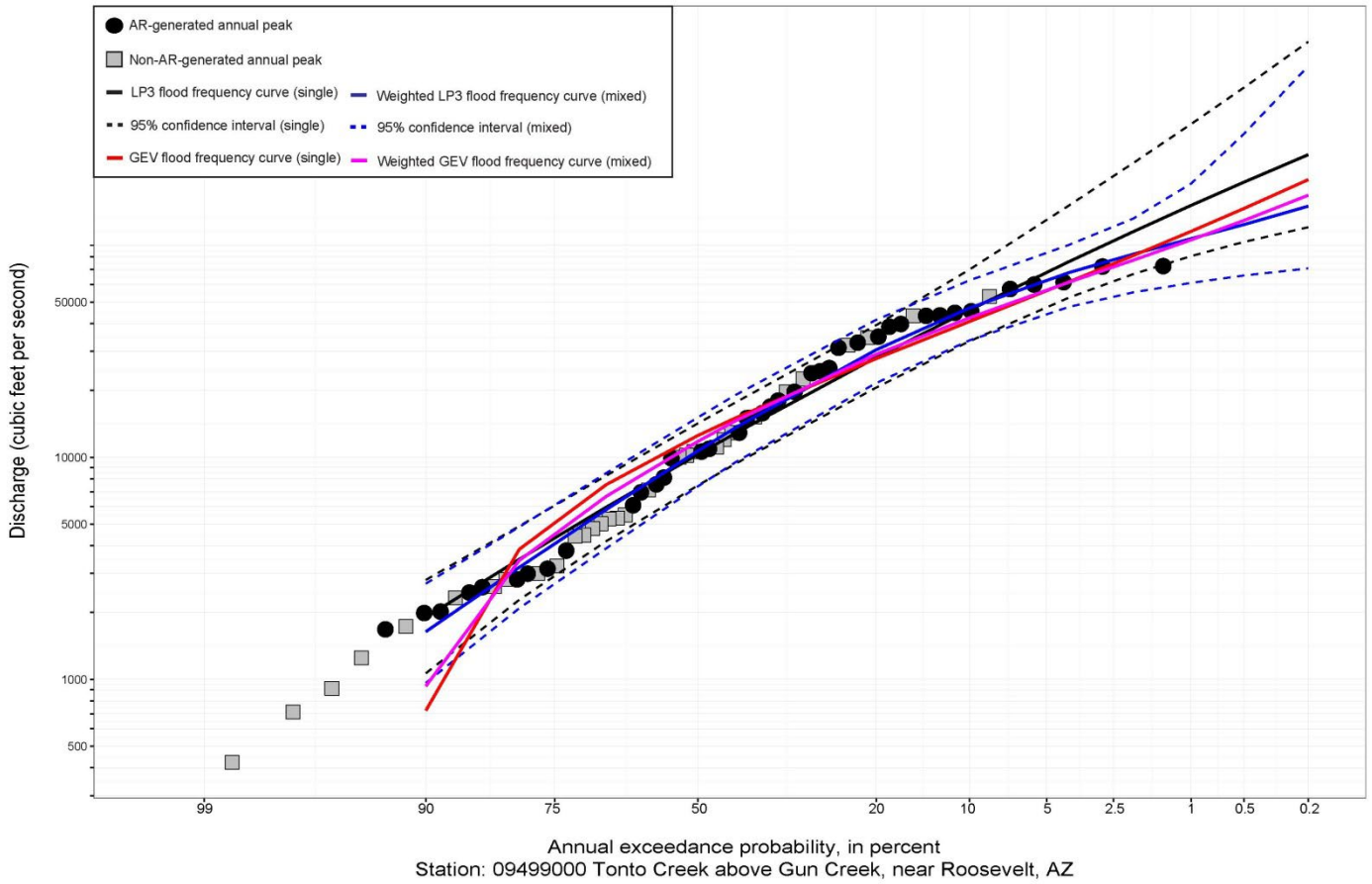

Figure S.54: Fitted frequency curves as described in figure caption S.44.

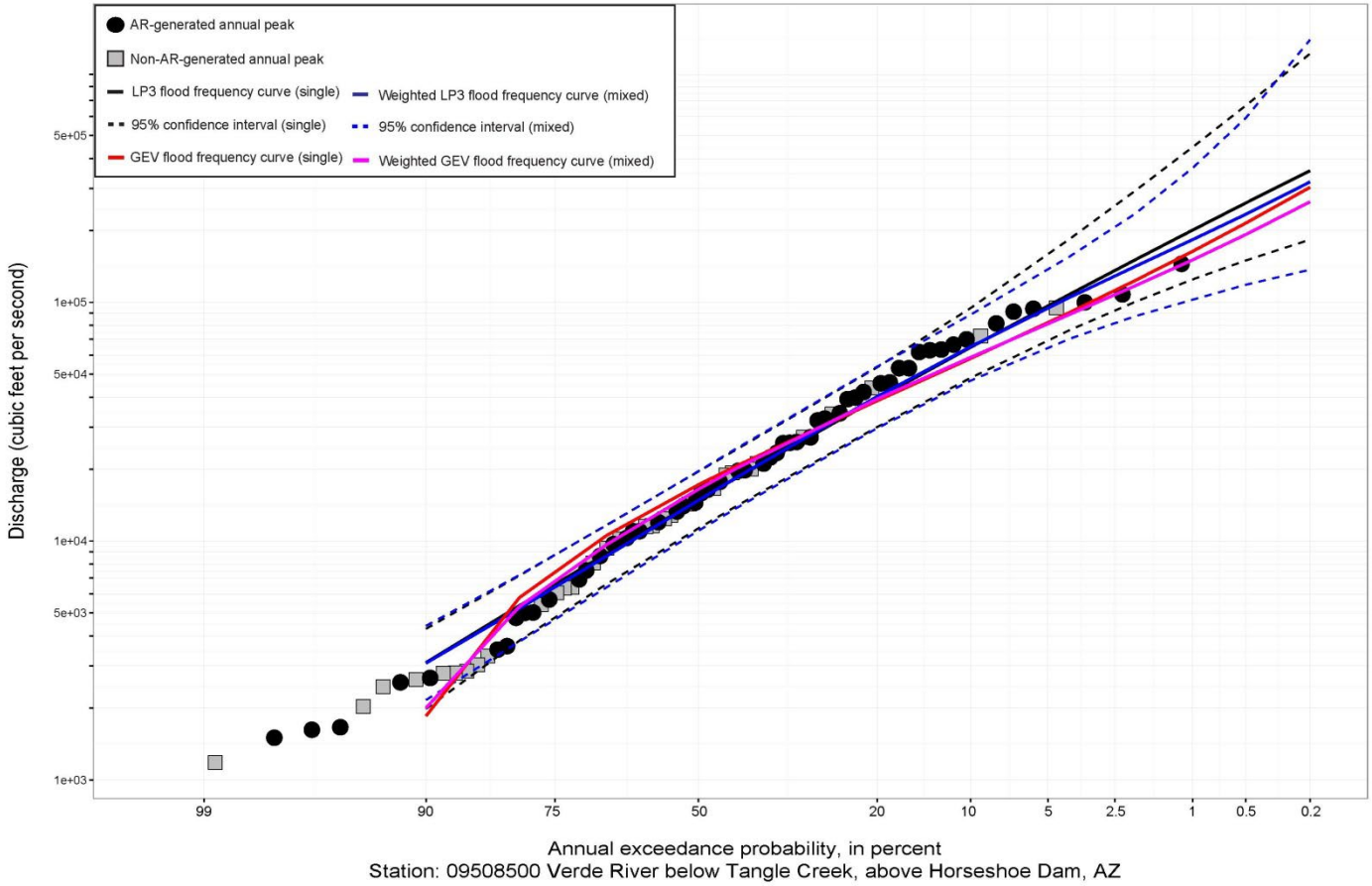

Figure S.55: Fitted frequency curves as described in figure caption S.44.

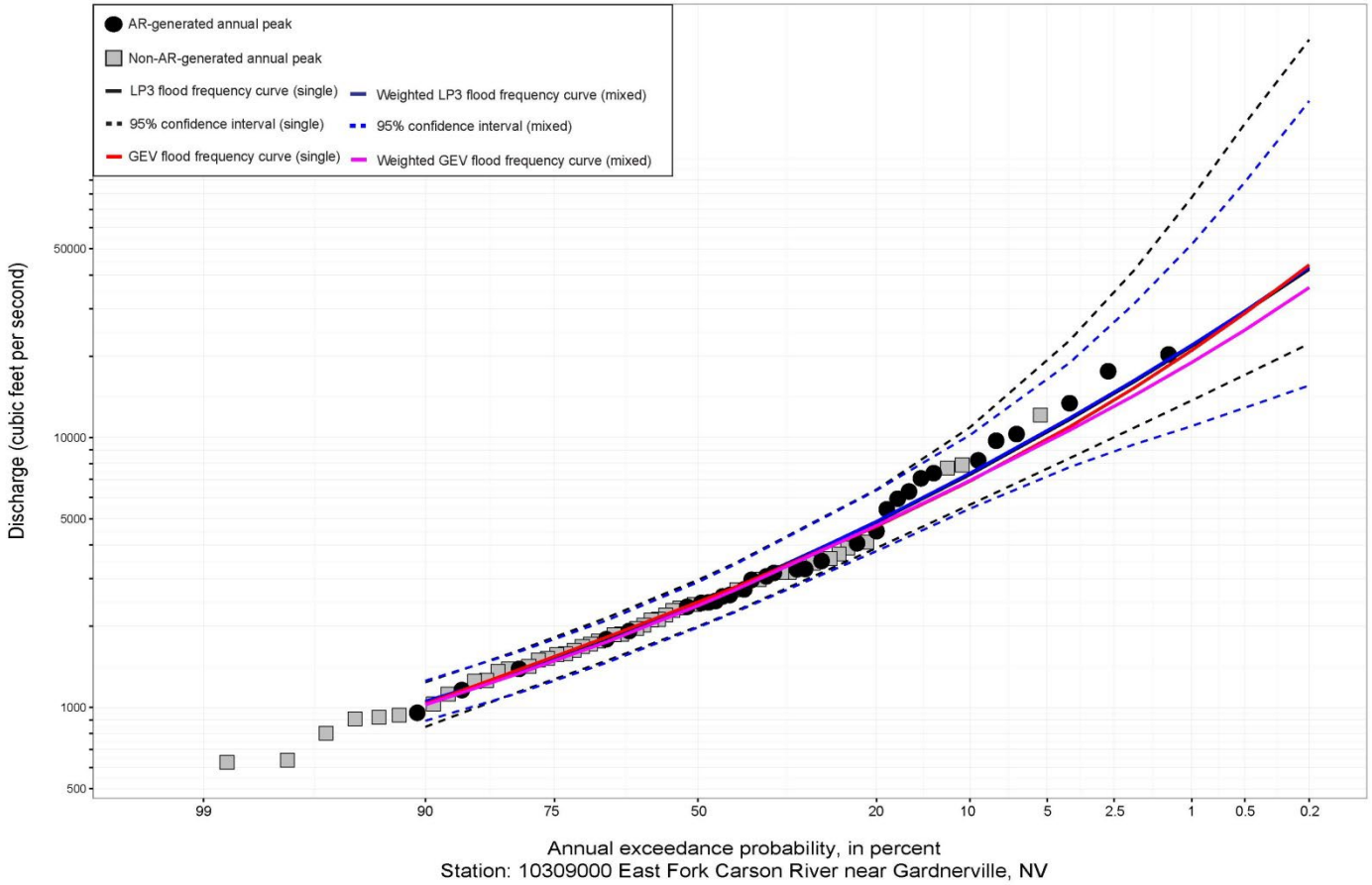

Figure S.56: Fitted frequency curves as described in figure caption S.44.

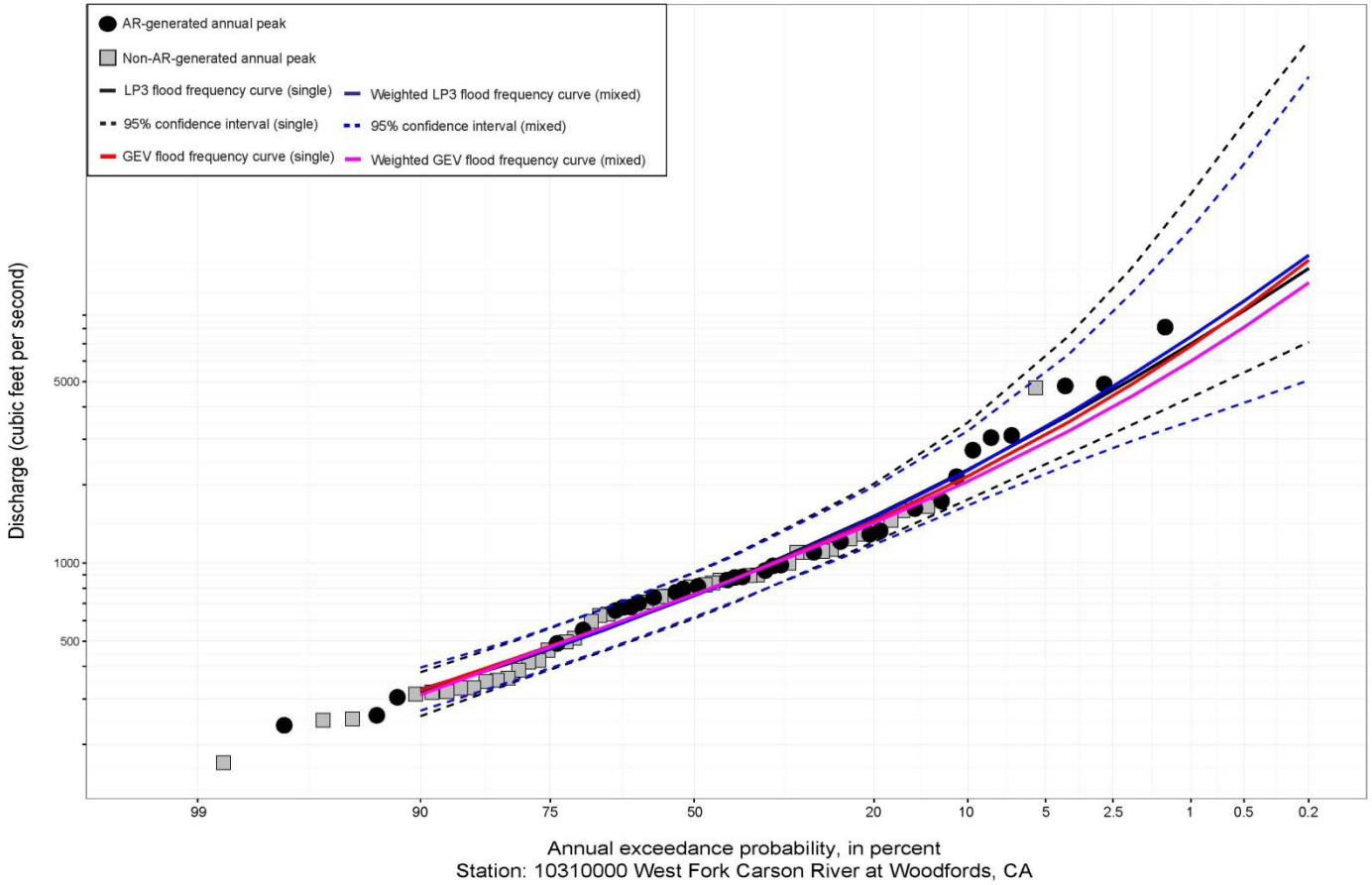

Figure S.57: Fitted frequency curves as described in figure caption S.44.

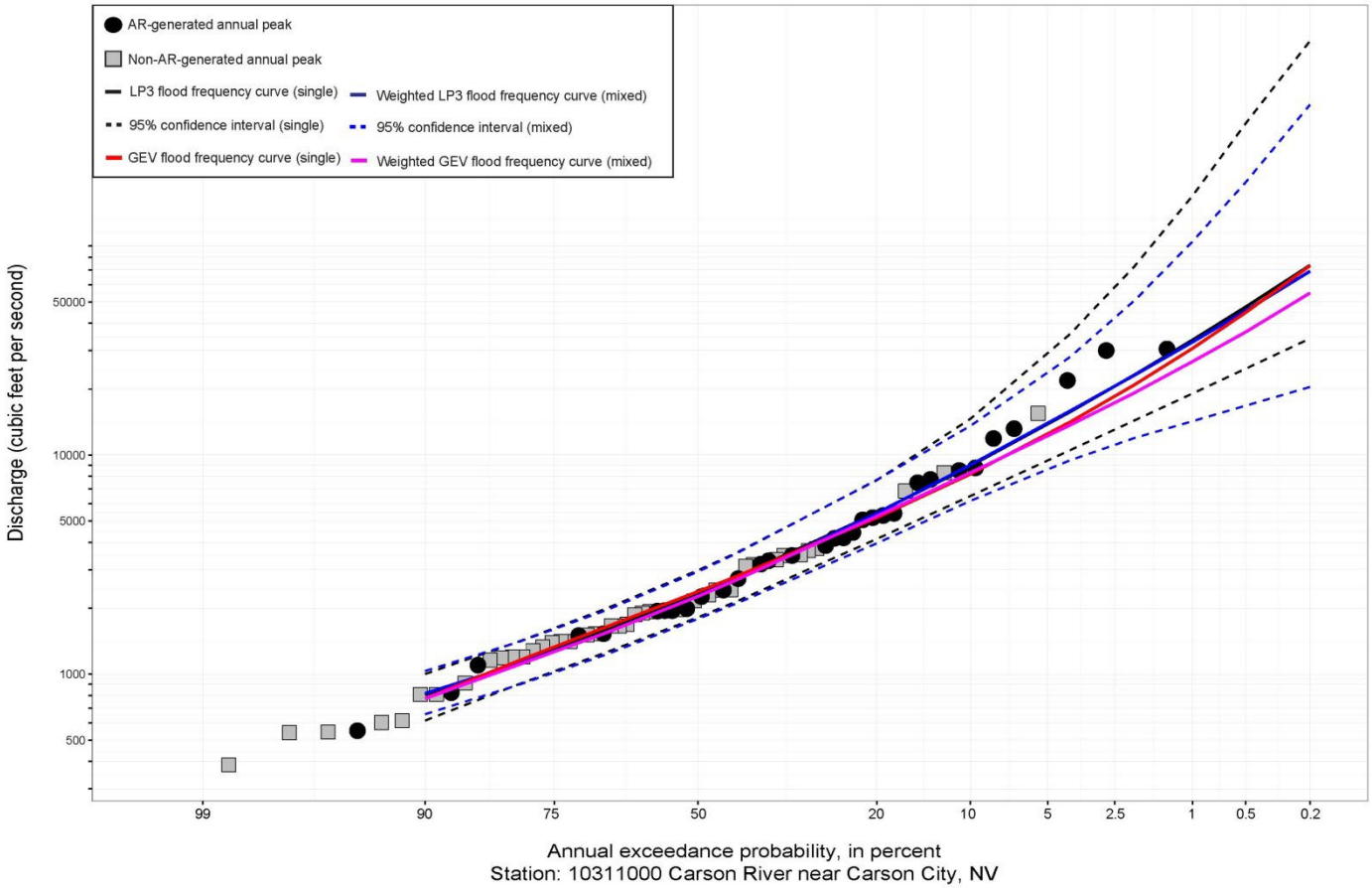

Figure S.58: Fitted frequency curves as described in figure caption S.44.

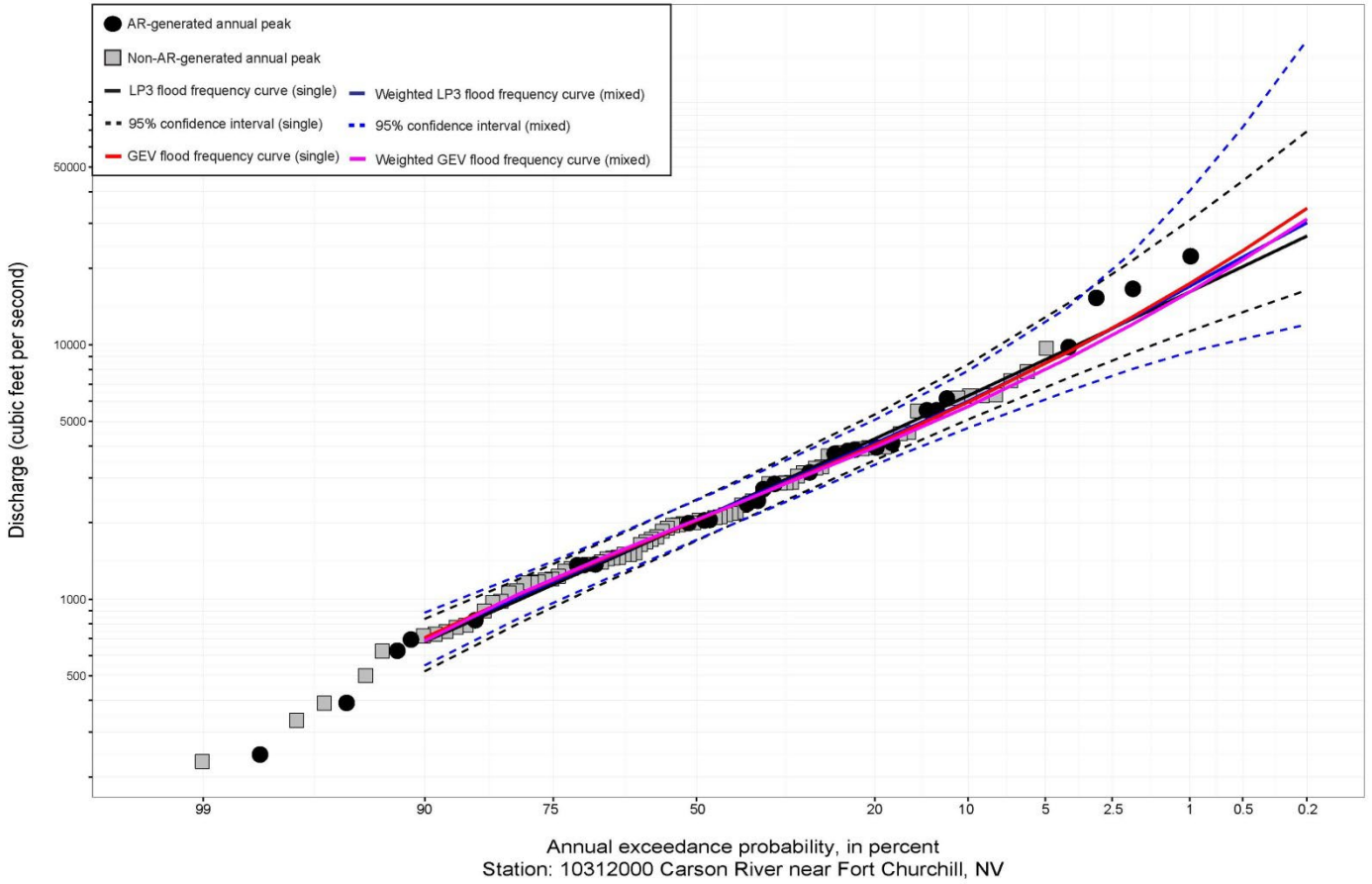

Figure S.59: Fitted frequency curves as described in figure caption S.44.

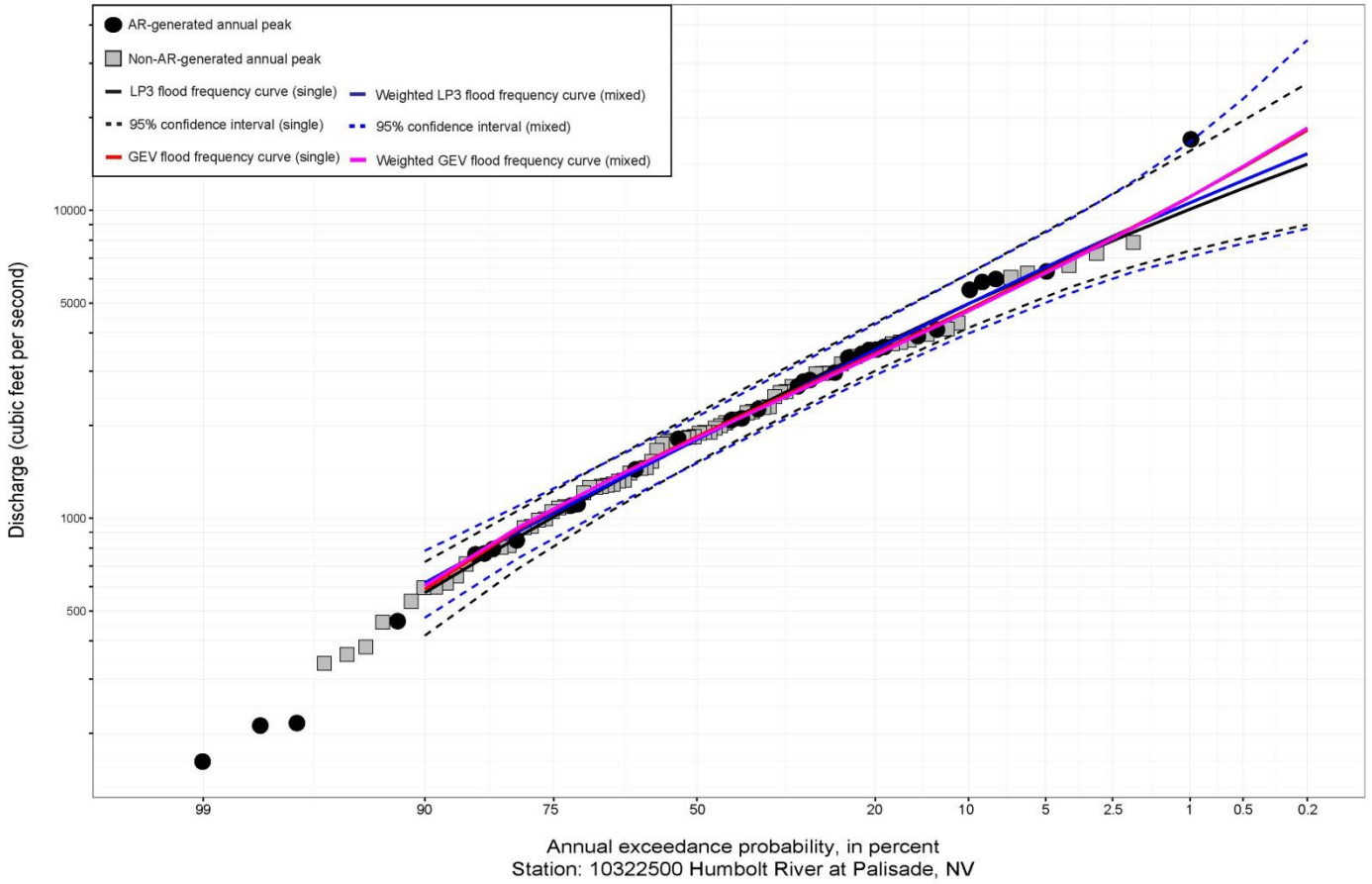

Figure S.60: Fitted frequency curves as described in figure caption S.44.

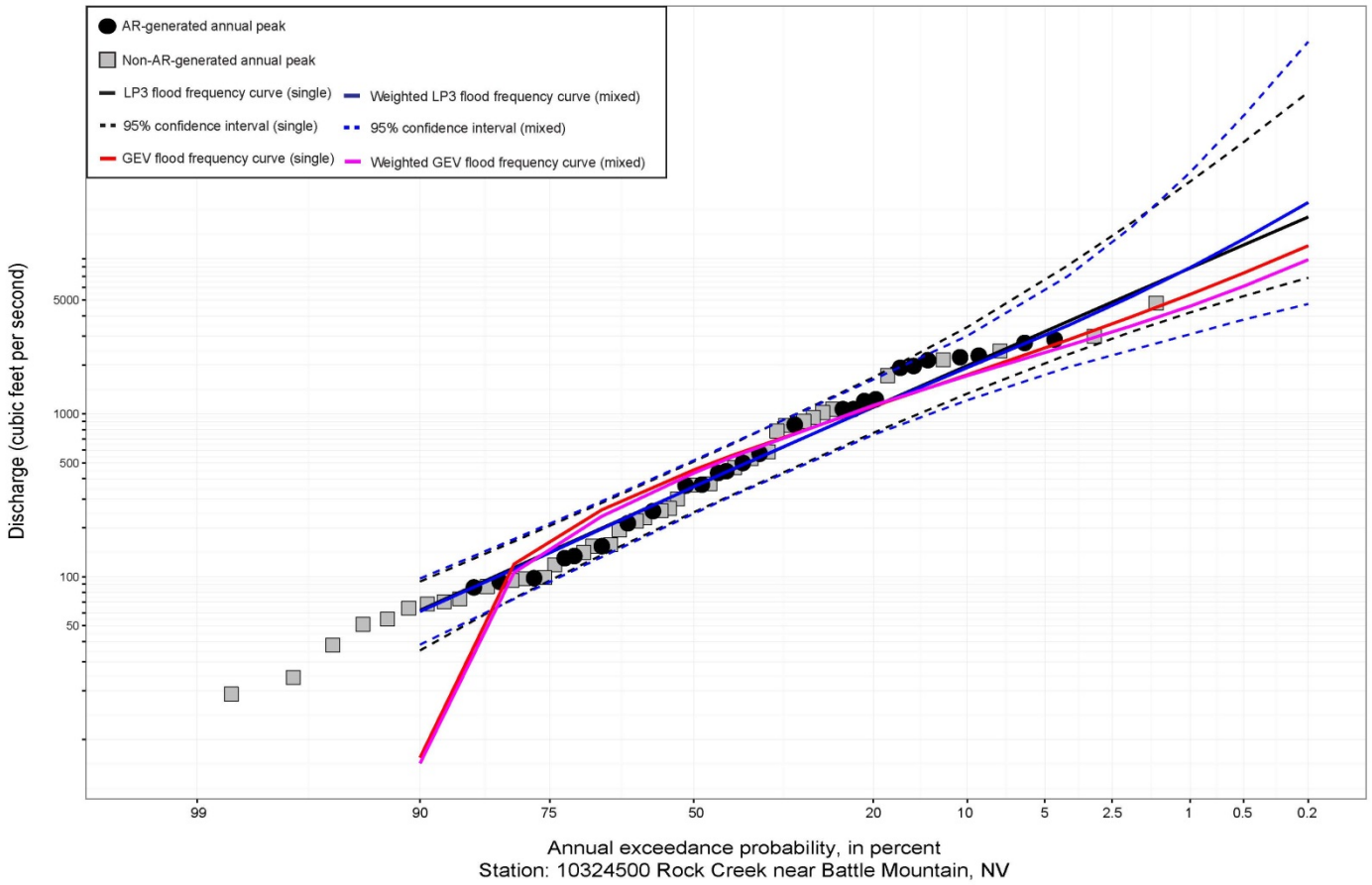

Figure S.61: Fitted frequency curves as described in figure caption S.44.

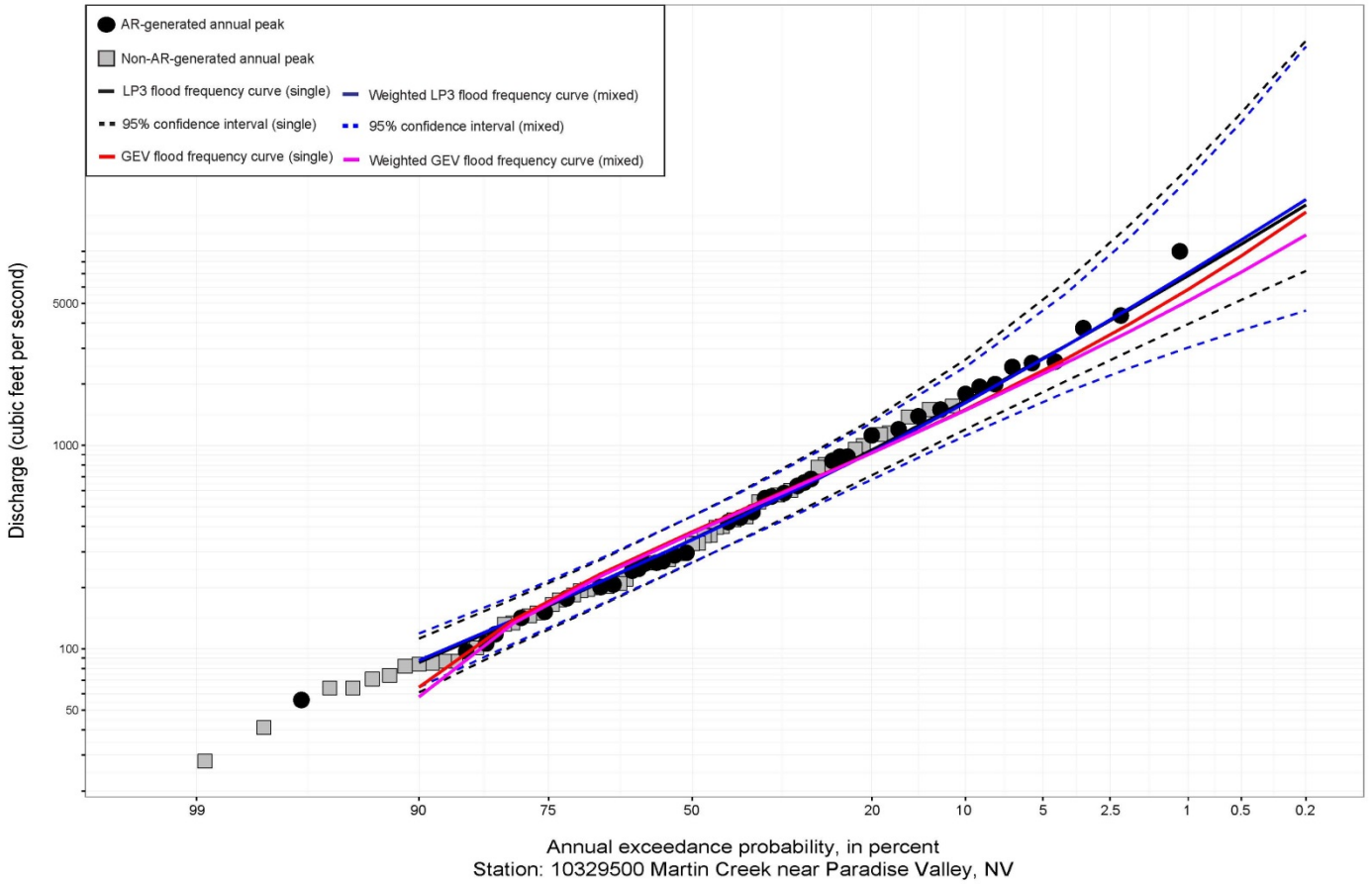

Figure S.62: Fitted frequency curves as described in figure caption S.44.

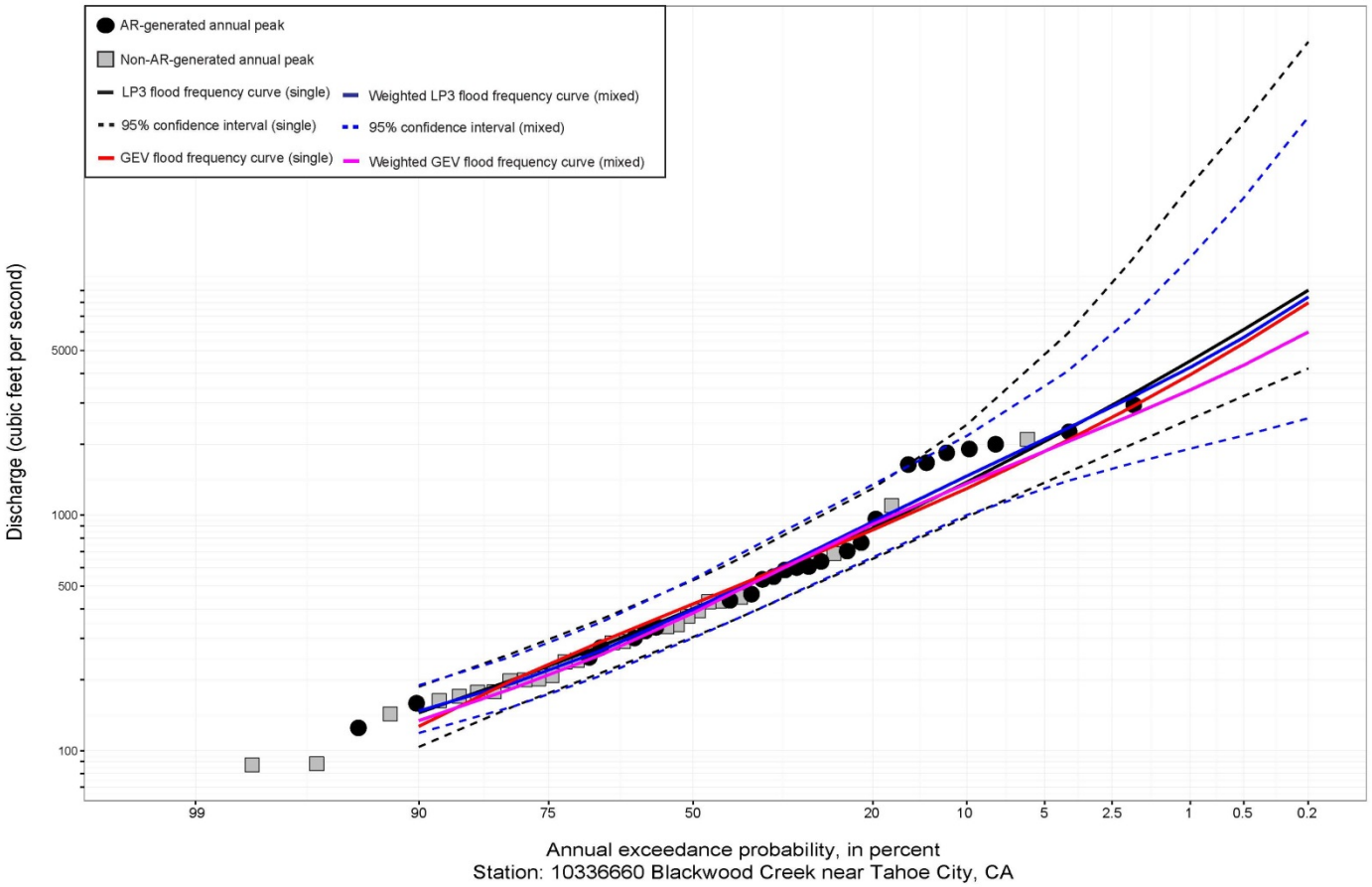

Figure S.63: Fitted frequency curves as described in figure caption S.44.

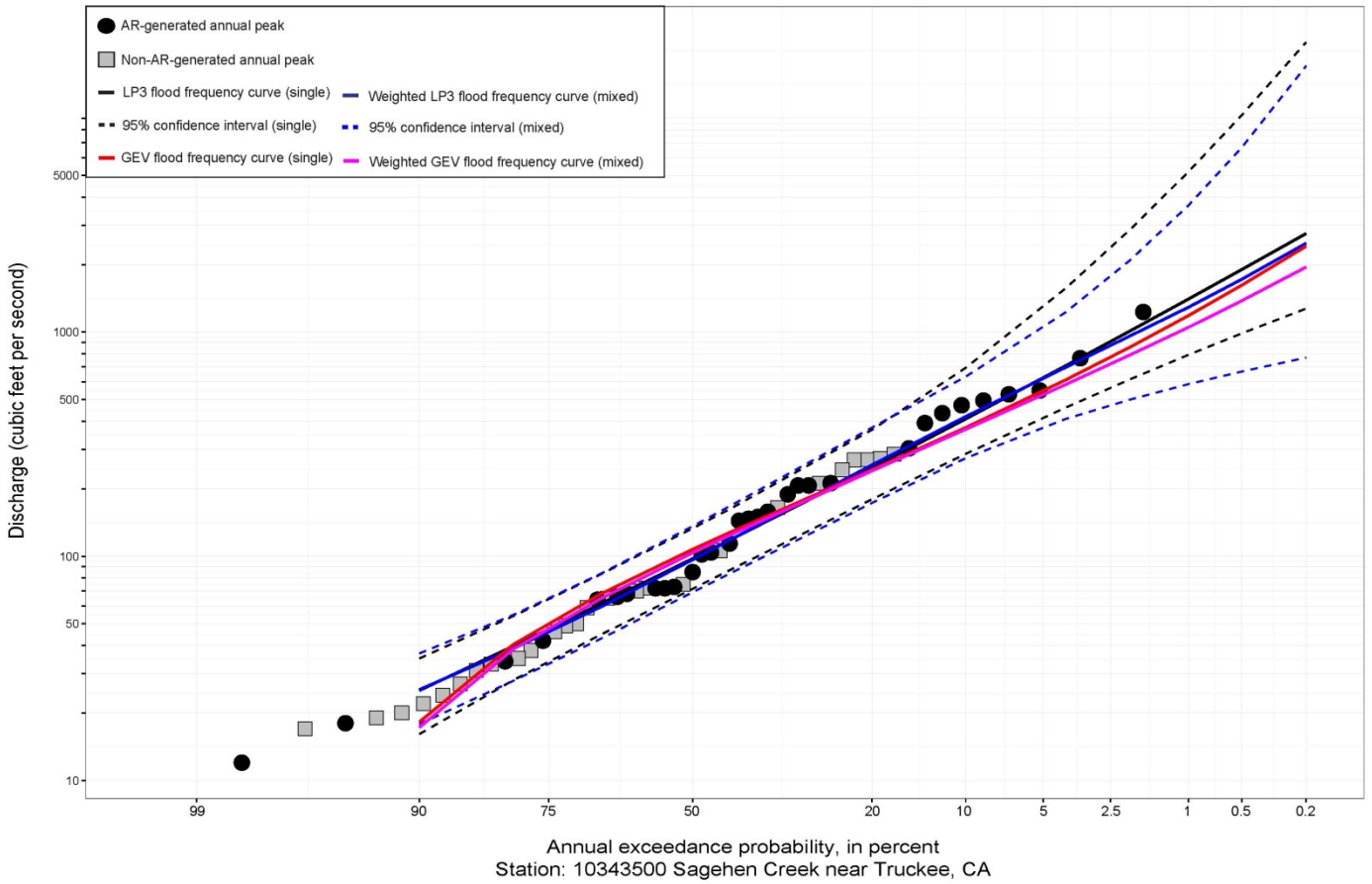

Figure S.64: Fitted frequency curves as described in figure caption S.44.

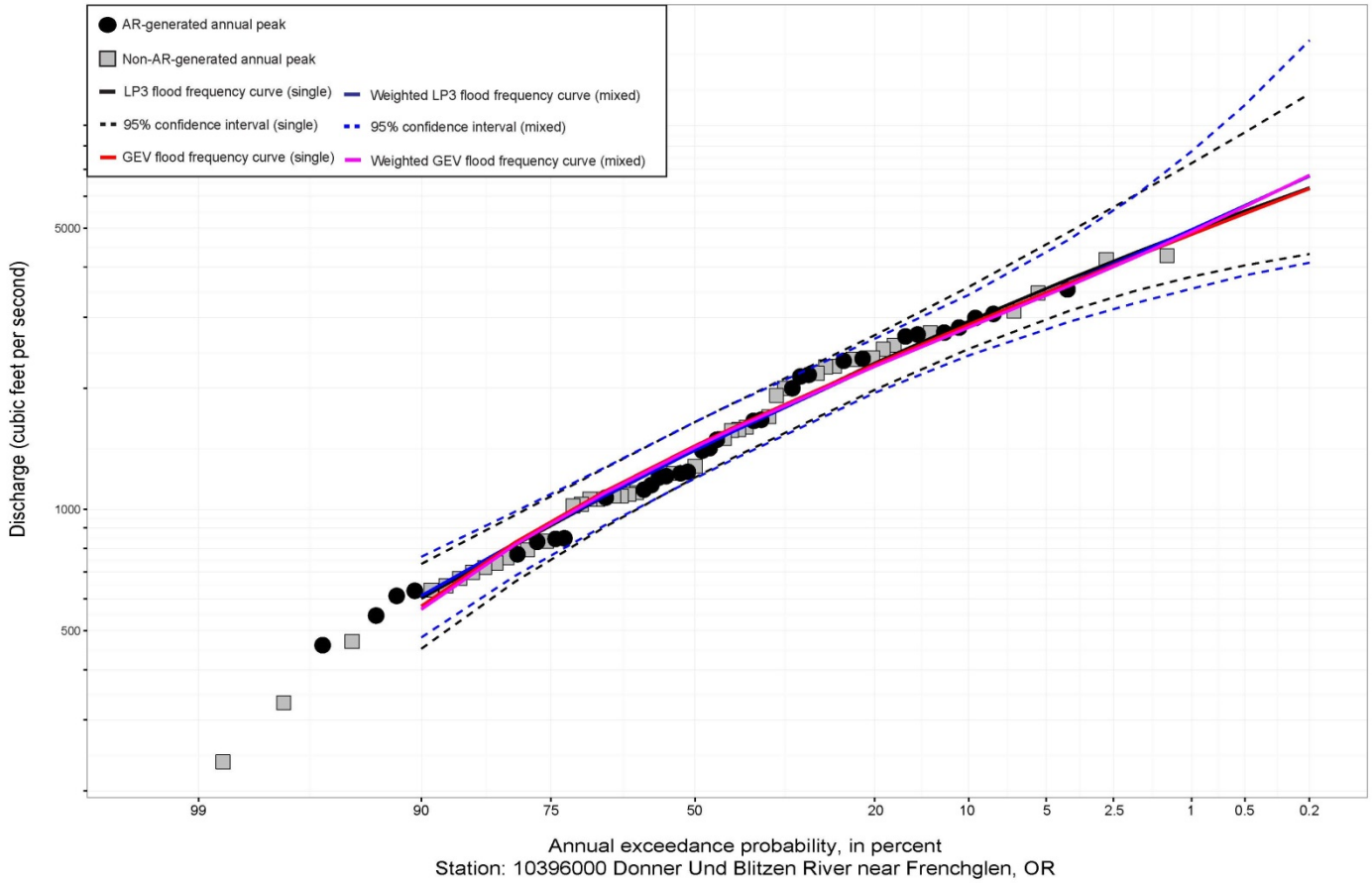

Figure S.65: Fitted frequency curves as described in figure caption S.44.

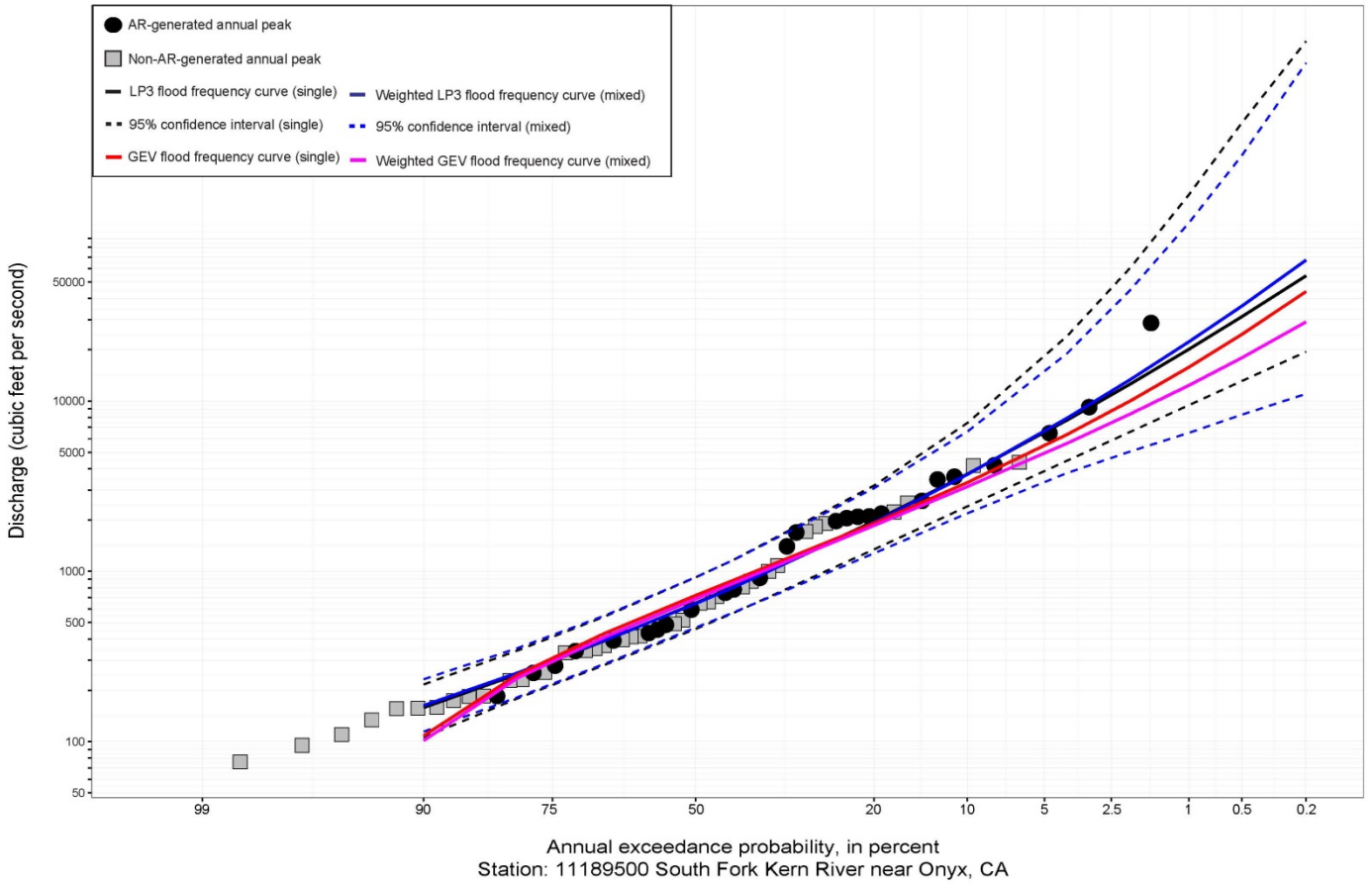

Figure S.66: Fitted frequency curves as described in figure caption S.44.

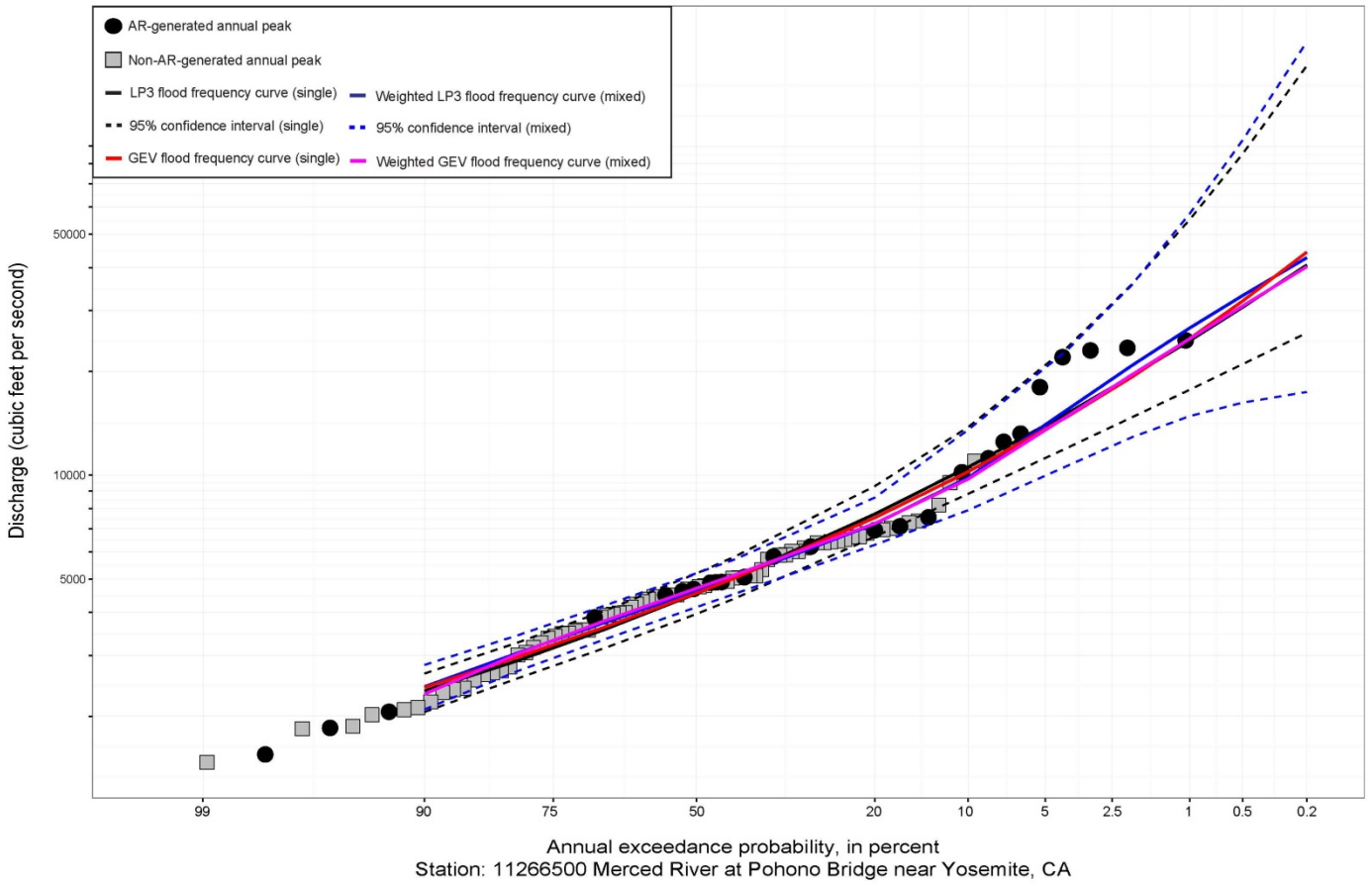

Figure S.67: Fitted frequency curves as described in figure caption S.44.

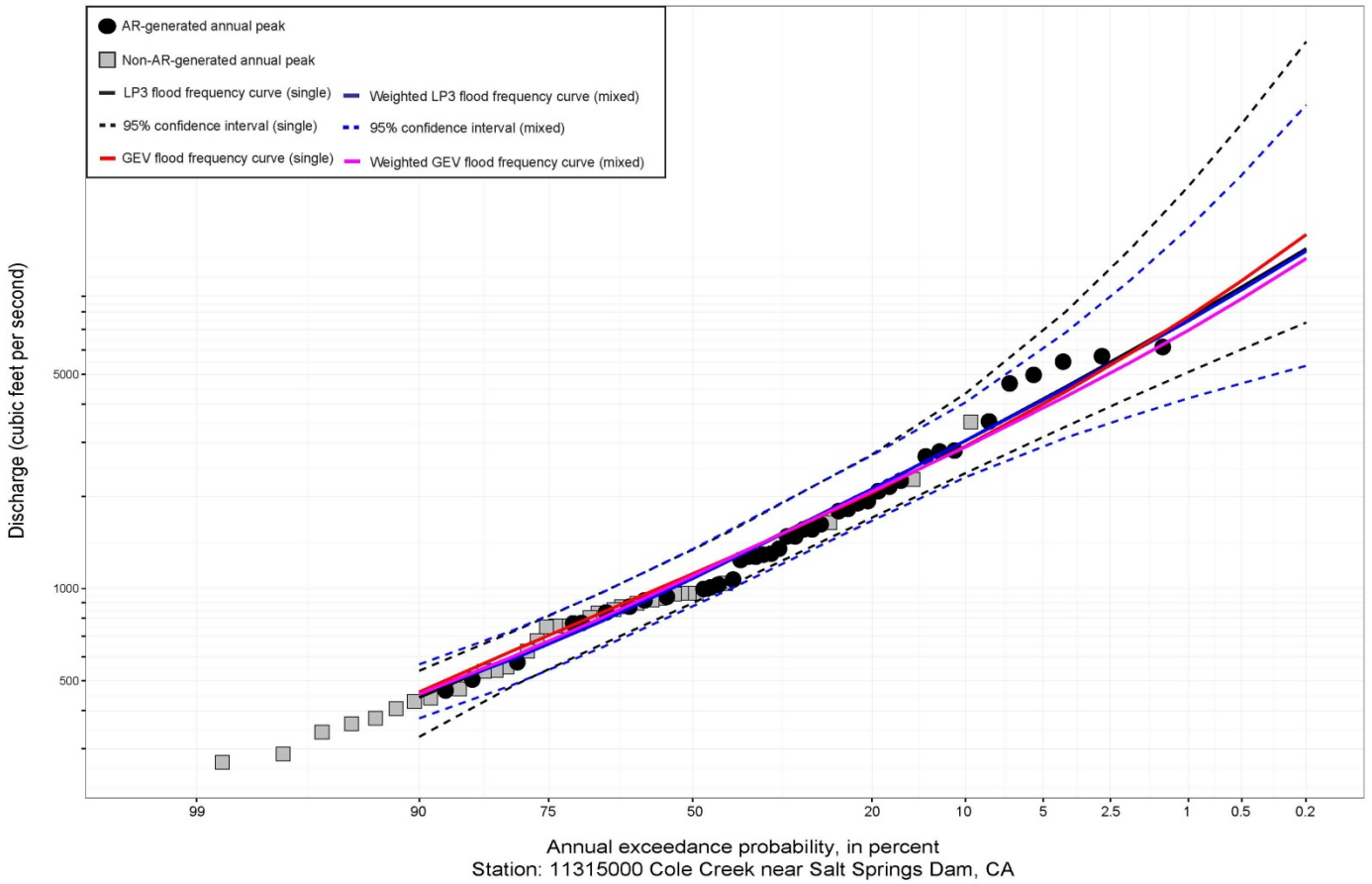

Figure S.68: Fitted frequency curves as described in figure caption S.44.

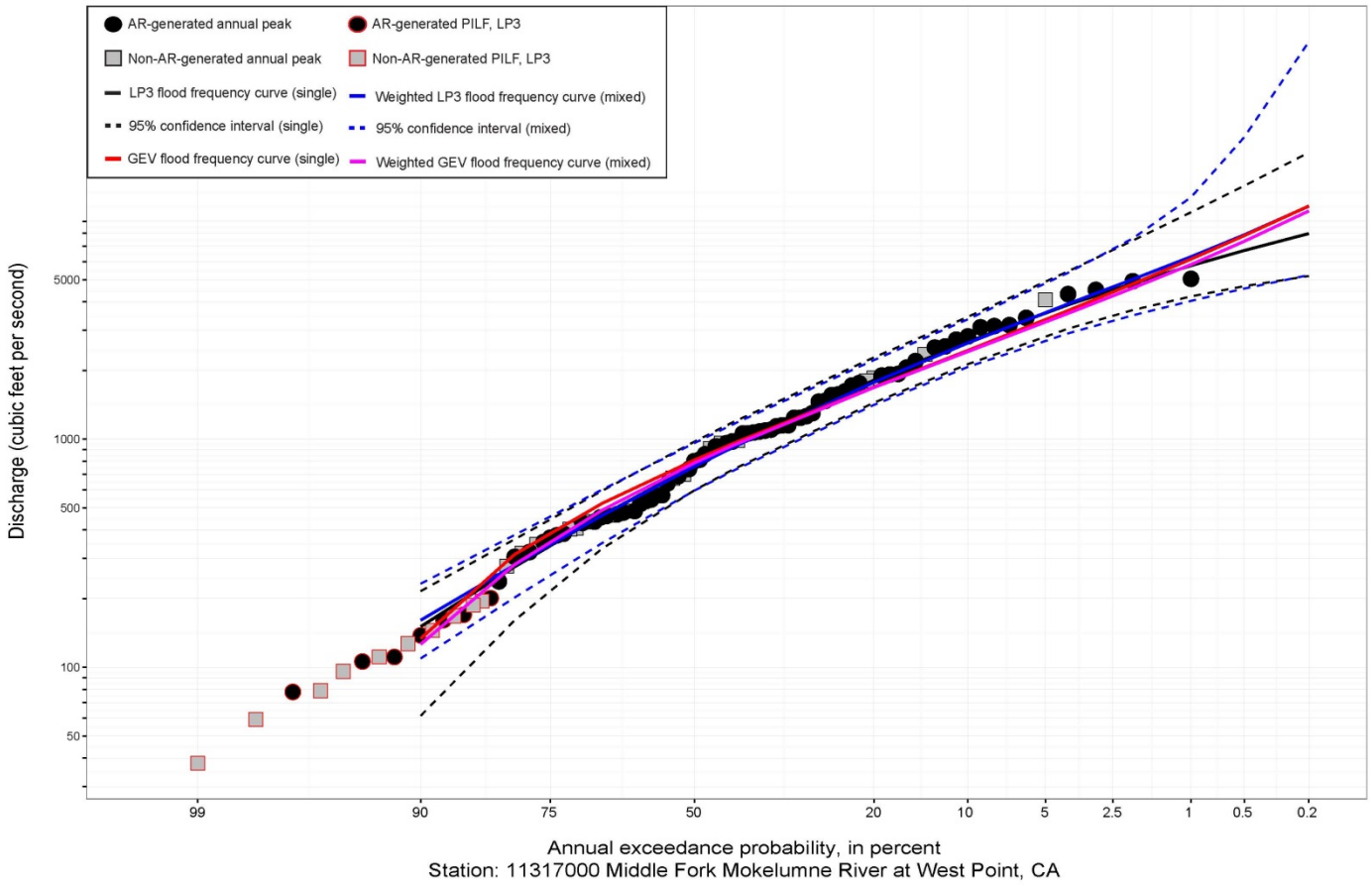

Figure S.69: Fitted frequency curves as described in figure caption S.44.

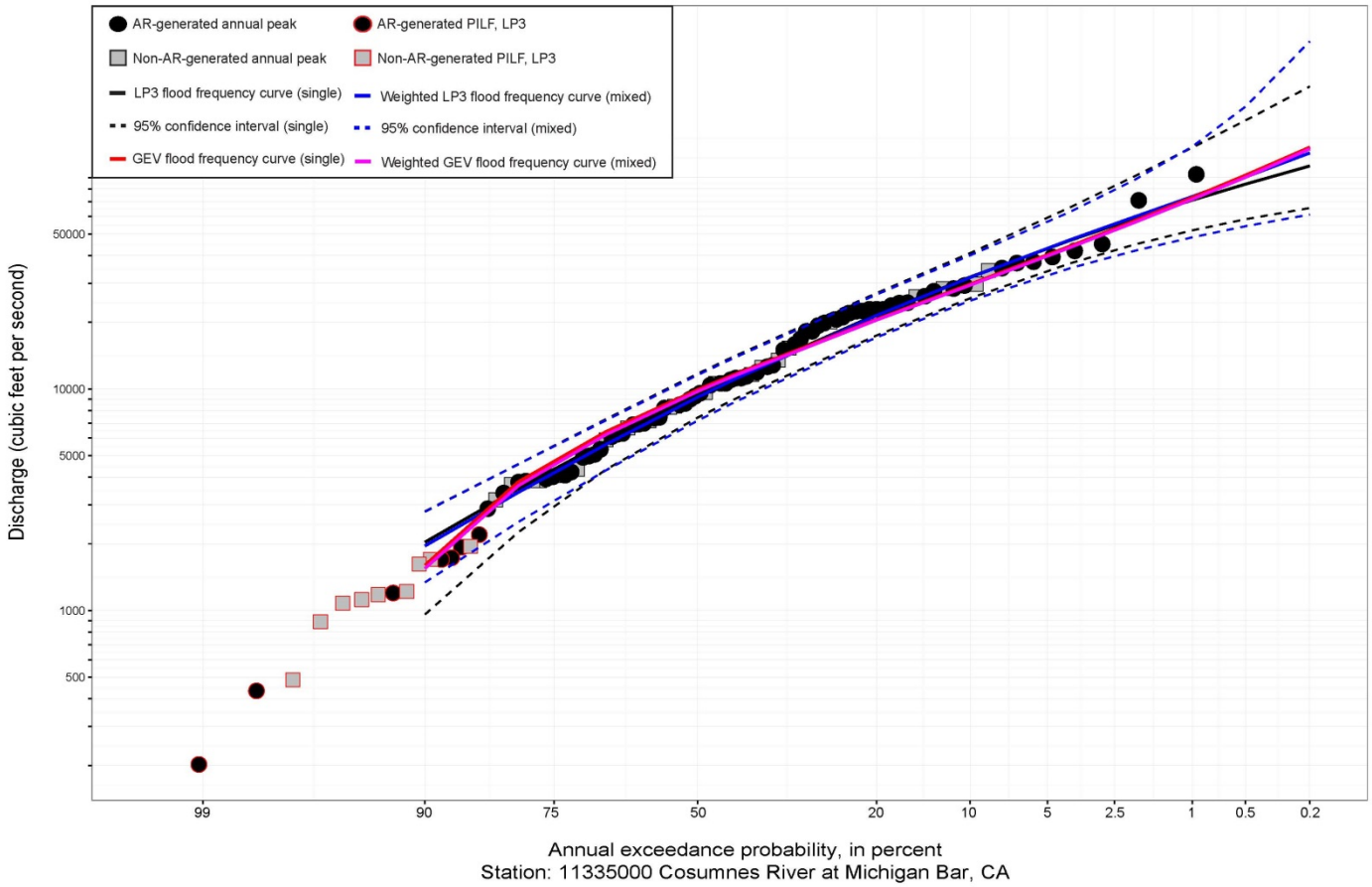

Figure S.70: Fitted frequency curves as described in figure caption S.44.

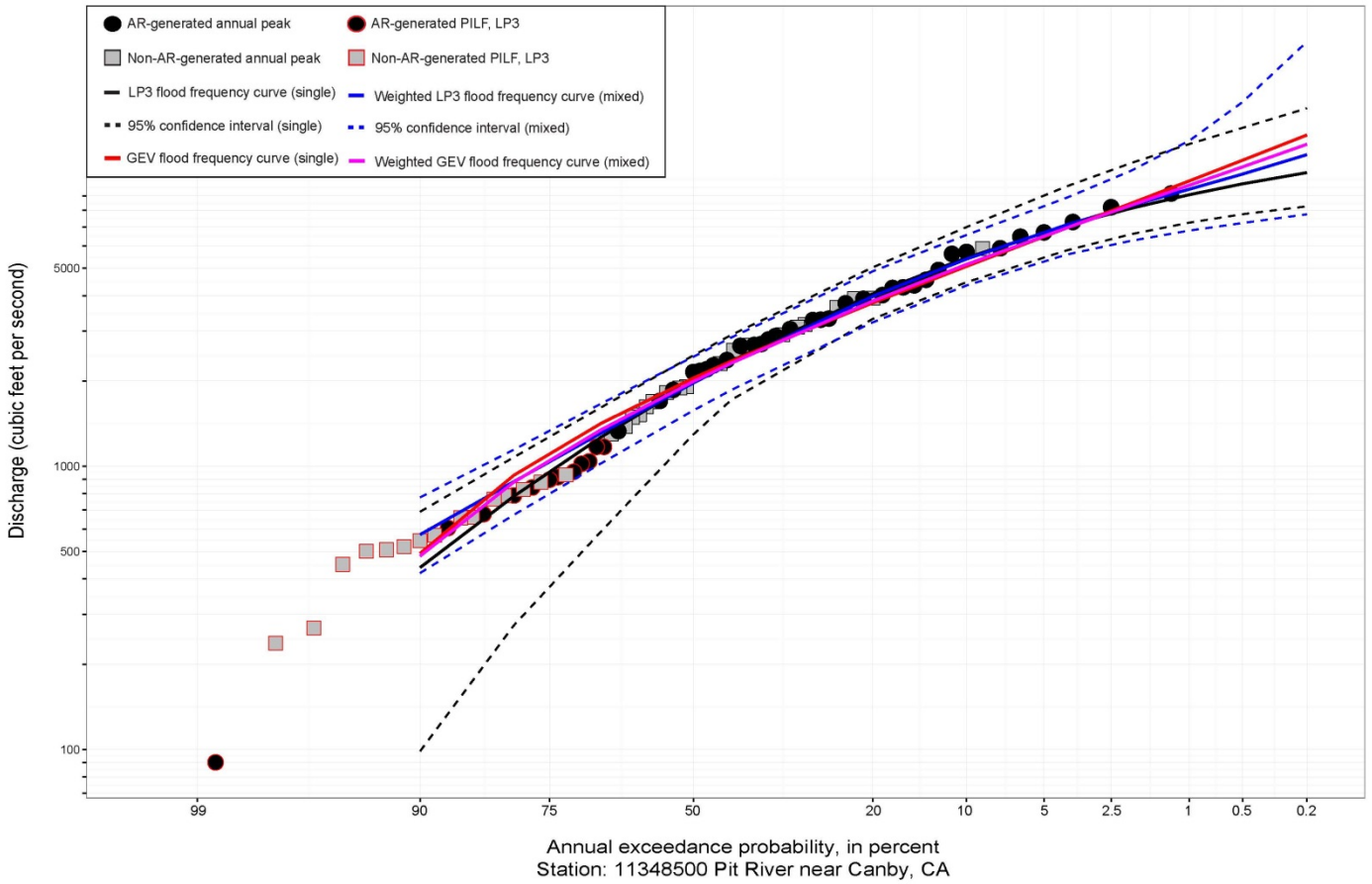

Figure S.71: Fitted frequency curves as described in figure caption S.44.

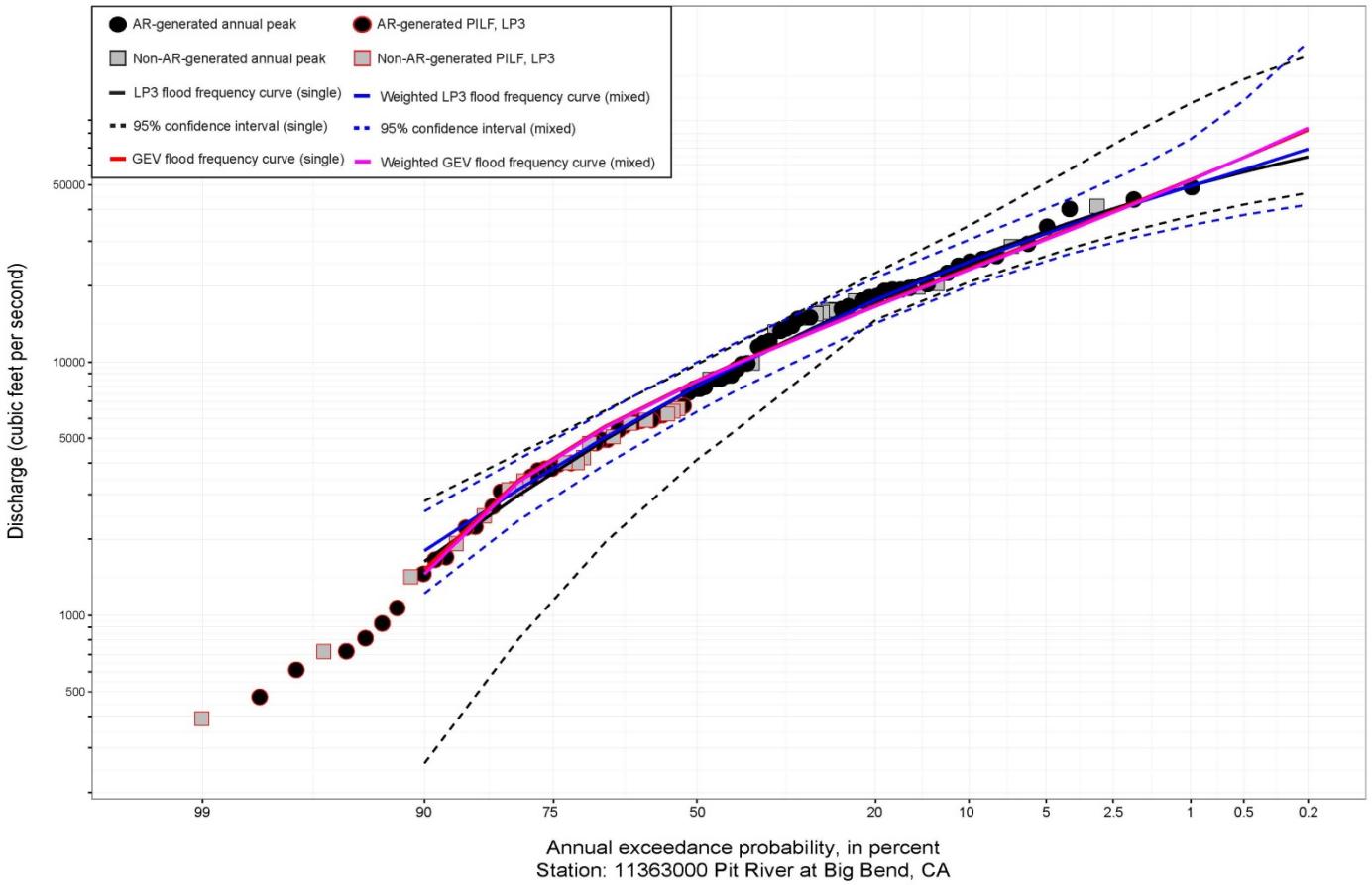

Figure S.72: Fitted frequency curves as described in figure caption S.44.

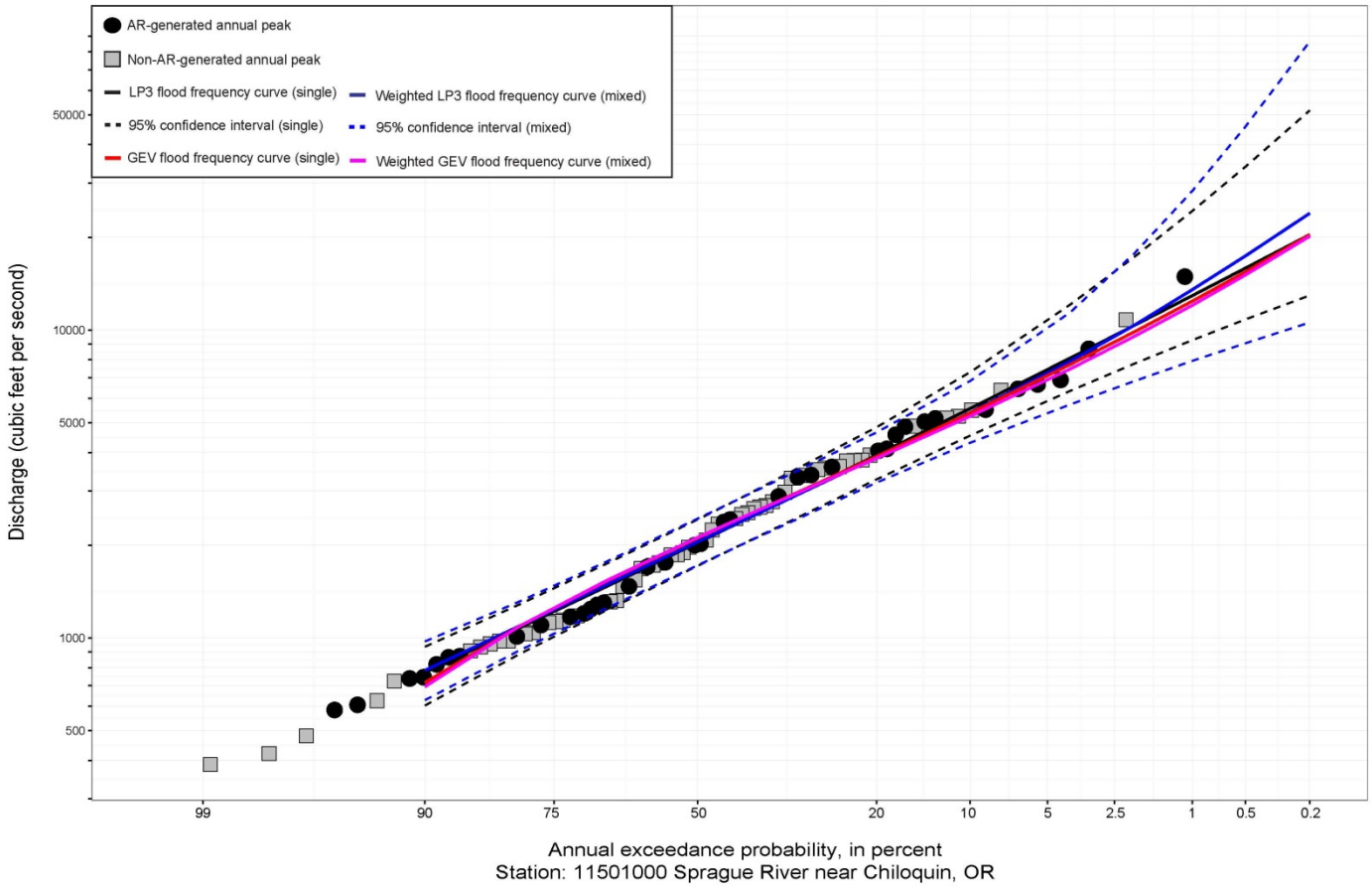

Figure S.73: Fitted frequency curves as described in figure caption S.44.

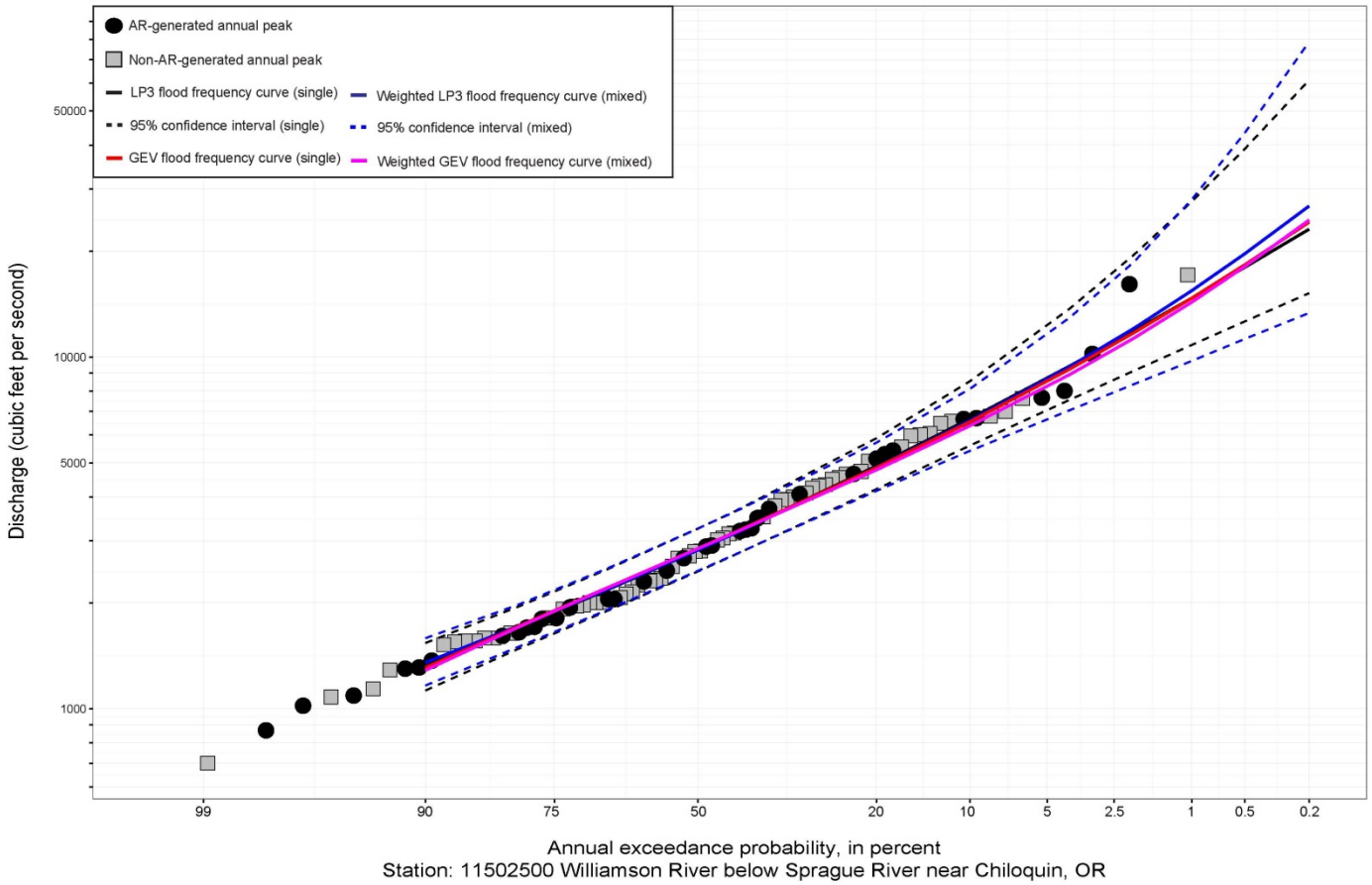

Figure S.74: Fitted frequency curves as described in figure caption S.44.

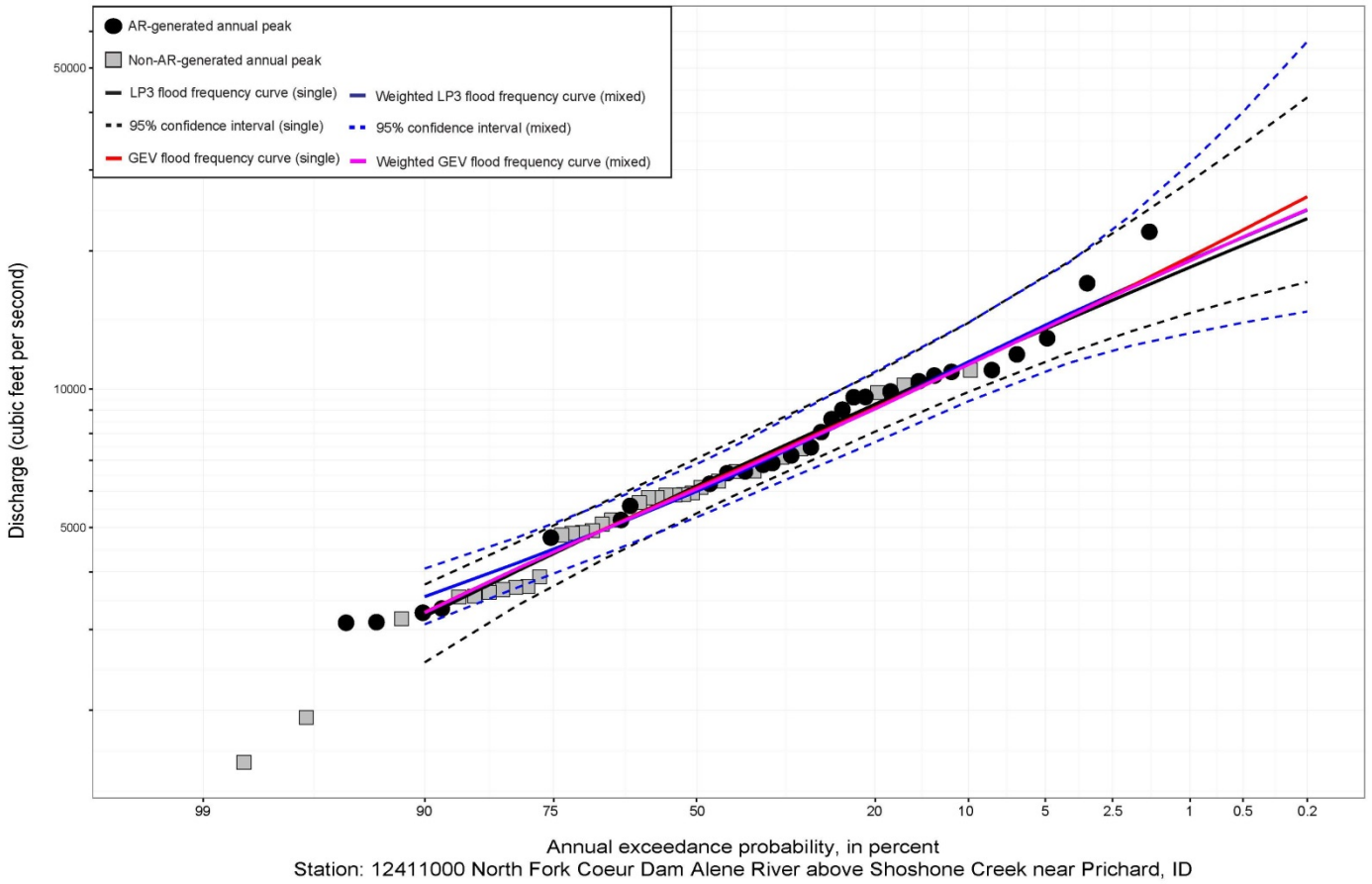

Figure S.75: Fitted frequency curves as described in figure caption S.44.

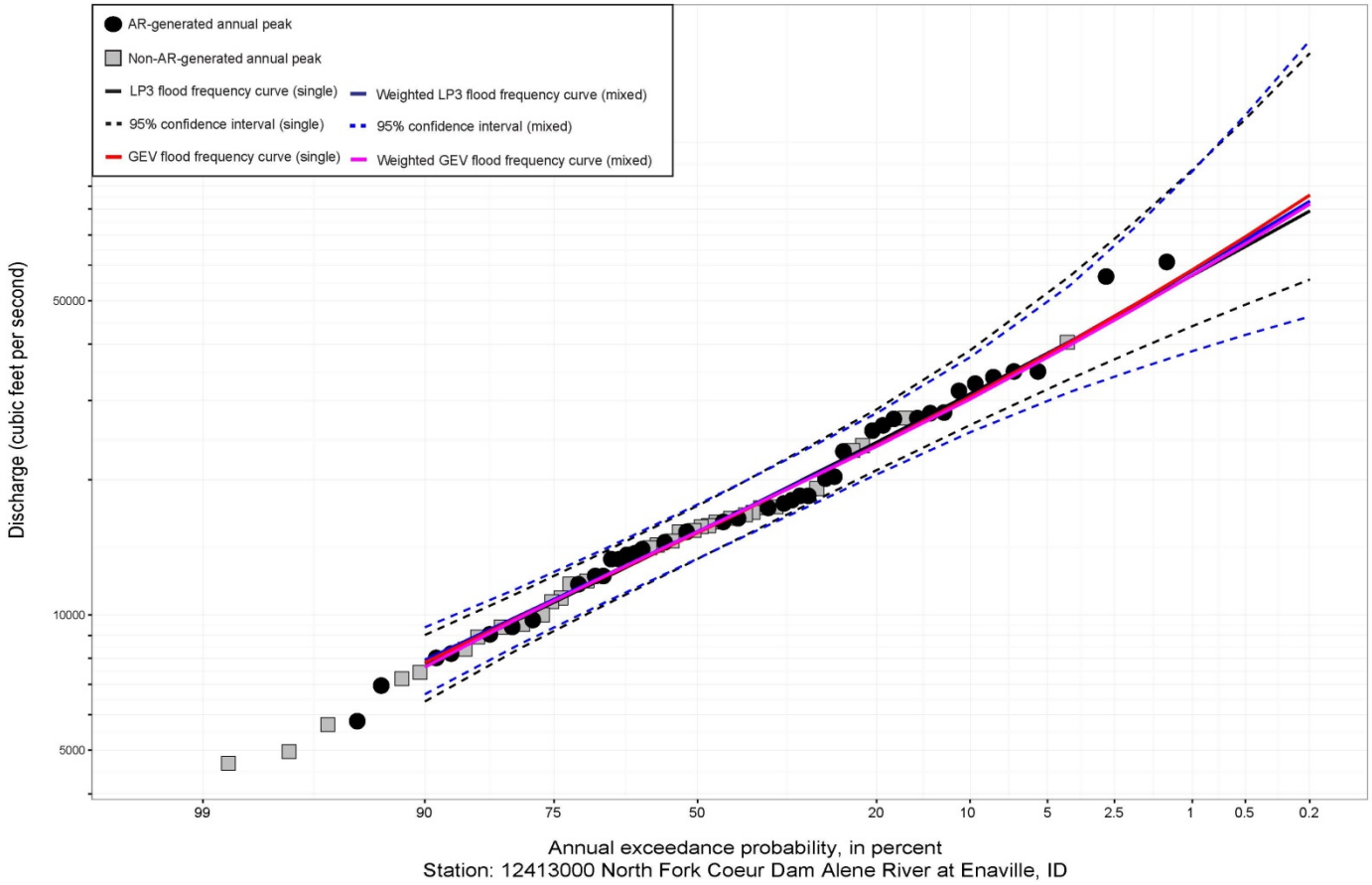

Figure S.76: Fitted frequency curves as described in figure caption S.44.

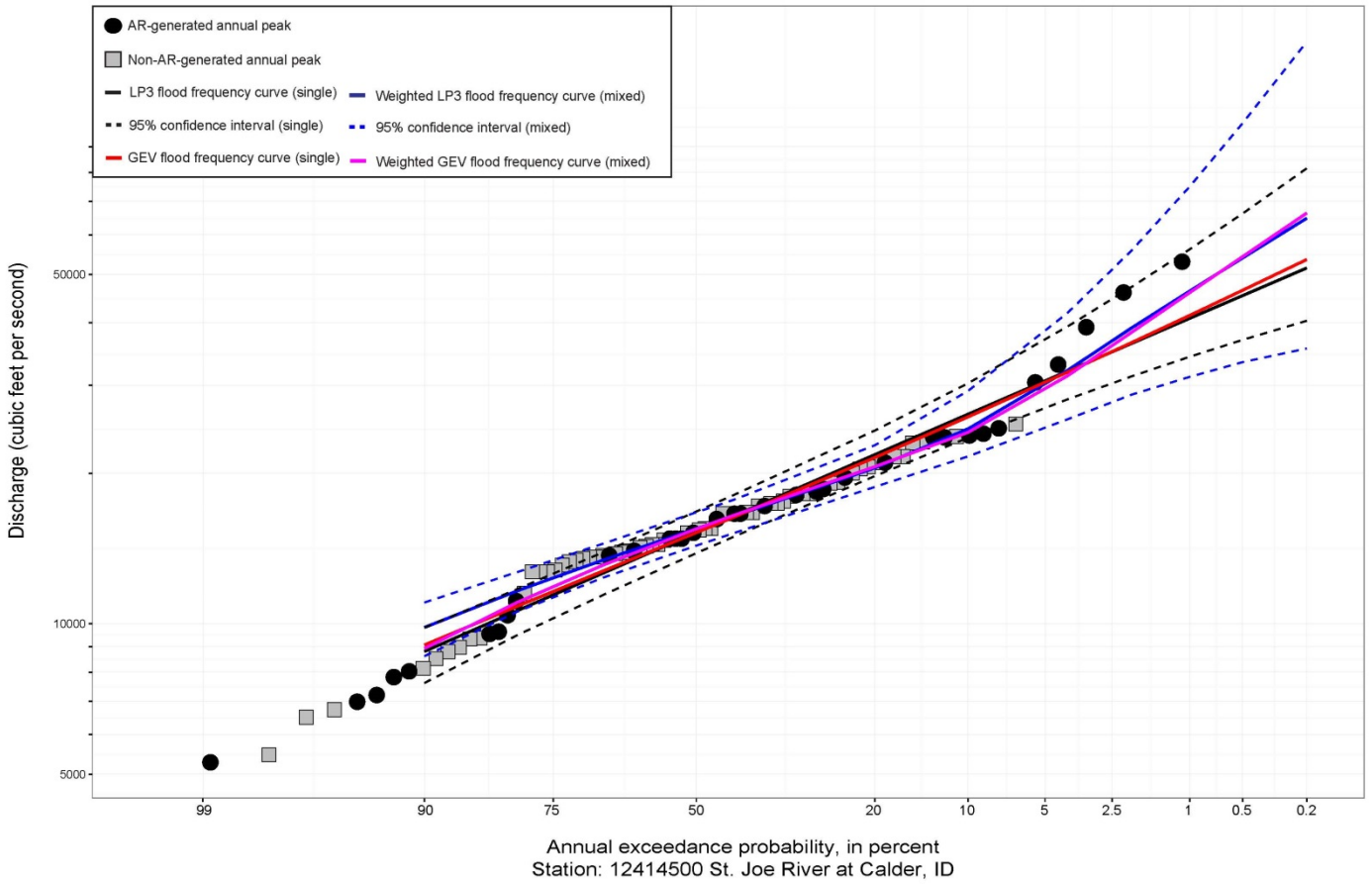

Figure S.77: Fitted frequency curves as described in figure caption S.44.

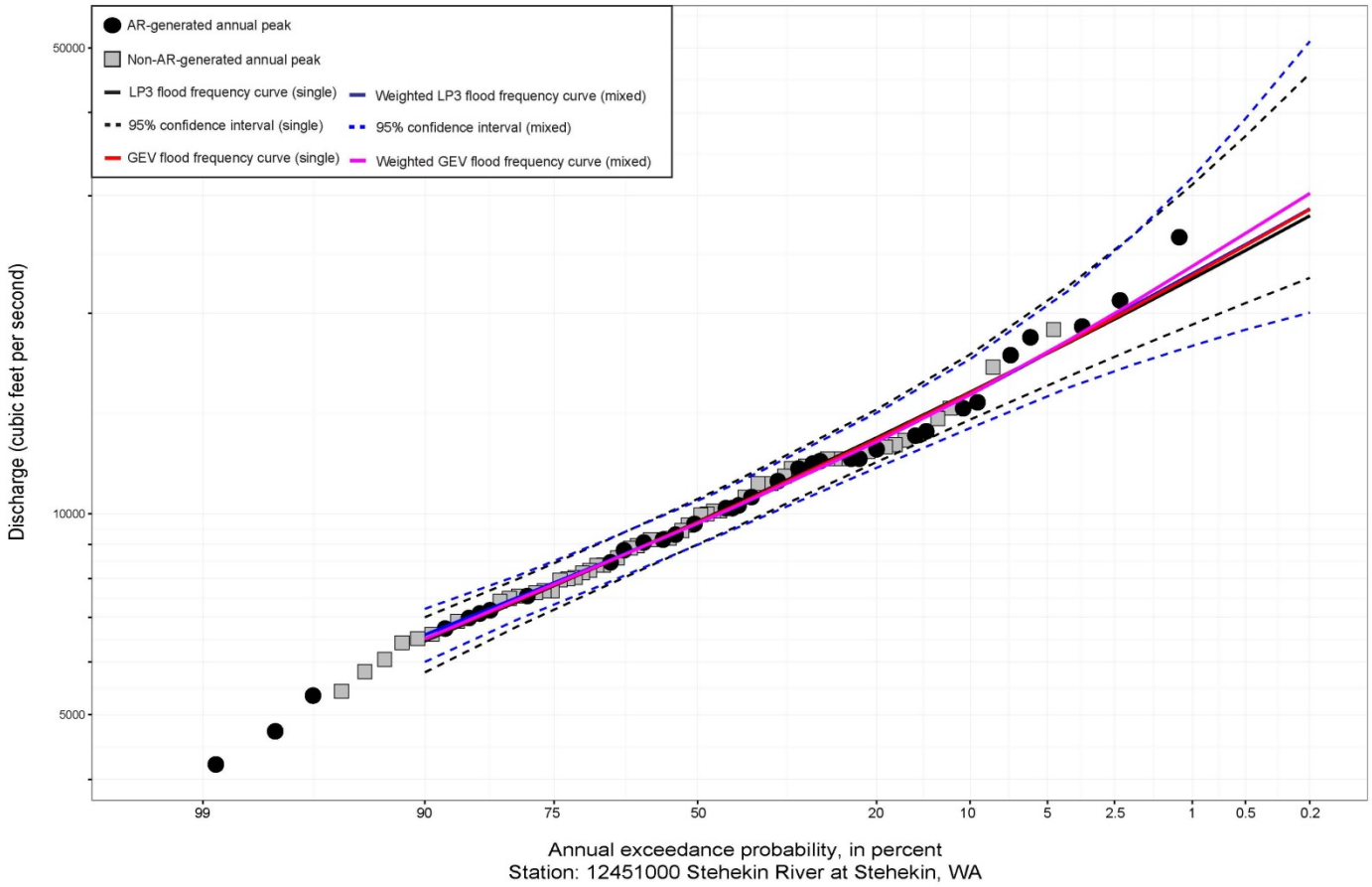

Figure S.78: Fitted frequency curves as described in figure caption S.44.

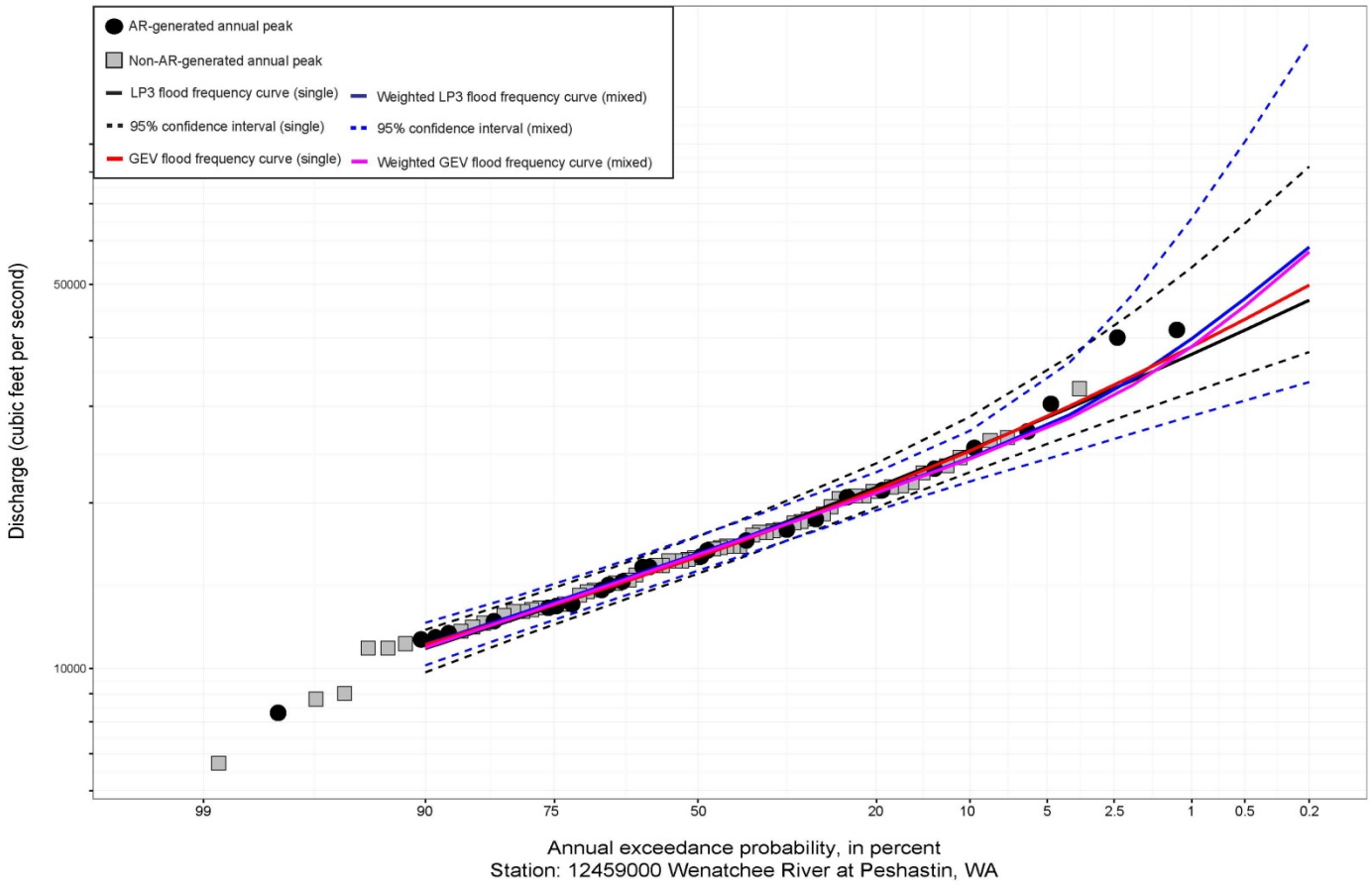

Figure S.79: Fitted frequency curves as described in figure caption S.44.

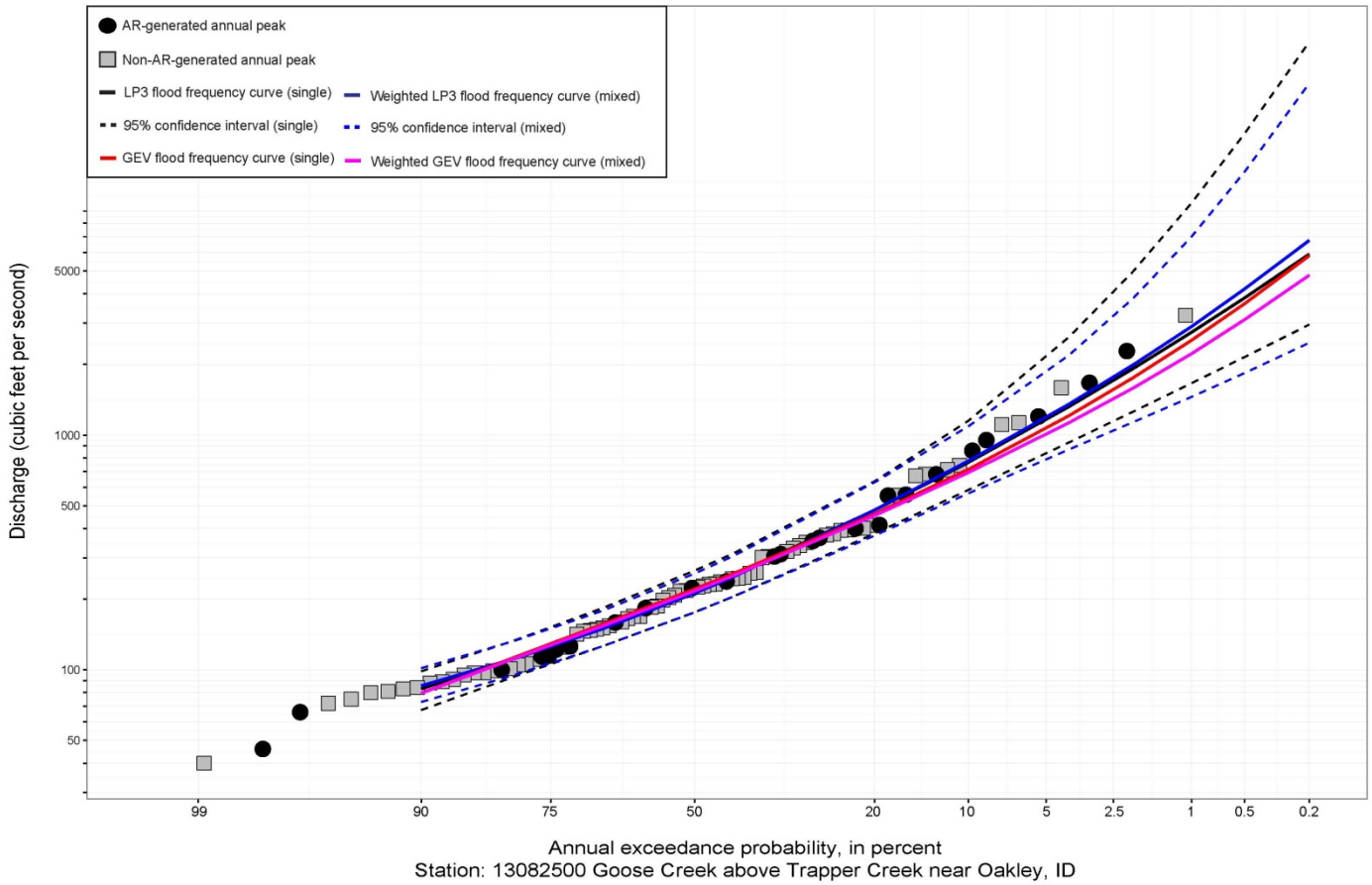

Figure S.80: Fitted frequency curves as described in figure caption S.44.

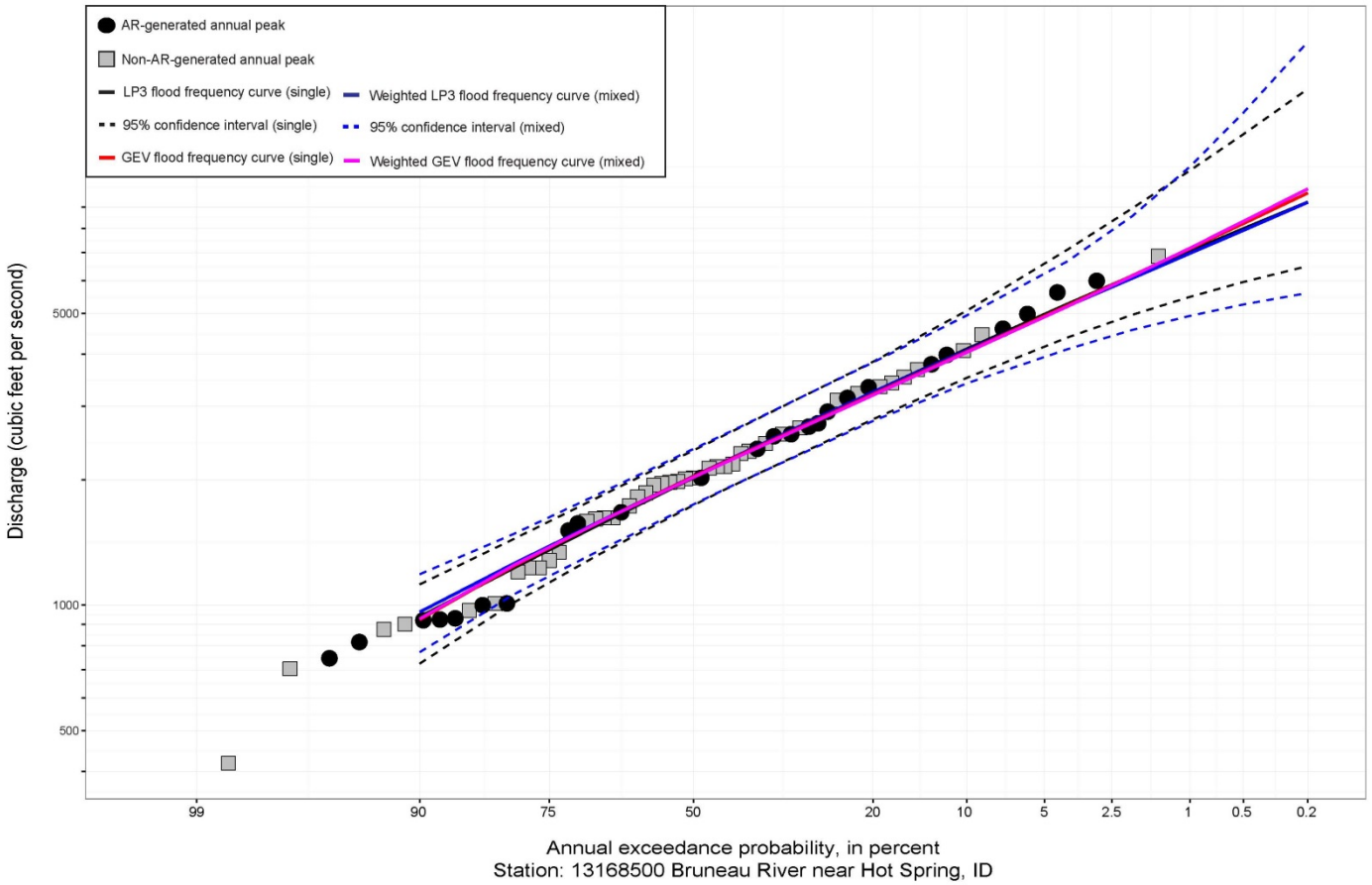

Figure S.81: Fitted frequency curves as described in figure caption S.44.

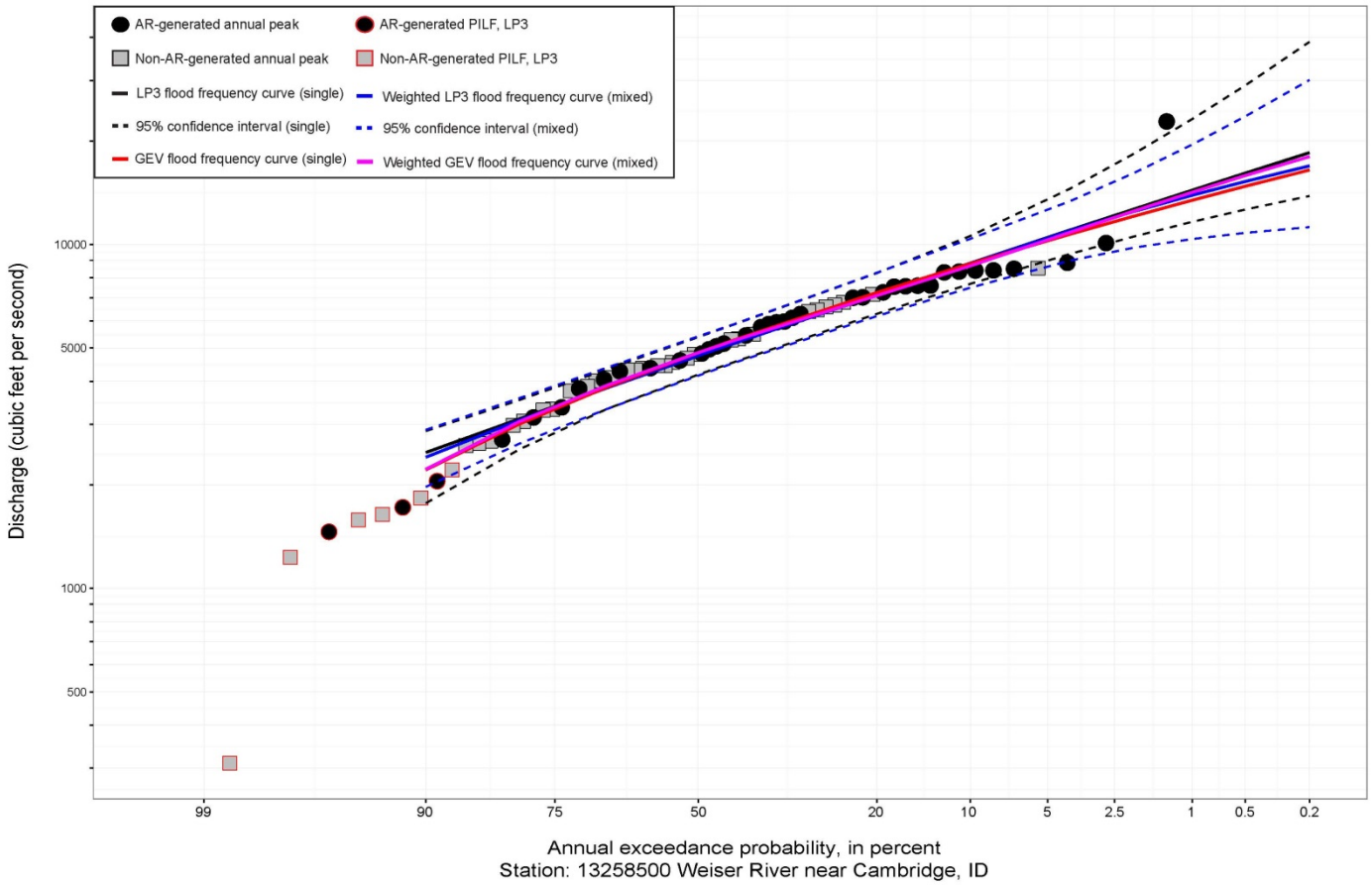

Figure S.82: Fitted frequency curves as described in figure caption S.44.

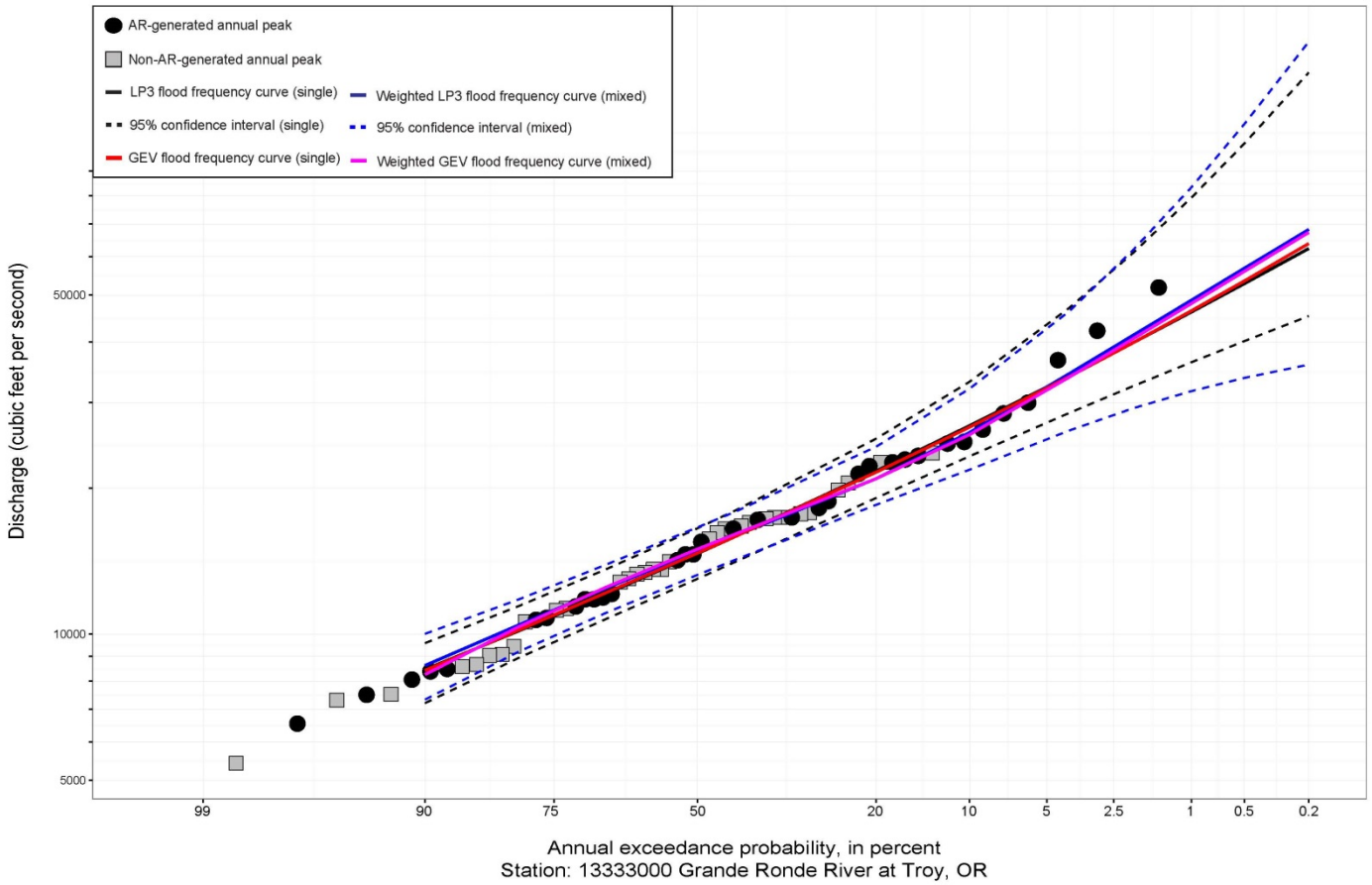

Figure S.83: Fitted frequency curves as described in figure caption S.44.

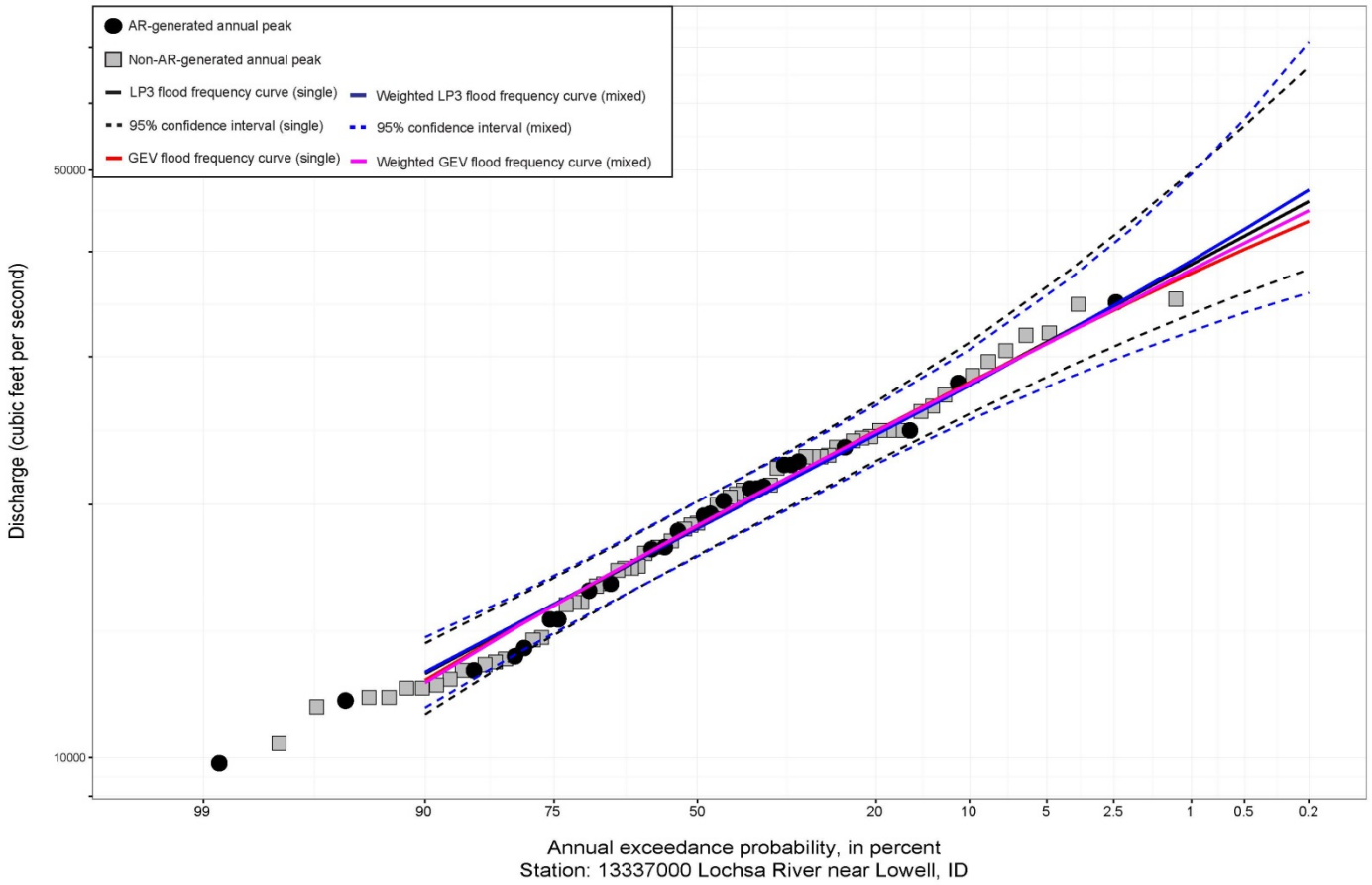

Figure S.84: Fitted frequency curves as described in figure caption S.44.

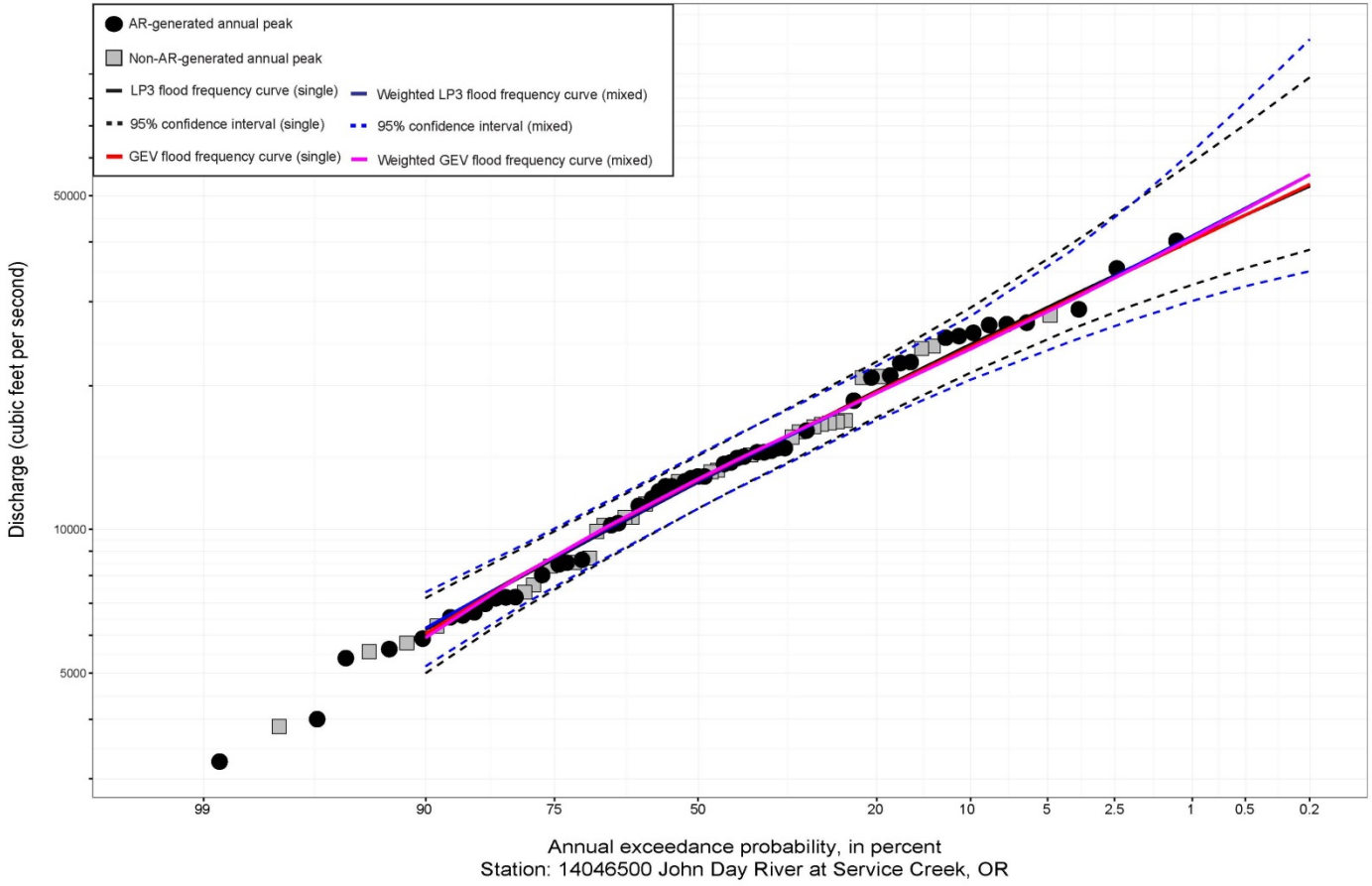

Figure S.85: Fitted frequency curves as described in figure caption S.44.

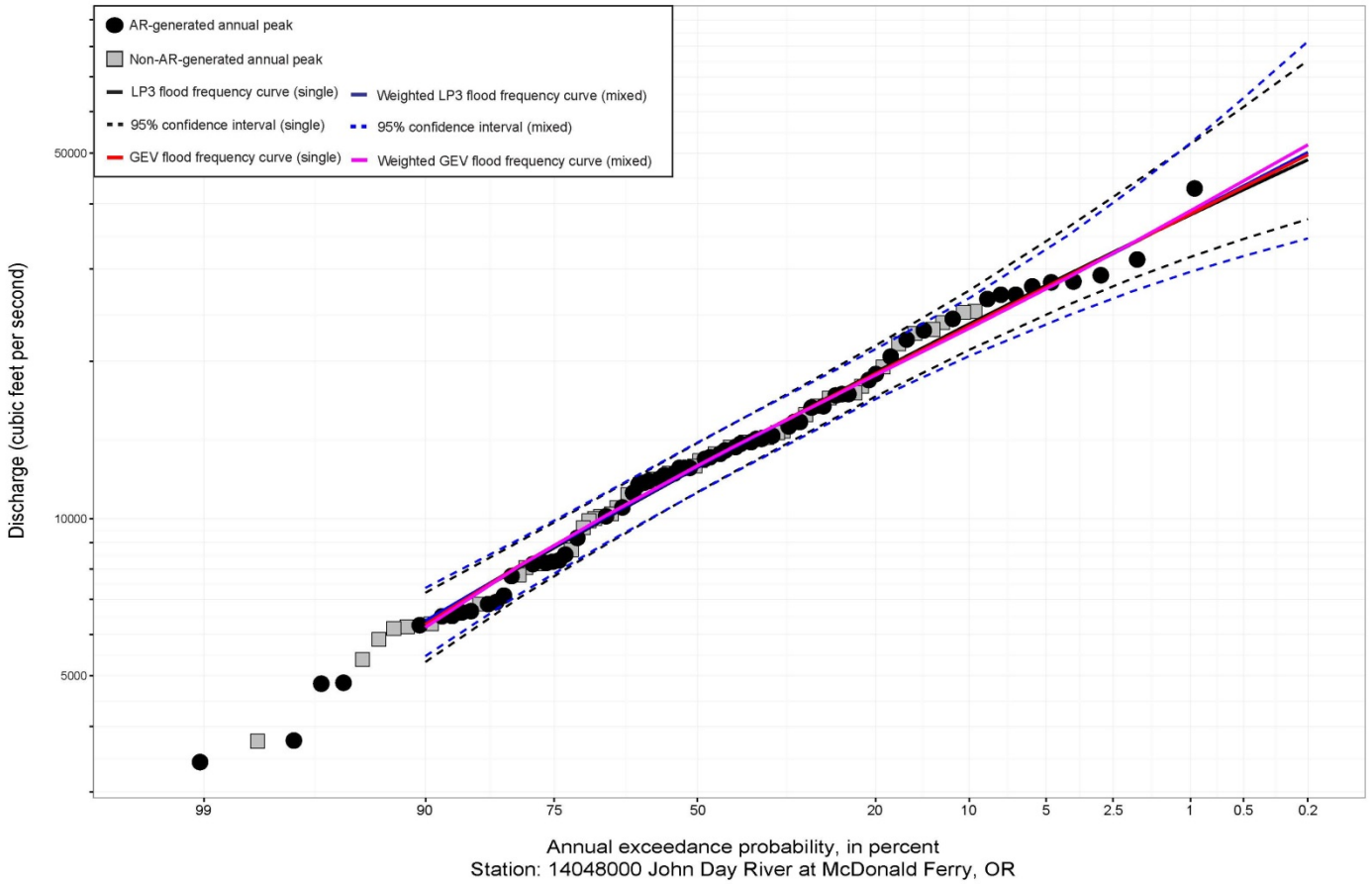

Figure S.86: Fitted frequency curves as described in figure caption S.44.