2-2022

Evaluation of GEV Over LP3 When Predicting Return Period Annual Exceedance For Santa Ana, San Gabriel and Urbanized Regions in California

Maria Beatriz de Paula Macedo
Loyola Marymount University, mdepaul2@lion.lmu.edu

Evaluation of GEV
Over LP3 When Predicting Return Period Annual
Exceedance for Santa Ana, San Gabriel and
Urbanized Regions in California

by

Maria Beatriz De Paula Macedo

A thesis presented to the

Faculty of the Department of
Civil Engineering and Environmental Science
Loyola Marymount University

In partial fulfillment of the
Requirements for the Degree
Master of Science in Engineering in Civil Engineering

August 14, 2023

I am submitting herewith a thesis written by Maria Beatriz De Paula Macedo entitled "Evaluation Of GEV Over LP3 When Predicting Annual Exceedance For the Santa Ana, San Gabriel and Urbanized Regions in Southern California". I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science in Engineering in Civil Engineering.
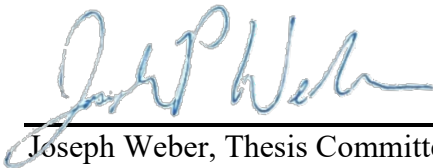
_____

Donald Kendall, Advisor

We have reviewed this thesis and recommend its acceptance:

_____

Jeremy Pal, Thesis Committee Member

_____

Joseph Weber, Thesis Committee Member

ABSTRACT

The objective of this present thesis was to determine whether GEV (Generalized Extreme Value) itself can be a more conservative distribution than LP3 (Log Pearson III) associated with other methods, such as the B17B weighting procedure with Single Grubbs-Beck (SGB) for low outliers, when determining the projected floods in a flood frequency analysis (FFA) for Santa Ana and San Gabriel regions and other urbanized stream gages present in California. In this work, USGS PeakFQ was utilized. From the results obtained, it was possible to state that GEV fitting results were directly affected by the length of the data. When the length of the record is short, it is not accurate to use a projection of 100-year return period, for example, to represent future projection. Comparing the LP3 and GEV CDFs, for the majority of the stream gages analyzed in this project, GEV proves to be the most conservative method, with smaller return periods.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABREVIATIONS

AEP – Annual Exceedance Probability

B17B – Bulletin 17B

CDF – Cumulative Density Function

CI – Confidence Interval

EMA – Expected Moments Algorithm

FFA – Flood Frequency Analysis

GEV – Generalized Extreme Value Distribution

IACWD – Interagency Committee on Water Data

LP3 – Log-Pearson Type 3

MGB – Multiple Grubbs-Beck

MLE – Maximum Likelihood Estimates

PDF – Probability Density Function

SGB – Single Grubbs-Beck

USGS – United States Geological Survey

WY – Water Years

UNITS

cfs – Cubic feet per second

k – Shape Parameter

σ – Scale Parameter

μ – Location Parameter

$\bar{y}$ – Mean Value

K – Frequency Factor

Gs – Coefficient of Skewness

S_y – Standard Deviation

T – Return Period

P – Percent Annual Exceedance (1/T)

1. INTRODUCTION

Over the years, different studies have been developed and published in order to refine the way of predicting flood's magnitude and frequency. Due to changes in climate, land use and increasing urbanization, extreme flood events tend to be more frequent.

Because of climate change, there will be an increasing variability in the future predictions for floods, meaning that extreme events, such as droughts and floods, will be more frequent, with longer duration and with higher magnitudes. According to England & Cohn (2007), floods could be significantly reduced through improved mitigation measures that can only be achieved with accurate flood frequency analysis (FFA). FFA is the most common technique used for at-site estimations on flood recurrence magnitude. (Farooq, Shafique, & Khattak, 2018).

The longer the data records are, the more nonstationary in climate need to be investigated. Many of the hydrological predictions are based on assumptions of stationary climate, instead of climate uncertainty, which can lead to biased predictions. The effects caused by that uncertainty need to be reflected in water infrastructure. Also, the use of some optimization models can be alleviated in favor of systematic analysis to achieve satisfactory results. (Stakhiv, 2011).

Statistical or risk analysis is the widely used way to predict flood events through reliable data records. Reliance on statistical FFA depends particularly on the selected distribution, on the correct estimation of the function parameters, on possible outliers and on the length of the observed flood series. (Farooq, Shafique, & Khattak, 2018 apud Saghafian et al., 2014).

In order to perform these statistical analyzes, it is undeniably important to check the data record and its sources. Most of the data used in recent studies come from streamgages. USGS systematic records, for example, come from gages that measure stages and discharges. Stages are the water level measurements (m or ft) above gage datum, generally at every fifteen minutes, which is hereafter combined with discharge measurements or streamflows to determine a relationship for each streamgage. There are systematic data records, historical data records and paleofloods (reconstruction of floods that happened in the past). Authors argue whether it is necessary to combine different information. According to Millington & Simonovic (2011), it is important to use historical data when predicting events, together with statistical distributions that are used to fit the data. Another important thing to consider is that the data records are provided in water years (WY).

The majority of studies focus on gages where the records are not affected significantly by urbanization, regulation, or diversions because these characteristics have a considerable impact on these predictions. The flood-frequency analysis in some desert regions, for example, is tricky due to many zero flows and a short-period data record. Also, there are factors behind the streamgage readings in high elevation areas, such as rainstorms and snowmelt runoffs. All these data records are important to be studied and compared for the sake of the regions where they are located. Ungaged sites, channelization and urban areas need to be given the same importance while considering flood frequency analysis.

In the United States and other countries, Log-Pearson Type III (LP3) has been the most conventional statistical method and largely used for flood frequency analysis (FFA). However, missing peaks are typically ignored when the LP3 method is used. Some papers, such as Bulletins 17 (B and C) and USGS Report 2012–5113, were important for the release of new publications geared towards the refinement of the FFA in the country. The main purpose of these bulletins, for example, was to provide a nationwide, uniform approach for FFA, using the LP3 distribution and method of moments for parametrization. (England & Cohn, 2007). The LP3 statistical method is being used with other manipulations and procedures, such as regionalization, detection, and removal of low outliers, among other manipulations in order to determine the peak flow for any recurrence period. However, missing peaks during periods of systematic data collection typically are ignored when the conventional LP3 method is used (Parrett and others, 2011).

While designing a flood protection system, it is important to select the best design probability of exceedance and magnitude of the events. Therefore, it is important to adapt the existing conventional methods and not just rely on the statistical method itself because it may ignore some rare and occasional events. A question can be brought up regarding the necessity of all these methods used and if they are conservative enough to be adopted in further assessments.

1. PURPOSE OF THE PROJECT

The primary objective of this thesis is to determine whether GEV itself can be a more conservative distribution than LP3 associated with other methods, such as the B17B weighting procedure with SGB, when determining the annual exceedance flows under a changing climate condition This will be done using the data records from streamgages also used in USGS Report 2012-5113 analysis.

2. STATISTICAL METHODS FOR FFA

Distributions are descriptions of the data as models. These descriptions can be the nature, shape or spread. Frequency distributions can help the understanding of many relations between data. Flood-frequency estimation can be done by fitting a known statistical distribution to series of annual peak flows. (Gotvald et al., 2012). A probability density function (PDF), for example, describes the occurrence probability of the records to obtain exact values. The cumulative density function (CDF), instead, is a way to describe how likely a random record will be less than some picked arbitrary value. CDF is the summation (for discrete data) or the integration (for continuous data) of all the values obtained from the PDF to a specific value of interest. Because of that, CDF is a good way to show how conservative is one distribution in relation to another.

Most of the hydrologic data does not fit normal (gaussian) distributions due to skewness of the data. The normal distribution is the one symmetrical around the mean. If not symmetrical around the mean, the data can also be uniform, skewed to the left or skewed to the right. Some natural events can be approximated to a normal distribution. Nevertheless, many of them require other distributions in order to be portrayed correctly. When the population skew is different than zero and the data does not fit a bell-shape curve, that characterizes a normal distribution, this data can be log-normally distributed. However, because hydrologic records are generally short, long-normal distributions do not always provide conservative predictions.

2.1. Log-Pearson Type 3 (LP3)

Log-Pearson Type III is a gamma distribution and probably the most used distribution for hydrologic frequency analysis. It is a three-parameter distribution based on the mean ($\bar{y}$), standard deviation (S_y) and skew (Gs). Naming the three parameters as: $\propto$, β and γ as the shape, scale and shift, respectively; it is possible to set some relations, where: $\propto = \frac{4}{Gs^2}$; $\beta = \left| \frac{S_y \times Gs}{2} \right|$ and $\gamma = \bar{y} - 2\left(\frac{S_y}{Gs}\right)$. Its CDF function can be expressed as:

$$F(x) = \Gamma D(x - \gamma, \propto, \beta) \quad for\ Gs > 0$$
$$F(x) = \Gamma D(\gamma - x, \propto, \beta) \quad for\ Gs < 0$$

Equation 1

3

Where,

$$\Gamma(s) = \int_0^\infty e^{-x} \, x^{s-1} \, dx = \int_0^\infty e^{-x} \, \frac{dx}{x} \quad for \; s > 0$$

Therefore, the sign of the skew defines whether the gamma distribution will be given as the first or the second option (Equation 1) for obtaining the CDF. The gamma function is defined to all complex numbers, except for non-positive integers. Also, the gamma function has no zeros. Therefore, its reciprocal is possible and it is an entire function. Because of that, when using the gamma function for LP3 with contrasting skews obtained for the different data records, it is necessary to stablish aconditional clause (such as what is shown in Equation 1.

The logharitmic transformation is effective in normalizing values that vary widely in magnitude. It is also important for preserving large peak values from dominating the calculation of the population parameters. However, the danger in log transformations is that low outliers are given a great weight. When large values are the focus, small values can be reported as zero if they fall below a certain threshold. (Stedinger, Vogel & Foufoula-Georgiou, 1993). According to Mays (2010), the LP3 frequency factor equation in terms of discharge (Q_T) based on a specific return period (T) can be shown below.

$$\log Q_T = \bar{y} + K(T, Gs) \times S_y$$

Where,

Gs – Coefficient of Skewness

K(T, Gs) or K_T – Frequency Factor

S_y – Standard Deviation

$\bar{y}$ – Mean Value

T – Return Period

Many studies, instead of calculating the actual log-data skew (Gs), use regional skews that are based on a log-skew value previous calculated for a certain region and provided in a map. The regional skew is calculated based on a huge number of streams that may better represent the focused area provided in Bulletin 17B for example. However, Parrett et al. (2011) found that the regional skews, for some reagions with not sufficient records, could not be reliably determined. Therefore, in this thesis, all skews were calculated by gage.

The estimation of station skews for streamgages with short-period of records is also biased because it brings on large sampling errors and, most of the time, it is necessary to deal with records with completely different lengths. For that reason, considering an ideal scenario, it would be necessary to discard several data based on their length and to use just data with the exact same length of record. However, doing that does not represent the real scenario and it does not consist in picking the data ramdonly, provoking biased analyzes regarding statistics.

The variable K (T, Gs), presented in Equation 3, has fundamental importance to the frequency factor equation of LP3. It is composed by *"z"*, which is equal to the data point (x) minus the mean (μ) divided by the standard deviation (σ): $z = \frac{(X-\mu)}{\sigma}$ or simply the inverse function of the standard normal cumulative distribution (norm.s.inv function in Excel) and *"k"* that is the Gs (skewness) over six: $k = \frac{Gs}{6}$. There is a K_T table that relates the skew coefficient and the exceedance probability. It simplifies when it is necessary to do quick analyzes. Part of this table can be seen below.

Table 1 - K_T values for LP3 (Mays, 2010).

	Recurrence interval (Yr)										
	1.0101	1.0526	1.1111	1.2500	2	5	10	25	50	100	200
	Exceedance probability										
Skew coeff.	.99	.95	.90	.80	.50	.20	.10	.04	.02	.01	.005
3.0	−0.667	−0.665	−0.660	−0.636	−0.396	0.420	1.180	2.278	3.152	4.051	4.970
2.9	−0.690	−0.688	−0.681	−0.651	−0.390	0.440	1.195	2.277	3.134	4.013	4.909
2.8	−0.714	−0.711	−0.702	−0.666	−0.384	0.460	1.210	2.275	3.114	3.973	4.847
2.7	−0.740	−0.736	−0.724	−0.681	−0.376	0.479	1.224	2.272	3.093	3.932	4.783
2.6	−0.769	−0.762	−0.747	−0.696	−0.368	0.499	1.238	2.267	3.071	3.889	4.718
2.5	−0.799	−0.790	−0.771	−0.711	−0.360	0.518	1.250	2.262	3.048	3.845	4.652
2.4	−0.832	−0.819	−0.795	−0.725	−0.351	0.537	1.262	2.256	3.023	3.800	4.484
2.3	−0.867	−0.850	−0.819	−0.739	−0.341	0.555	1.274	2.248	2.997	3.753	4.515
2.2	−0.905	−0.882	−0.844	−0.752	−0.330	0.574	1.284	2.240	2.970	3.705	4.444
2.1	−0.946	−0.914	−0.869	−0.765	−0.319	0.592	1.294	2.230	2.942	3.656	4.372
2.0	−0.990	−0.949	−0.895	−0.777	−0.307	0.609	1.302	2.219	2.912	3.605	4.298
1.9	−1.037	−0.984	−0.920	−0.788	−0.294	0.627	1.310	2.207	2.881	3.553	4.223
1.8	−1.087	−1.020	−0.945	−0.799	−0.282	0.643	1.318	2.193	2.848	3.499	4.147
1.7	−1.140	−1.056	−0.970	−0.808	−0.268	0.660	1.324	2.179	2.815	3.444	4.069
1.6	−1.197	−1.093	−0.994	−0.817	−0.254	0.675	1.329	2.136	2.780	3.388	3.990
1.5	−1.256	−1.131	−1.018	−0.825	−0.240	0.690	1.333	2.146	2.743	3.330	3.910
1.4	−1.318	−1.168	−1.041	−0.832	−0.225	0.705	1.337	2.128	2.706	3.271	3.838
1.3	−1.383	−1.260	−1.064	−0.838	−0.210	0.719	1.339	2.108	2.666	3.211	3.745
1.2	−1.449	−1.243	−1.086	−0.844	−0.195	0.732	1.340	2.087	2.626	3.149	3.661

(*Continued*)

However, instead of using table 1, is is much more precise to calculate K_T through the Equation 4 below.

$$K(T, Gs) = z + (z^2 - 1)k + \frac{1}{3}(z^3 - 6z)k^2 - (z^2 - 1)k^3 + zk^4 + \frac{1}{3}k^5$$

Equation 4

5

2.2. Generalized Extreme Value Distribution (GEV)

The GEV distribution is a probability distribution within extreme value theory to combine Frechet (1927), Weibull (1951) and Gumbel (1958) families of distributions. It is used as an approximation to model the extrema or largest or smallest values from long sequences. According to Stakhiv (2011), GEV gives more weight to extreme events, being considered a "fat-tailed" distribution. Like LP3, GEV is also a three-parameter distribution. It consists in three basic parameters: shape (k), scale (σ), location parameters (μ). (Hajani & Rahman, 2018). Depending on the shape parameter, it can be either Frechet, Gumbel or Weibull distributions:

k = 0 (type I GEV or Gumbel distribution) – EV1

k > 0 (type II GEV or Frechet distribution) – EV2

k < 0 (type III GEV or Weibull distribution) – EV3

The probability density function (PDF) for GEV can be expressed by Equation 5.

$$F(x|k, \mu, \sigma) = \left(\frac{1}{\sigma}\right) t(x)^{k+1} \ e^{-t(x)}$$

Where,

$$t(x) = \begin{cases} \left[1 + k\left(\dfrac{x - \mu}{\sigma}\right)\right]^{-1/k} & for \ k \ \neq 0 \\ e^{-\frac{(x-\mu)}{\sigma}} & for \ k = 0 \end{cases}$$

Equation 5

Figure 1 shows a graph of PDFs versus z-score for the three different types GEV.
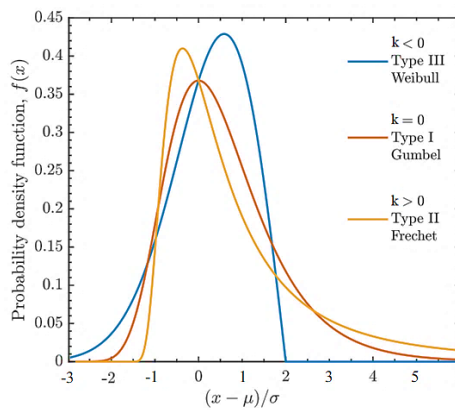
Figure 1 - Graph of PDFs of the different types of GEV based on the shape parameter k.

For Weibull's type III, the shape parameter is less than zero and equal to ⁻½ and for Frechet's type II, the shape parameter is higher than zero and equal to ½ (in figure 1). The skewness is a measure of the lack of symmetry in a distribution. A positive skewness means extreme events that occure in the right-side tail of the distribution (Millington, Das & Simonovic, 2011).

6

Gumbel, also called EV1 distribution, has its shape parameter equal zero. Because of that, Gumbel becomes a two-parameter distribution. According to The MathWorks (2022), distributions whose tails decrease as a polynomial, such as student's T, result in a positive shape parameter. When the tails decrease exponentially, such as normal, result in a zero-shape parameter. And when the tails are finite, such as beta, result in a negative shape parameter.

The general mathematical form and cumulative density function (CDF) which incorporates GEV types I, II and III is shown in Equation 6 below.

$$F(x|k,\mu,\sigma) = \exp\left\{-\left[1 + k\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/k}\right\}$$

Equation 6

According to Mays (2010), the equation to predict flood magnitude based on a specific return period and the three parameters can be seen below (Equation 7).

$$Q_T = \mu + \binom{\sigma}{k} \times \left\{1 - \left(-log\left(\frac{T-1}{T}\right)\right)^k\right\}$$

Equation 7

where,

k – Shape Parameter

σ – Scale Parameter

μ – Location Parameter

T – Return Period

Figure 2 shows the behavior of the GEV distributin based on the three parameters.
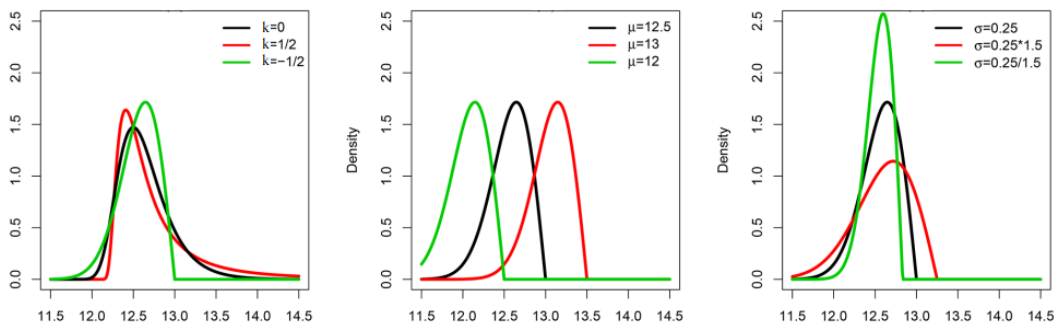
Figure 2 - Behavior of the distribution considering the three GEV parameters. (Adapted from Rohmer et al., 2020).

The shape parameter (k) defines the distribution classification. The location parameter (μ) is responsible for shifting the distribution to the left or to the right and the scale parameter (σ) is responsible for stretching or compressing the distribution. The smaller the μ value, the more shifted to the left is the distribution and the smaller the σ value, the more stretched is the distribution.

2.3. Comparison between the Distributions in the Literature

In the work of Millington, Das & Simonovic (2011), three distributions were used to estimate the probability of future maximum occurrences: GEV, LP3 and EV1. The authors evaluated datasets from the Upper Thames River Watershed with two basic procedures: the goodness of fit tests, used in FFA to estimate best distribution to fit an observed data (Farooq, Shafique, & Khattak, 2018), and the L-Moment Ratio Diagrams (based on the probability-weighted methods and on linear combinations in ascending order) to define how appropriate was each distribution.

The GEV distribution proved to be the strongest fitting distribution while LP3 distribution showed to be the second best fit. Gumbel or EV1 distribution showed to be the worst fit of all. They concluded by saying there was a necessity for more studies with the application of GEV on other watersheds in Canada to confirm its countrywide applicability.

According to Vogel & Wilson (1996), the LP3 distribution was a standard model in more countries than GEV was and some of the countries reported the use of more than one model as a standard in that time. For them also, countries should reevaluate their standards when choosing a suitable distribution for FFA. The authors used a bibliography showing the use of goodness of fit test for many of the models and concluded that three parameter log-normal (LN3), the LP3 and the GEV were all acceptable models to use for FFA in the country while other two and three-parameter models were not acceptable for the entire continent. The study also revealed that annual minimum flows were best approximated by Pearson III distribution. Using the L-Moment diagrams they also reveled that LP3 were a flexible distribution, being able to fit many series of annual maximum, average and minimum streamflows in US.

Considering maximum streamflows, Stakhiv (2011), in his work for the Journal of the American Water Resources Association (JAWRA), presented that a flood estimated to have a 100-year return period using the LP3 distribution, had only a 47-year return period using the GEV distribution (figure 3).
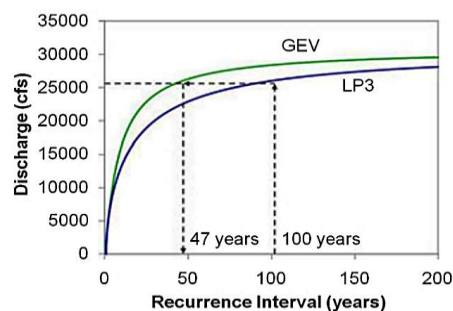

Figure 3 - Comparison of GEV and LP3 Probability Distributions for Flood Frequency Analysis. (Stakhiv, 2011).

8

Figure 3 shows that the same discharge, in cfs, is predicted to happen much sooner when using the GEV distribution instead of LP3. Stakhiv (2011) finishes his work by saying that is necessary, considering the climate change's assumption to change evaluation procedures and mentions the USACE's proactive adaptive management approach, such as replacing LP3 and applying GEV probability distribution for flood frequency analysis (FFA).

Farooq, Shafique & Khattak (2018), in their work, used different two and three-parameter statistical distributions: Generalized Extreme Value (GEV), Log Pearson 3 (LP3), Gumbel Max, and Normal to hydrological stations in Pakistan, where floods are among the most devastating and recurring natural hazards. Also, they utilized a software package, "EasyFit", to apply goodness of fit tests, such as Anderson-Darling (AD), Kolmogorov-Smirnov (KS), Chi-Squared ($\chi 2$), among others at a 5% significance level ($\alpha = 0.05$) to the observed data. The chi-squared test is not considered a highpower statistical test. Anderson-Darling test, in the other hand, is responsible for comparing the fit of an observed CDF to an expected CDF (figure 4).
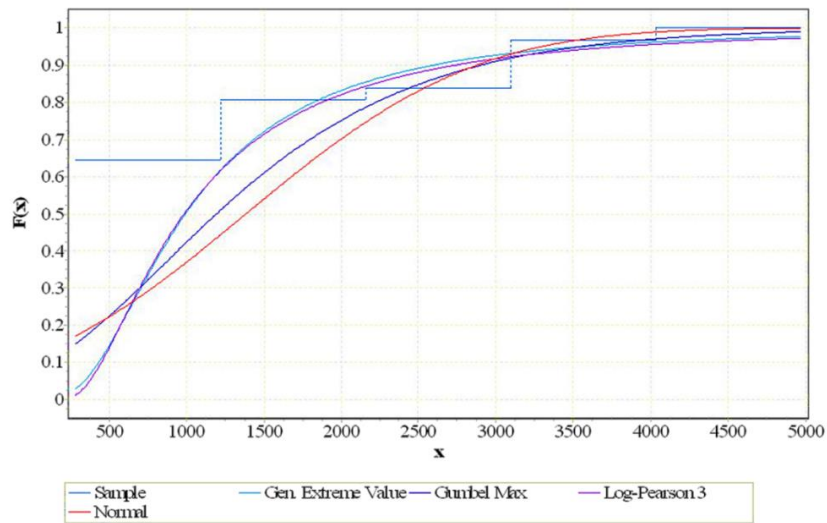
Figure 4 - CDF curves for the distributions of GEV (light blue), LP3 (purple), Gumbel Max (dark blue) and Normal (red) and the sample data as a stair graph.

In figure 4, the stairs graph represents the historical data and the other curves are the observed CDFs. The results, for their work, indicate that LP3 and GEV were ranked top two distributions at all locations while Gumbel Max and Normal were the least fitted. Two-parameter distributions have smaller standard error, but larger bias than distributions with more parameters. The study showed that LP3 was ranked 1 by all three tests even for smaller sample sizes (the length of the data is extremely important in the consideration of which distribution to use). Goodness of fit has less significance if one is assuming a changing climate where trends in the mean and variance may be present. GEV was ranked 1 for Chakdarra and Munda

Headwork while LP3 distribution was ranked 1 for Khwazakhela and Panjkora gauge stations. GEV was more suitable for moderate slope regions in the Swat valley and LP3 more suitable distribution at steep valley. They also concluded that a single distribution cannot be specified as the best fit distribution for all locations. (Farooq, Shafique & Khattak, 2018).

2.4. Outliers

In the hydrological scenario, there are peaks that are considerably higher or lower than the rest or the peaks recorded by a certain gage. In statistics, an outlier is a value or occurrence that notably differs from the rest of the data record. Low outliers are unusually small observations. (Stedinger, Vogel & Foufoula-Georgiou, 1993). Most of the outliers can be detected when all the data values are plotted to a graph.

The outliers can indicate experimental errors, variability in the measurements or unusual events. According to Parret et al. (2011), low outliers are peak records that are significantly smaller than the others, having a large effect on the Log Pearson Type III distribution fit to all the recorded data. In the same study, the authors identified all the low outliers that have large influence in the upper tail of the curve formed by the fitted data. The low-outlier censoring threshold is a strategy that can eliminate values that can change the fitting curve for one that is not very representative for the rest of the peak values.

2.4.1. Grubbs-Beck Test

The Grubbs-Beck test is recommended for finding low outliers that lead to influences in the fitting curve and, mostly, in the upper tail fitting that means larger flows with smaller annual exceedance probabilities. (Gotvald et al., 2012). It is used to detect an outlier in a data set (like in figure 5) that follows approximately a normal distribution.
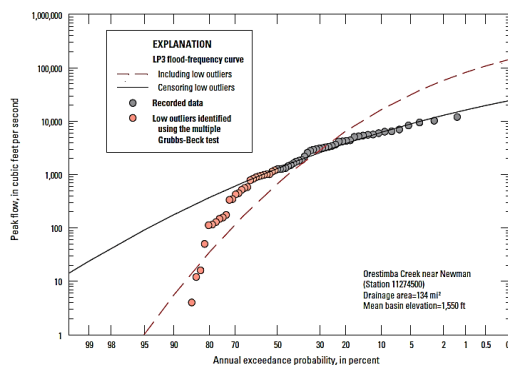

Figure 5 - Example showing the effects of including or censoring potentially influential low outliers identified from the multiple Grubbs-Beck test. (Gotvald et al., 2012).

Figure 5 shows an example of the application of multiple Grubbs-Beck test that detects and excludes multiple low outliers that could pottentialy affect the fitting curve. It is possible to see that the curves in the figure are very different. The orange curve will present higher peak flow values as the AEP gets lower, in other words, while the return periods get higher, the fitting curve will not provide values as high as the ones provided in case the outliers are included.

2.5. Trend Analysis

Trends are generally a problem for further use of the data. They can lead to a wrong analysis depending on the time frame analyzed. Parametric and non-parametric tests are usually applied to detect trends in data. The parametric test is used in independent and normally distributed data. The non-parametric does not need the data to be normally distributed but requires only independency in data and allowence of outliers. It is possible to detect trends using hypothesis tests, present in inferential statistics, also called tests of significance. They are procedures for testing claims about a property of some data. (Triola et al., 2006).

2.5.1. Mann-Kendall Test

It is a complex method to detect linear or other types of trends. When the data presents trends, it can lead to errors in the fitting analysis. If a sample is not selected randomly, it is possible that it is biased in some way, not representing the situation correctly. The Mann-Kendall test examines whether to reject the null hypothesis H_0 (no monotonic trend) and accept the alternative hypothesis H_A (monotonic trend, not necessarily linear). In MatLab, if the test returns $H = 1$ that indicates a rejection of the null hypothesis.

The null hypothesis H_0 means that there is no monotonic (increasing or decreasing) trend and that the time series values are independent. In Matlab, using the hypothesis test, no trend is obtained where $H = 0$ that indicates a failure to reject the null hypothesis at a certain alpha significance level. The alpha value is commonly 0.05 because it is compared to the p-value.

2.5.2. P-Value

The p-value shows how likely it is to get the result obtained. If the p-value is greater than the alpha value, there is insufficient evidence to reject the null hypothesis. Therefore, the null hypothesis exists, meaning no trend. When the p-value is less than or equal to 0.05, trends are considered to be significant. A p-value of 0.05 indicates that there is 5% of chance that the trend test used will identify a trend even when there is no actual trend present.

3. PROCEDURE AND TOOLS

The main tools used for this project consisted in: the annual peak data gathered from the USGS Water Data for USA (2022) posteriorly modeled and manipulated; Excel for calculations and the organization of the gages in data lengths, latitudes and longitudes, gages with missing years or zero floows, log data, LP3 skews and other important analyzes; PeakFQ for obtaining the LP3 fitted curve and the annual exceedance probability for the data (using B17B global analysis method with Single Grubbs-Beck test option) and MatLab for the GEV test (doing trend analysis, such as the Mann-Kendall and using the GEV functions for estimation reasons). The final procedure was to compare CDFs between LP3 and GEV distribution for the data records analyzed. Also, the streamgages chosen were also analyzed by USGS report, by Gotvald et al. (2012), that provides some approaches to determine FFA for streamgages with ten or more years of annual peak-flow record in southeastern California and for eight other selected streamgages affected by urbanization.

3.1. Focused Region

The focused region, chosen to be analyzed, was Santa Ana and San Gabriel rivers (figure 6). Their streamgages were also mentioned in USGS Scientific Investigation 2012-5113 Report by Gotvald et al. (2012).
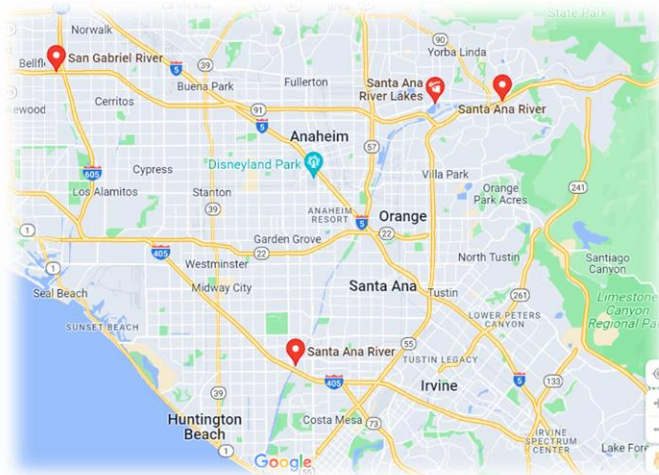
Figure 6 - Location of San Gabriel and Santa Ana rivers pinpointed in red. (Google Maps, 2022).

Both of these regions, presented in figure 6, are densely populated but with no legend of urbanization in the data records gathered from USGS website for these areas. San Gabriel is located at 33° 53' 16.134" N and 118° 6' 25.722" W and Santa Ana downstream at 33° 42' 22.3308" N and 117° 56' 0.3264" W. All the streamgages data gathered and analyzed are in the buffer area of around 60 km of radius.

Even before the growing urbanization in the region, Santa Ana River was already considered as a river with a potential for producing extreme floods and have historical dangerous paleofloods that have occurred in intervals of approximately 30 years in the region (1780, 1825, 1862, 1867, 1884, 1891, 1910, 1916, 1938, 1969 and 1995) and with the increase in urbanization, the flood threat was increased as well (Clarke, 1996). According to Guinn (1890), the flood of 1825 changed the course of the Santa Ana River and he mentions other episodes in 1832 and after that responsible for changing the countour of the south of the city and the drainage that changed the vegetation in the region and that was lately followed by droughts. Also, according to the same author, the flood of 1884 cut a channel to the sea. All of the changes in channels that happened through the years, in the region, are because of the formation of deltas (formed by the set of lowlands originated from the accumulation of alluvial materials). There are some concerns near these regions. According to Orsi (2004), in 1920s, there was a construction of a high San Gabriel dam to solve flooding with a single block of concrete. This was a trial that failed because the hydrology and drainage in the location were not correctly understood.

Figure 7 shows only the map ID streamgages considered in the analysis (table 1, in the next page, shows all of the analyzed data records). They were considered based on the length of years representing the record, if there were missing periods of systematic data.
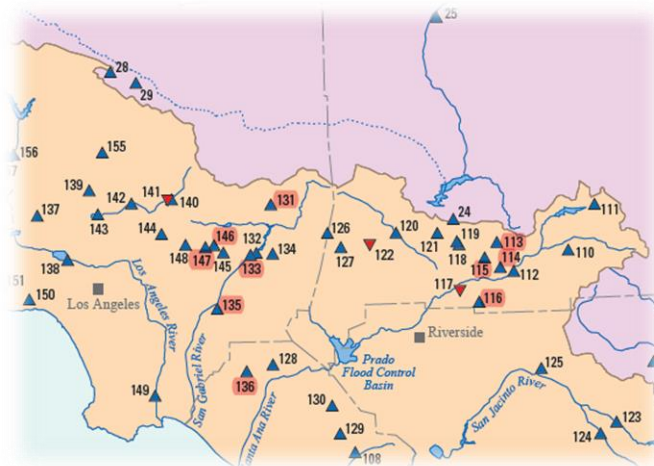
Figure 7 - Streamgages with data collected from USGS website. (Adapted from the map of Gotvald et al., 2012).

Figure 7 presents a 1:100,000-scale map, projected as UTM, Zones 10 and 11 by USGS, Gotvald et al. (2012). These selected final streamgages were considered because there were no long missing periods of data record (table 1). There were no indications of thresholds for outliers (section 3.3) in the data and no perceived trend (section 4 – results).

3.2. Assembling the Data

The updated data was gathered from USGS Water Data for USA (2022) considering the selected streamgages (in red) and the information in annual peak flows. The process of gathering and analyzing the characteristics of the data starts by gathering the data as a table-separated file and peakfq file according to the site's identification number (ID) according to what is shown in figure 8A and B.
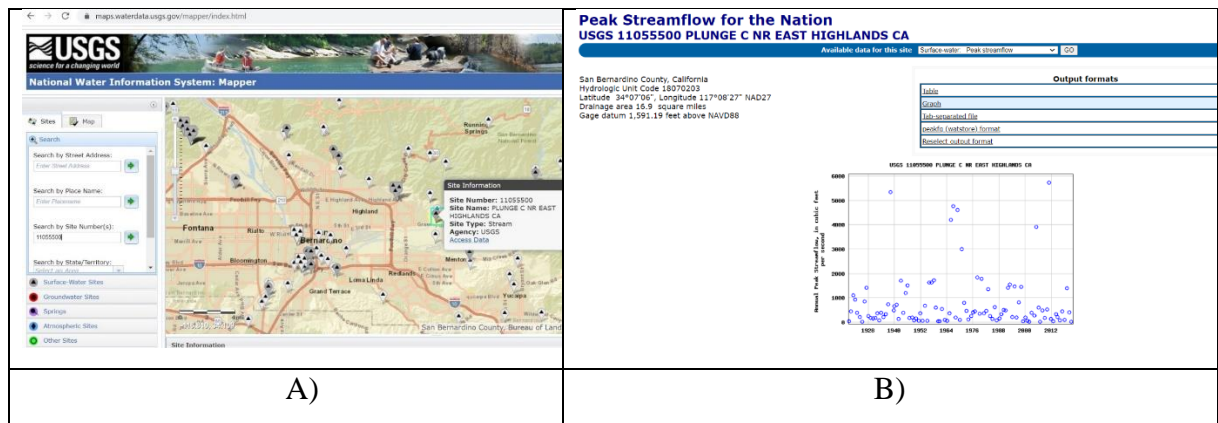
Figure 8 - Process of gathering the data records and other important information in USGS Water Data for USA (2022) website by using the streamgage ID.

In the website map, shown in figure 8A, it is possible to see where the streamgage selected is located, for example, which river is providing information on peak flows for a certain streamgage or if it is closed to urbanized spots, affected by channelization, among other factors (all streamgages analyzed in this thesis are affected, in some way, by urbanization). The name on the streamgage is generally referred to the region or lake which it is inserted. Figure 8B shows the type of information as it is possible to see in figure 9 that shows the information of gage 114 to use in the PeakFQ software.

Figure 9 - Data record from 114 streamgage for PeakFQ.

All the data gathered have more than 10 years of record. The gages marked with asterisk, in the table 2 below, were analyzed but not considered since there were missing periods that could potentially interfere in the whole analysis.

# of the Gage (From when to when)	ID	Number of Water Years on the record	Mean Stream-Flow for the Annual Peak Records (cfs)	Maximum Stream-Flow for the Annual Peak Records (cfs)	Minimum Stream-Flow for the Annual Peak Records (cfs)	Standard Deviation of the Raw Data	Log Skew
113 1960-1973	11055300	13	179.15	620	12	190.98	-0.2483
114 1919-2021	11055500	103	770.55	5740	9.7	1148.51	-0.2128
115 1989-2021	11055801	33	1330.37	9900	11	2404.15	0.0797
116 1927-1979	11057000	51	1253.88	14000	24	2376.39	0.1899
118 1920-2021	11058500	100**	528.34	6000	7.1	886.89	0.1056
127 1928-1975	11073470	46	668.56	10300	9.9	1629.75	0.2429
128 1950-1961	11075740	11	192.52	935	2.3	302.59	-0.1162
131 1960-1973	11081200	13	516.39	2080	8	698.38	-0.1982
135 1965-1978	11086990	14	6602.86	11100	3580	2843.88	0.3184
136 1932-1969	11089000	38	473	3700	20	699.59	0.1325
144 1914-2020	11098000	107**	1136.84	8620	12	1618.40	-0.3909
146 1916-1970	11100000	54	669.59	7000	17	1204.54	-0.0129
147 1916-1962	11100500	46	66.25	536	2.3	95.82	0.3057
148 1916-1965	11101000	50	356.16	2400	4	454.59	-0.4974

Legend: ** - Missing data records in some of the years.

Also, another completely different analysis was done for the 8 urbanized streamgages chosen in USGS 2012-5113 Report. According to Gotvald et al. (2012), the reason for choosing these 8 streamgages was due to trend analysis and data quality review. The information regarding these eight streamgages selected, including the station name that mentions the region in which the streamgage is located, station ID and map identification number, can be seen in figure 10.

Figure 10 - Information about the 8 selected gages in USGS Report 2012-5113 by Gotvald et al. (2012).

Even the report mentioning that there was data qualiy review, when gathering the updated annual peak flow for these gages, there are mentions of urbanization affecting part of the data: "C → All or part of the record affected by urbanization, mining, agricultural changes, channelization, or other". Table 3, below, shows the same information provided in table 1 but considering only these eight gages.

Table 3 - General information and characteristics belonging to each of the eight separated streamgage records.

Gage	ID	Number of Water Years on the record	Mean Stream-Flow for the Annual Peak Records (cfs)	Maximum Stream-Flow for the Annual Peak Records (cfs)	Minimum Stream-Flow for the Annual Peak Records (cfs)	Standard Deviation of the Raw Data	Log Skew
772	11023330	23	1294.78	4990	102	1269.64	0.2241
773	11023340	57	1779.03	5730	49	1579.74	-0.443
774	11047200	11	1458.55	5150	425	1428.89	0.719
775	11120000	80	2032.50	10200	2.1	2196.40	0.030
776	11162720	32	1761.94	3560	610	833.78	0.644
777	11162800	38	243.05	644	16	143.76	-0.734
778	11182500	68	410.06	1600	9.96	406.61	-0.611
779	11447360	41**	1435.80	3450	312	696.68	-0.842

Legend: ** - Missing data records in some of the years.

Because of the characteristics of the missing records in both tables 2 and 3 and considering that the objective is to investigate which type of distribution is more conservative, the streamgages with missing records were analyzed but not considered for the final results.

3.3. PeakFQ

PeakFQ is a program that follows Bulletins 17B guidelines of the Interagency Advisory Committee on Water Data with implementations from Bulleting 17C and provides estimates for flood magnitudes for different AEPs using the LP3 method and graphs as outputs with fitted

frequency curve, low outliers, systematic, censored and historic peaks, thresholds and CIs. Figure 11 shows an example of running a streamgage data record in PeakFQ.
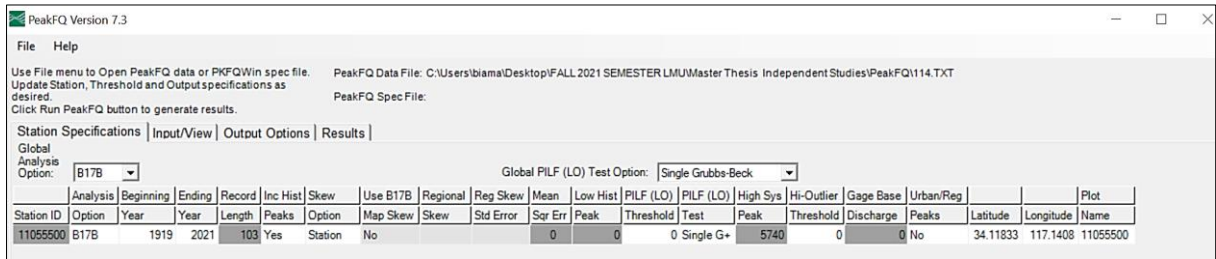

Figure 11 - Running data from the streamgage 114 in PeakFQ using station skew and B17B and Single Grubbs-Beck as the test option.

As it is possible to see in figure 11, the station skew was used for all the streamgages because regional skew tends to generalize the analysis by region, not portraying precisely what happens in particular with each river or water body analyzed.

B17B weighting procedure, used in the analysis (figure 11) employs the guidelines from Bulletin 17B. According to England & Cohn (2007), despite the utilities and guidance that this weighting method has provided over the years because of its consistency, newer research in the field shows that if the method was revised more accurate frequency estimations could be obtained. The B17B confidence interval method is based on applying a 2-parameter log-Normal distribution procedure to the LP3 method. Therefore, when the third parameter (which is the skew) is different than zero (not turning into a 2-parameter log-Normal) the CIs would be too narrow for the characterization of the data. This is an important information when comparing to the wider 95% CIs obtained from the GEV distribution in MatLab.

In the PeakFQ section of USGS website it is possible to find a message warning that sequencing and computational issues have been identified in the software and Bulletin 17C guidelines (B17C). Therefore, the USGS, in collaboration with the U.S. Army Corp of Engineers, is trying to update these methods.

The Expected Moments Algorithm (EMA) is a special analysis option for the application of LP3. It detects multiple potentially low outliers when fitting the LP3 distribution. EMA uses interval discharges when characterizing missing data in periods of systematic collection, floods of unknown magnitude that exceed some value (binomial censored), floods of unknown magnitudes that are less than some value (censored from below) and floods with magnitudes described by a range (interval censored). It uses data that includes multiple thresholds for low outliers and uncertainty with ranges of floods. (England & Cohn 2007).

The use of Single Grubbs-Beck option for detecting potentially influential low-flow (PILF), in PeakFQ, was the only detection method considered when applying the B17B as the global analysis option. It is interesting that when trying to run EMA instead of B17B with the use of SGB for some of the streamgages data with no missing records, the confidence intervals of 95% got higher, reducing accuracy.

Since, the objective of this thesis is to determine how conservative is GEV when compared to LP3, there was no use of unknown data records, no attempt to address low outliers or zero flows, EMA has not been given focus. In addition, when trying to run EMA for gages with missing systematic record, the program fails to use a known threshold in the range of zero to infinity and fails run as well. It displays an attention window indicating the use of an alternative analysis option, such as "skip" or "B17B" instead of running EMA with that data.

PeakFQ provides two types of output results: as a graph (annual peak discharge in cfs by AEP in %), results included in Appendix I, or as a text output (figure 12) showing also the confidence intervals provided by the test methods used with the LP3 distribution.

Figure 12 - Estimated peak using B17B estimation method and its 95% confidence intervals for 100-year return.

Figure 12 shows an example of results obtained running the data from streamgage 114. One of the goals of this thesis is to compare the 100-year peak flow obtained from in from LP3 using PeakFQ (with B17B as the Global Analysis Option) with the 100-year peak flow obtained using the GEV in MatLab and also to stablish in what year the prediction would have provided the same result.

3.4. MatLab

After treating and filtering the data records, by analyzing their caractheristics and parameters, it is necessary to upload the data to MatLab (figure 13) to evaluate whether the data have any trend. Because, as mentioned in section 5, the data with trend is biased and they can lead to a wrong conclusion.

Figure 13 - Uploading the filtered data records into MatLab. Example of selecting the data from streamgage 114.

The answer for trends or not in the data came from downloading the Mann Kendall Test and using the following code considering a small alpha value of 0.05, already discussed in section 5.1: [H, p_value]=Mann_Kendall(GageNumber,.05).

The answer of 0, obtained for the majority of the data, means that the test identified no trend under the assumption of the alpha value used. The answer of 1 means that trend was detected. Also, it is important to consider the p_values obtained in this test. If the p_value is bigger than the alpha value, no trend should be detected while a smaller p_value than the alpha means that there is a possible trend in the data.

MatLab was also used for obtaining the GEV parameters, one of the most important steps in this whole analysis, since they provide the fitting of the data when using a GEV distribution. This parameterization process defines not only the parameters necessary to obtain a relevant model but it presents the behavior of the curve regarding the GEV distribution.

The parmhat = gevfit(X) returns maximum likelihood estimates of the parameters for the generalized extreme value (GEV) distribution given the annual peak flows data in a specific streamgage record. This code returns: $1 - k$ (shape parameter), $2 - \sigma$ (scale parameter) and $3 -$

μ (location parameter). The other option is to use parmhat = gevfit(X) and [parmhat,parmci] = gevfit(X,alpha) for obtaining not only the parameters but also the 95% confidence interval (CIs) for each of the streamgages data.

The script below shows a basic path to define the parameters and the projection of the maximum peak streamflow for 20 years.

Figure 14 - Matlab GEV code lines used for parameterization and peak streamflow projection for Gage 113.

To compare peak values at specific return periods, in this project, the function "gevinv" (the inverse of the GEV) was utilized. The function R0lml = gevinv (figure 14) returns the GEV predicted flow for a specific return period.

The standard normal variable (z-score) is used in the hypothesis testing, meaning that the bigger the z-score, the more distant from the mean the value is, under the area of the probability curve. The area under the curve goes from 1 to 100% and the inverse of the z-score provides information on the standard normal probability. The cumulative area is $1 - 1/T$, where T is the return period. In Excel, the use of the command norm.s.inv (0.99) means the need of a projection considering a return period of 100-years and provides the z = 2.32628.

In figure 14, for example, the use of 0.95 means that 20-years is being considered for the return period. The rest of the information in the same function are the parameters obtained in the parametrization for the gage 113 in the example. Therefore, this function can also be used when comparing the results obtained using LP3 distribution to GEV, making it possible to know in what return period the second distribution can be compared to the 100-year return period of the first one.

4. RESULTS AND DISCUSSIONS

4.1. Trend Test Results

The results for the non-parametric trend test, Mann Kendall, which detects trend for all types of distributions, performed in MatLab for the ten main streamgages, can be seen in table 4.

Table 4 - Results from the Mann Kendall trend test performed in MatLab for the main streamgages.

| Gage Map ID | # WY | Mann Kendall Trend Test | |
		H (0 = no trend and 1 = trend)	p-value
Gage 113	13	0	0.2224
Gage 114	103	0	0.5477
Gage 115	33	0	0.3139
Gage 116	53	0	0.5642
Gage 131	13	0	0.5022
Gage 133	45	0	0.9051
Gage 135	14	0	0.7426
Gage 136	38	0	0.0067
Gage 146	54	0	0.286
Gage 147	46	0	0.6839

The same trend analysis was done for the USGS selected streamgages, except for streamgage 779, which was not considered because of the number of missing periods of data record. Also, gages were evaluated using both: the updated data record and the data record allowed for generating results in PeakFQ as well as what has been done by USGS.

Table 5 - Results from the Mann Kendall trend test performed in MatLab for the selected streamgages in the USGS report.

| Gage Map ID | # WY | Mann Kendall Trend Test | |
		H (0 = no trend and 1 = trend)	p-value
Gage 772 – Until 1993	23	1	0.0140
Gage 772 - Until 1983	13	1	0.0060
Gage 773 – Until 2021	57	1	0.0021
Gage 773 – Until 1983	19	0	0.0589
Gage 774 – Until 1981	11	0	0.1611
Gage 775 – Until 2021	80	1	0.0014
Gage 775 – Until 1957	16	0	0.7187
Gage 776 – Until 1996	32	1	0.0001
Gage 776 – Until 1977	14	0	0.3244
Gage 777 – Until 1996	38	0	0.7343
Gage 778 – Until 2020	68	0	0.5856

From table 5, it is possible to see that, when the p-value is less than or equal to 0.05, trends are considered to be significant. Also, when some of the streamgages had their data not updated yet, most of them passed in the trend analysis, except for gage 772 that has trend considering both the updated and the non updated data records. Therefore, it is possible to infer that 4 (four) out of 7 (seven) streamgages, using their updated records, will provide a biased analysis when considering trends in the data records.

4.2. The 100-year Expected Floods with 95% CIs for LP3 in PeakFQ

In Appendix I, it is possible to visualize the fitted frequency curves with CI curves for annual peak discharge (cfs) by AEP (%), using LP3 distribution with B17B global analysis method and SGB obtained in PeakFQ for the streamgages with no trend detected. For the data records used, there were no perception thresholds detected in PeakFQ. The results for the 100-year projection for expected floods with 95% CIs can be seen in Table 6 (for the main streamgages).

Table 6 – 100-year expected flood with 95% CIs, using LP3 distribution with B17B and SGB in PeakFQ for the main streamgages.

Gage Map ID	# WY	100-year expected flood LP3 Distribution using B17B SGB in PeakFQ	Cis	
			5% Lower	95% Upper
Gage 113	**13**	1,660	665.7	8,680
Gage 114	103	6,862	4684	10,910
Gage 115	**33**	21,800	9518	71,570
Gage 116	53	14,940	8405	32,230
Gage 131	**13**	9,706	2660	101,800
Gage 133	45	3,974	2275	8,410
Gage 135	**14**	17,910	13130	31,430
Gage 136	**38**	7,133	3057	36,160
Gage 146	54	6,975	4073	14,170
Gage 147	46	481.5	294.6	929.7

The bolded WY in table 6 are the shortest lengths of records. Figure 13 shows a graph with the 100-year flood projections with 95% CIs as error bars using the LP3 with SGB in PeakFQ. In this figure, it is possible to see that gage 115 with 33 (tirty three) years of record, gage 131 with 13 (thirteen) years of record and 136 with 38 years of record present the wider CIs.
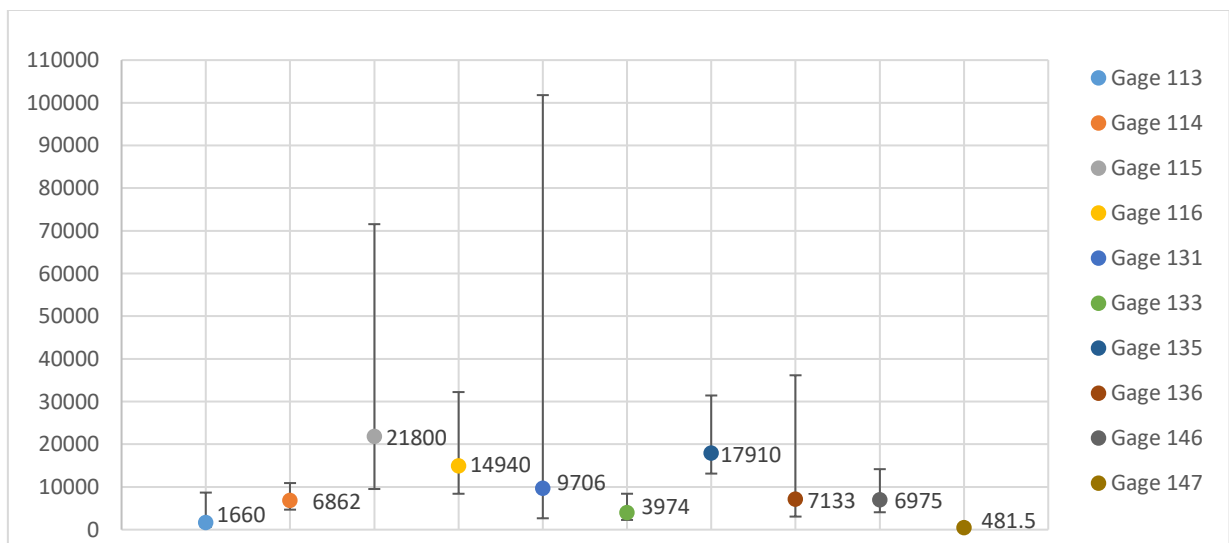

Figure 15 – Graph of the 100-year flood projection with 95% CIs as error bars using LP3 in PeakFQ.

In addition to the results provided in table 6 and in figure 15, it was possible to run two EMA examples just to compare the results with the B17B global analysis method. For gage 113, the 95% upper limit got higher than 8680 (B17B with SGB): 18000 (using EMA with SGB) and 14380 (using EMA with MGB). For gage 114, the CIs got wider as well with the 95% lower limit dropping to 4393 and upper limit rising to 12930 (EMA with SGB) and lower limit dropping to 4375 and upper limit rising to 12820 (EMA with MGB).

Table 7 presents the same 100-year projection for expected floods with 95% CIs for the separated selected by USGS streamgages.

Table 7 – 100-year expected flood with 95% CIs, using LP3 distribution with B17B and SGB in PeakFQ for the USGS selected streamgages.

Gage Map ID	# WY	100-year expected flood LP3 Distribution using B17B SGB in PeakFQ	CIs 5% Lower	CIs 95% Upper
Gage 772 – Until 1993	23	11,640	4,939	56,120
Gage 772 – Until 1983	13	11,640	4,939	56,120
Gage 773 – Until 2021	57	10,950	5,145	36,800
Gage 773 – Until 1983	19	10,950	5,145	36,800
Gage 774 – Until 1981	11	10,800	5,242	46,380
Gage 775 – Until 2021	80	12,770	6,486	34,580
Gage 775 – Until 1957	16	9,573	3,862	45,920
Gage 776 – Until 1996	32	3,931	2,784	7,383
Gage 776 – Until 1977	14	3,931	2,784	7,378
Gage 777 – Until 1996	38	658.4	513.4	922.6
Gage 778 – Until 2020	68	2,366	1,624	3,794

It is possible to see that, using the same LP3 distribution and global analysis method, in PeakFQ, the results for most of the streamgages do not change for the updated records. There was one rising 100-year flood projection for gage 775 but, considering that it fails the trend test for the updated record, it is just possible to infer that the projection and CIs are now higher.

4.3. Parametrization considering GEV Distribution

The parametrization for each streamgage considering the GEV distribution can be seen in the table 8 below (just the streamgages with no trend were considered). The errors up and down considering the 95% CIs will be provided in section 4.5 for GEV.

Table 8 - Parametrization for all the streamgages using GEV.

Gage ID	Shape Parameter (k)	Scale Parameter (σ)	Location Parameter (μ)
Gage 113	1.9856	41.4961	31.4257
Gage 114	0.8987	247.9207	207.5725
Gage 115	1.3323	266.4215	182.3505
Gage 116	1.0531	318.3207	263.5669
Gage 131	2.1084	86.9863	47.0667
Gage 133	0.9078	161.0979	129.2919
Gage 135	0.0745	620.0019	1354.9161
Gage 136	1.0260	124.5994	114.7762
Gage 146	1.1112	185.8892	147.9083
Gage 147	0.6476	25.9856	23.8519
Gage 773 – until 1983	0.8407	483.6010	424.3141
Gage 774	1.7296	242.0899	549.8886
Gage 775 – until 1957	0.9267	244.0264	192.6360
Gage 776 – until 1977	0.3433	327.7984	943.0939
Gage 777	-0.0272*	114.6851	179.5366
Gage 778	0.5685	182.2710	175.3703

This parametrization, indicated in table 8, performed in MatLab according to section 3.4 can be considered the most important step performed in this thesis because a parametrization done correctly can avoid major mistakes while fitting the data records in the GEV distribution and also comparing to the LP3 results.

Considering the shape (k) parameters in table 8, responsible for the classification of the distribution, Frechet (EV2) is the distribution detected for fitting all of the streamgages data. The only exception of gage 777, which has its shape parameter approximating its distribution curve to a Weibull (EV3). However, considering the CIs obtained also from the parametrization, and its positive interval, this distribution can also be evaluated as an EV2.

The scale parameter provides characteristics on the stretchness of the distribution. The smaller the scale parameter is like in gages 113, 131 and 147, the more stretched the distribution is. The location indicates how shifted the distribution is to the right based on its value.

4.4. GEV Return Period for the same 100-year Discharge using LP3

The GEV results obtained using MatLab provided very interesting results for the different streamgages. In Table 9, below, it is possible to see the return periods for the main streamgages when the 100-year flood projection using LP3 in PeakFQ has the same value using the GEV distribution.

Table 9 - GEV return period for the same discharge of 100-year return period using LP3 in PeakFQ for the main streamgages.

Gage Map ID	# WY	GEV return period (for the same discharge from 100-year in LP3)	Relation 1/[WY]	Relation 1/[GEV return period]
Gage 113	13	10	0.077	0.1000
Gage 114	103	40	0.0097	0.0250
Gage 115	33	35	0.0303	0.0286
Gage 116	53	50	0.0189	0.0200
Gage 131	13	14	0.0769	0.0714
Gage 133	45	32	0.0222	0.0313
Gage 135	14	20	0.0714	0.0500
Gage 136	38	55	0.0263	0.0182
Gage 146	54	30	0.0185	0.0333
Gage 147	46	50	0.0217	0.0200

When considering gage 114 from the results provided, in table 9, one question can be brought up regarding the number of water years. If it has a length higher than 100 years it is naively expected that the projected flood considering 100-year return period will be comparable to the maxium peak flow from the entire record. However, this is not a correct assumption to make. When dealing with GEV (explained in detail in section 2.2), it is known that extreme events are given a greater weight. In addiction, even if considering an unchanging climate, this distribution considers that peaks will occur every 40 years.

From table 9, it is possible to see that the frequency obtained from using the period in WY and the frequency using the GEV return periods for the same discharge obtained using LP3 with 100-year, have approximate values. Table 10 shows the same things for the selected streamgages, considering only the ones with no trend.

Table 10 - GEV return period for the same discharge of 100-year return period using LP3 in PeakFQ for the selected by USGS streamgages.

Gage Map ID	# WY	GEV return period (for the same discharge from 100-year in LP3)	Relation 1/[WY]	Relation 1/[GEV return period]
Gage 773 – Until 1983	19	34	0.053	0.029
Gage 774 – Until 1981	11	12	0.091	0.083
Gage 775 – Until 1957	16	50	0.063	0.02
Gage 776 – Until 1977	14	63	0.071	0.016
Gage 777 – Until 1996	38	100	0.026	0.01
Gage 778 – Until 2020	68	40	0.015	0.025

The frequency values considering the water years available for each of the records and the GEV return periods are also very similar (table 10). This means that GEV distribution respects the real length of each data record. Gage 777 was the only gage that the 100-year LP3 matches the 100-year GEV the other streamgages seem to be very conservative regarding the return period for the same discharge.

4.5. GEV Expected Floods for the Return Periods from Table 7 with 95% CIs

The 100-year discharges with 95% CIs, for GEV, were calculated in MatLab and they are presented in Appendix II. However, the main focus are the same discharges, considering the 100-year LP3, found for GEV in different return periods. Table 11, below, provides the discharges for GEV with the 95% CIs for the main streamgages compared to the LP3 CIs.

Table 11 – Discharges and 95% CIs considering the same GEV return periods indicated.

Gage Map ID	WY	Return Period	Discharges for the return periods indicated using GEV	CIs 5% Lower	95% Upper
Gage 113	13	10	1,833	16.8	358,430
Gage 114	103	40	7,440	2,684.1	22,202
Gage 115	33	35	22,433	2,515	238,690
Gage 116	53	50	15,823	3,360	84,782
Gage 131	13	14	10,416	152.51	989,660
Gage 133	45	32	3,988	780.63	38,863
Gage 135	14	20	18,066	5,302.1	188,600
Gage 136	38	55	7,413	1,142.5	53,696
Gage 146	54	30	7,252	1,369.5	45,516
Gage 147	46	50	485.88	116.49	2,795

It is possible to see, from table 11, that gages 113, 131 and 135, with lengths of 13, 13 and 14 water years and return periods for GEV, when compared to the LP3 discharges, presented in the table, they present high values for the 95% upper CI. From that, it is possible to infer that, the shorter the period, the wider and least precise are the CIs.

26

From table 11, it was possible to see the CI limits for GEV. For gage 114, with more than 100 years of record, considering the 100-year GEV (Appendix II), there is huge projection because it is representing the fitted distribution instead of the historical record itself.

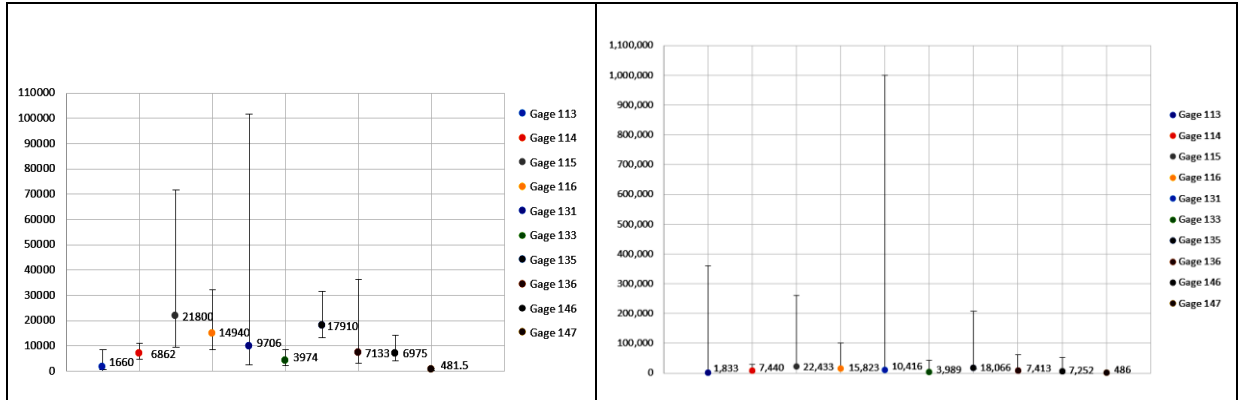Figure 16 compares the CI limits for LP3 (section 4.2: figure 15) and GEV.

Figure 16 - Comparisson between LP3 (in the left) and GEV (in the right) CIs.

It is possible to see that GEV CIs are wider than LP3 CIs. However, the widest CI ranges for GEV are for the streamgages with the shortest lengths of WY, showing that the accuracy is lower when considering short records, which is a good thing to consider. LP3 also shows wider CIs for some of the short length records but it is not the only reason for the wider CIs in these streamgages.

According to Stedinger, Vogel & Foufoula-Georgiou (1993), in log transformations, low outliers in the data record are given a great weight (that is why SGB was used in PeakFQ with B17B for obtaining the LP3 projected 100-year discharges). Considering LP3, when large values are the focus, some of the small values can be reported as zero if they fall below a certain threshold.

4.6. Comparisson between LP3 and GEV CDFs for each Streamgage.

Considering the parametrization obtained in section 4.3, the equation 6 from section 2.2 and the data records for each streamgage, the CDFs for GEV were displayed and compared to the CDFs for LP3, using equation 1 from section 2.2 and the LP3 results for each streamgage. The results for all the original streamgages with no trends detected can be evaluated and discussed below.

The figures below show the GEV and LP3 CDFs plotted to the same graphs for each of the streamgages with no trends in the data.
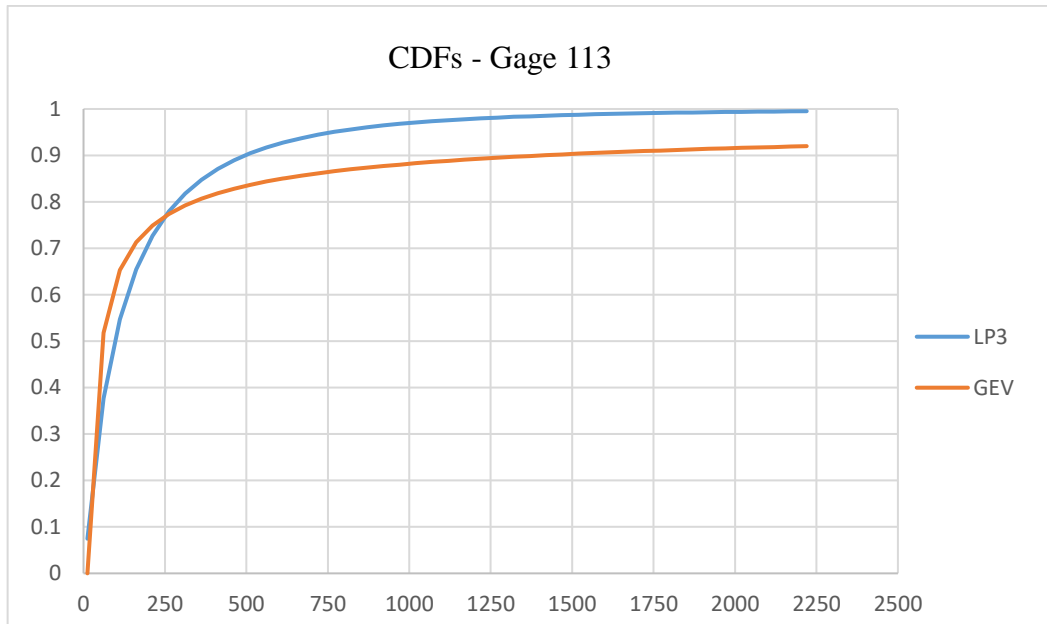
27

Figure 17 - GEV and LP3 CDFs for gage 113.

The curve above the other means that there is a larger percentage of data values in the selected interval of discharge values, in cfs, provided in the x-axis. GEV curve, below LP3 curve after the interval of 0 to 250 cfs (figure 17), shows that for higher discharge values, GEV is more conservative than LP3, presenting a lower percentage of these values in the ending part of its CDF curve and meaning that it considers a smaller return period. It also demonstrates how GEV takes the length of the records a lot in consideration.
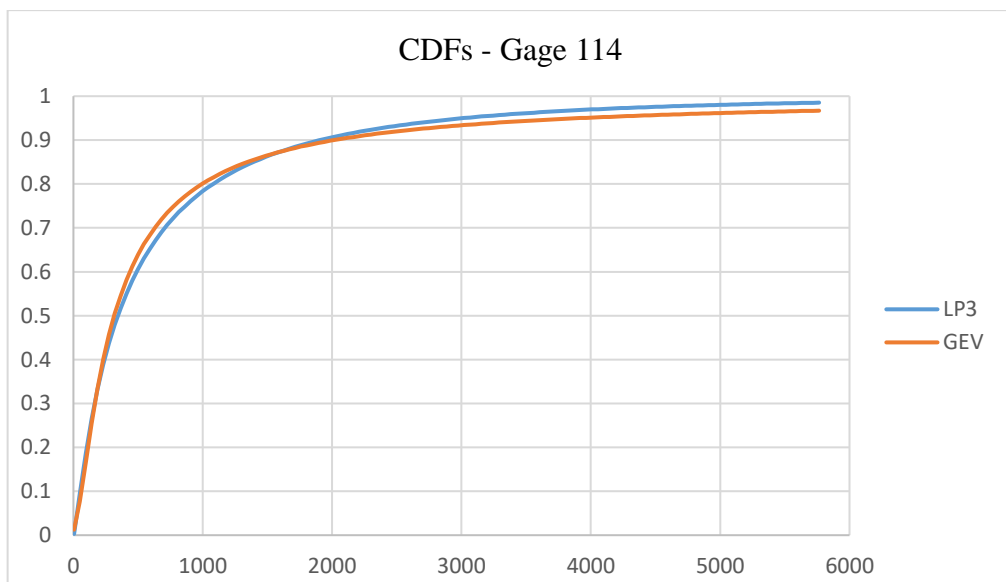
Figure 18 - GEV and LP3 CDFs for gage 114.

For the streamgage 114, the results for both LP3 and GEV CDFs were really mainly very similar. Until the interval of 0 to 1000 cfs, LP3 shows to be more conservative than GEV and, then, for the next higher discharge values, GEV presents a smaller return period.
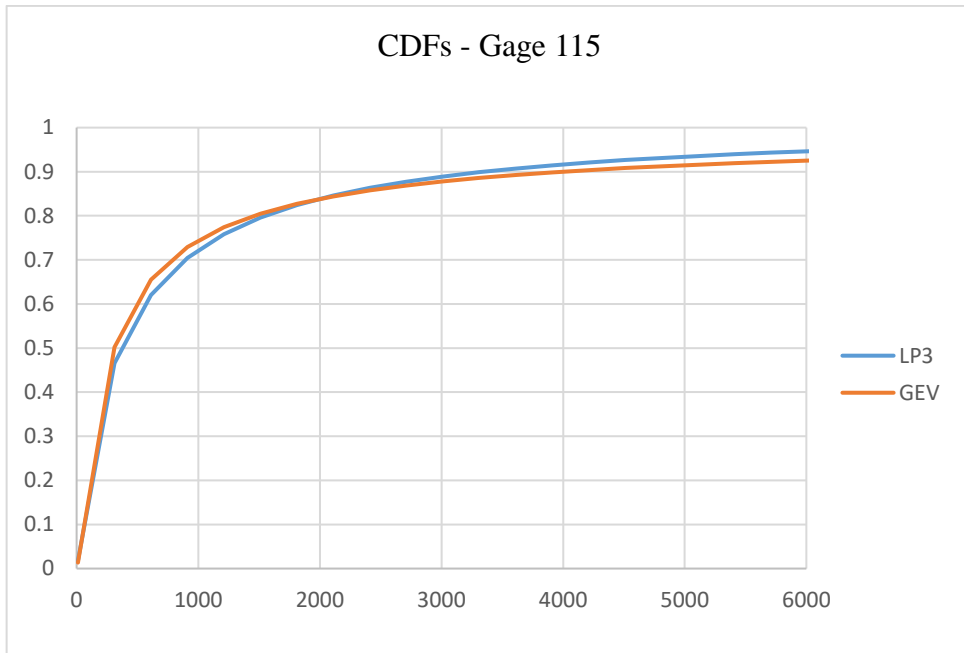
28

Figure 19 - GEV and LP3 CDFs for gage 115.

Considering gage 115 (figure 19), until the interval of 0 to 2000 cfs, LP3 shows to be more conservative than GEV and, then, for the next higher discharge values, GEV presents a smaller return period.
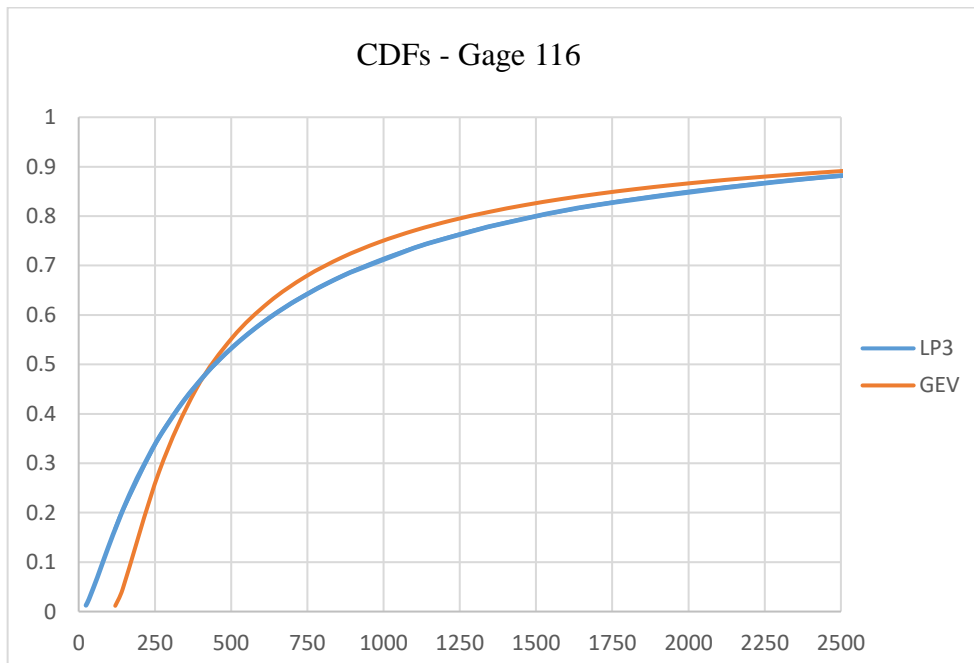

Figure 20 - GEV and LP3 CDFs for gage 116.

For gage 116 (figure 20), LP3 shows to be more conservative for all cases after the interval of 0 to 375 cfs.
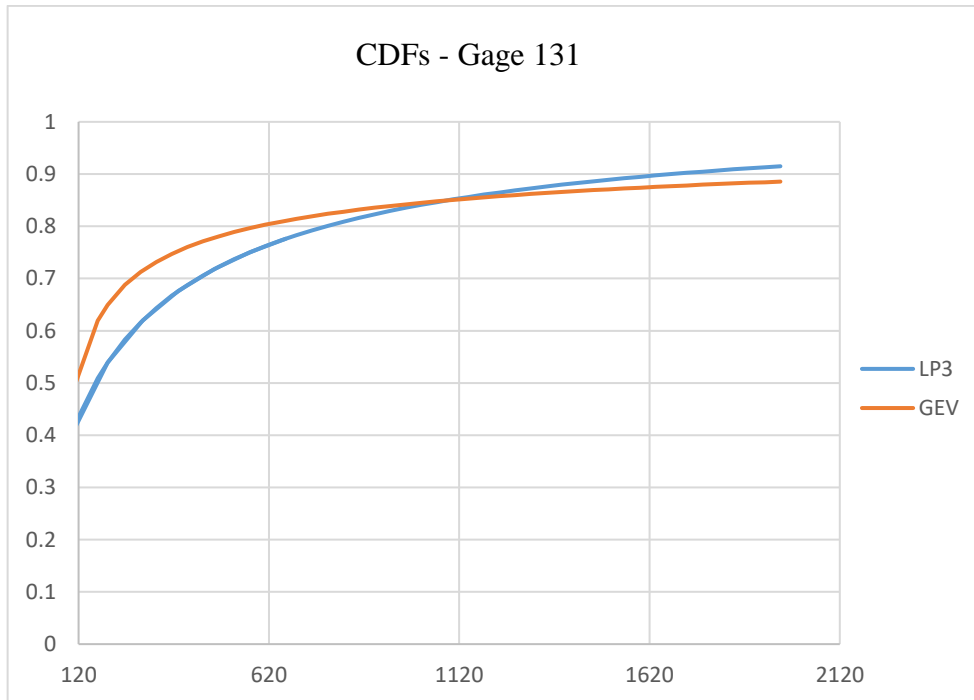
Figure 21 - GEV and LP3 CDFs for gage 131.

Considering gage 131 (figure 21), LP3 shows to be more conservative until 1120 cfs. Then, for the next higher discharge values, GEV presents a smaller return period, showing to be more conservative.
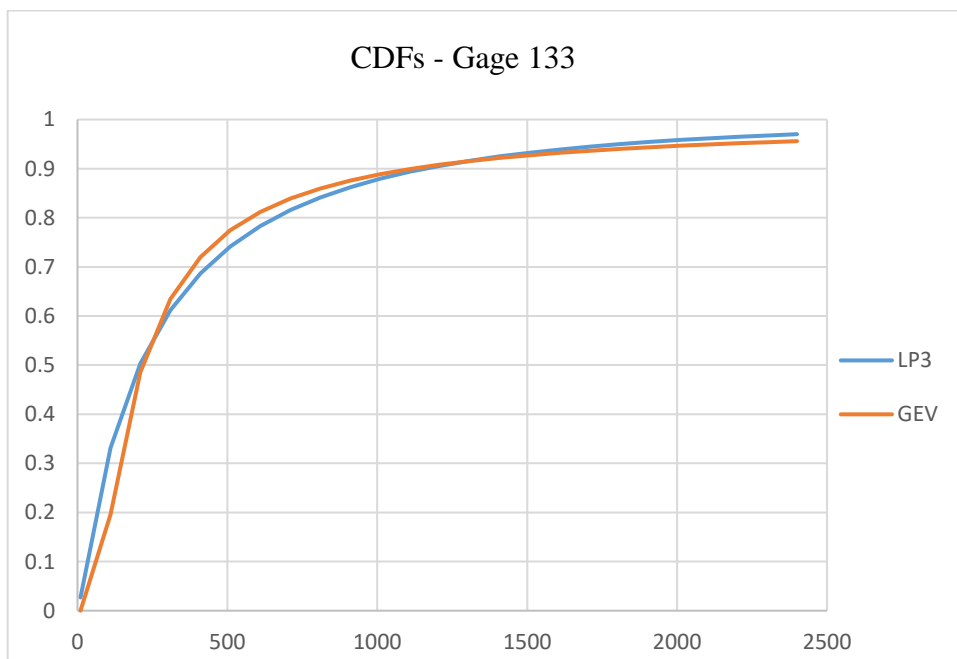

Figure 22 - GEV and LP3 CDFs for gage 133.

Considering gage 133 (figure 22), LP3 and GEV alternate being the best of fitness for one interval or another.
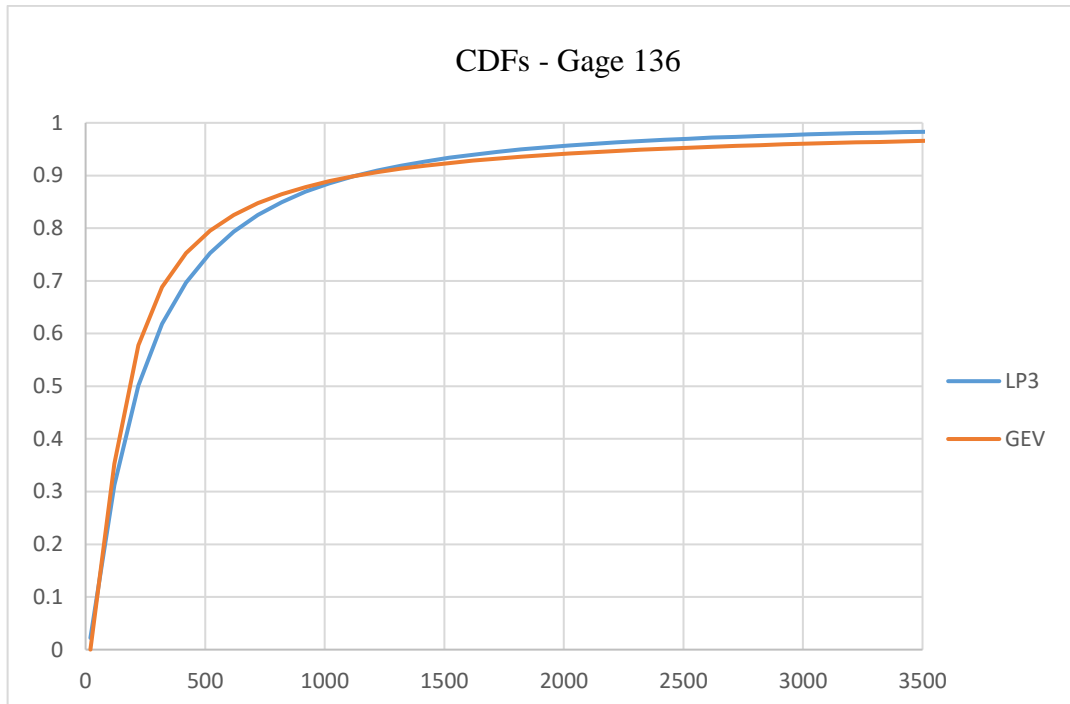
Figure 23 - GEV and LP3 CDFs for gage 136.

For gage 136, it is possible to see that LP3 is more conservative until around 1000 cfs and then GEV gets more conservative for higher discharges.
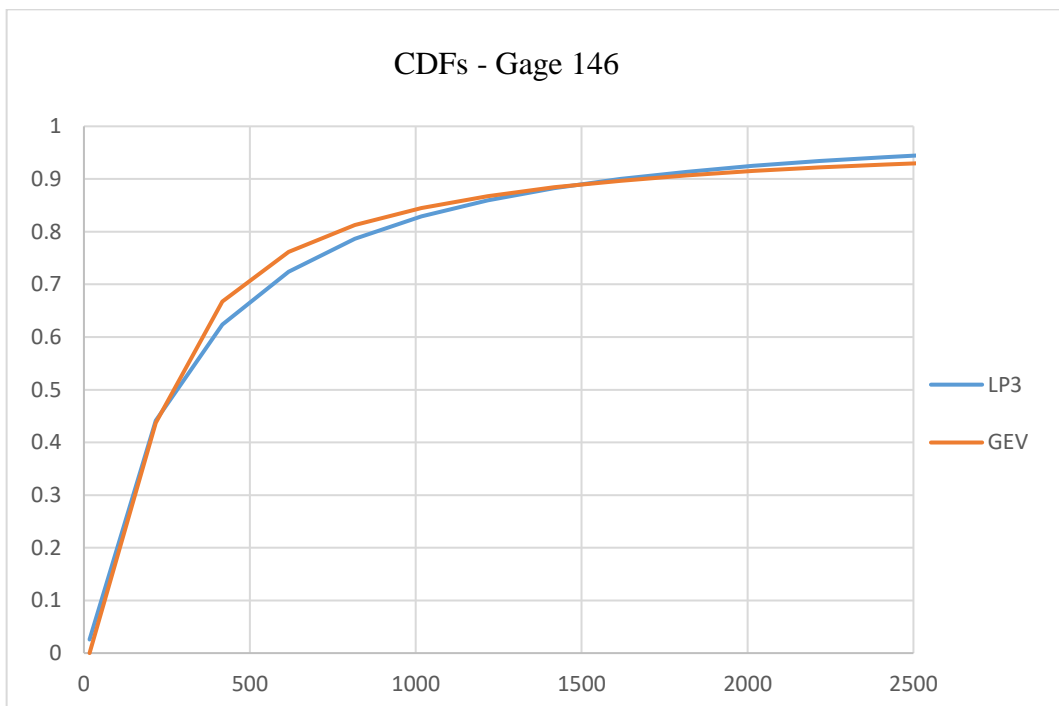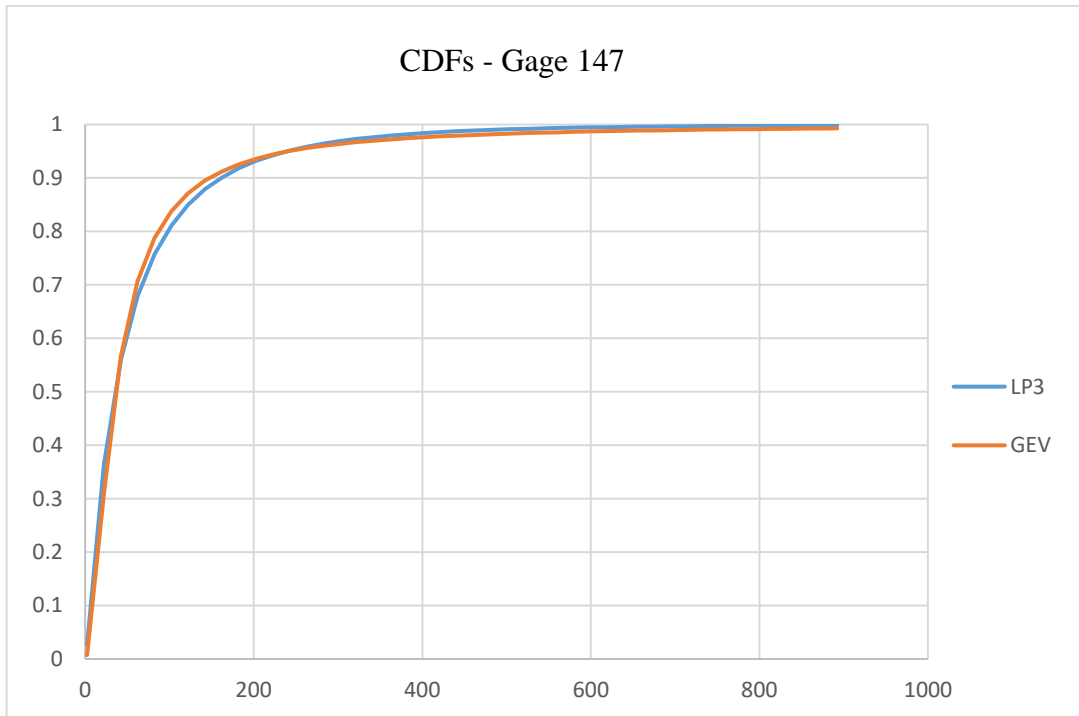

Figure 24 - GEV and LP3 CDFs for gage 146.

For gage 146, it is possible to see that LP3 and GEV behave similarly until before 250 cfs. LP3 is more conservative until around 1500 cfs and then GEV gets more conservative for the highest discharges.

Figure 25 - GEV and LP3 CDFs for gage 147.

For the streamgage 147, the results for both LP3 and GEV CDFs are mainly identical. It is possible to see also, from table 6, table 11 and Appendix II that the GEV discharge and CI results are not that different from LP3 results considering the differences portraid in the results for the other main streamgages.

For the USGS selected streamgages, the gages have their CDFs evaluated when the original updated records had no prior trends. Taking only gage 778 as an example (figure 25) to show both CDFs with the data plotted, it is possible to see how the data follows each of the CDF curves and how proximate it is to one or the other curve.
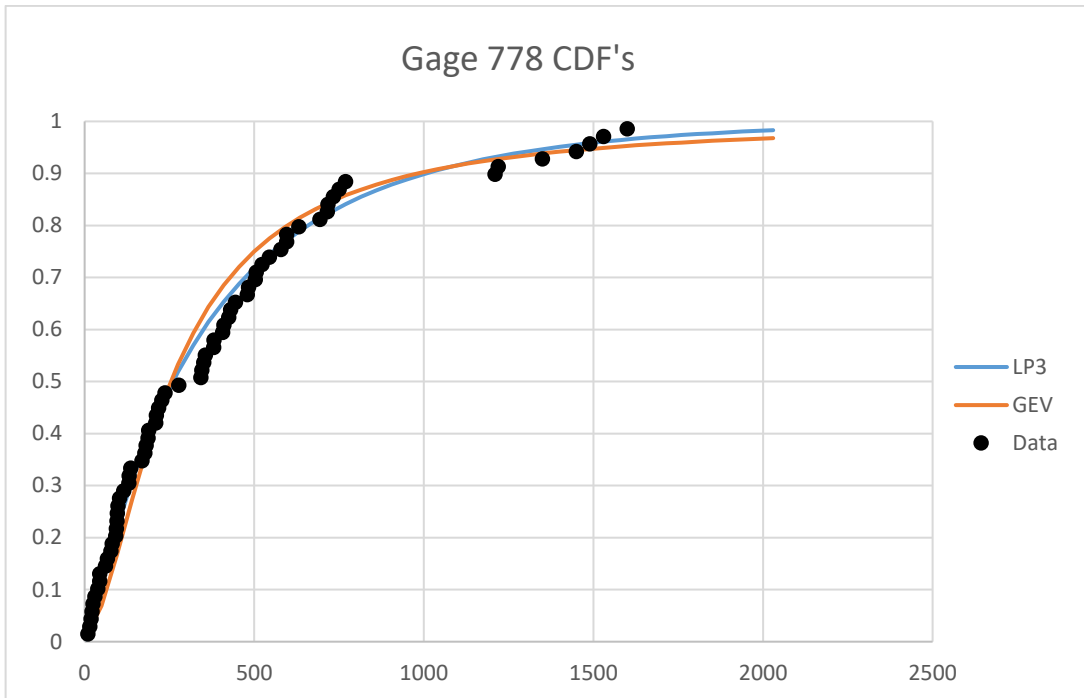
Figure 26 - GEV and LP3 CDFs with the historical data plotted for gage 778.

For this same gage, a plot with just the GEV curve and the historical data stair graph was done using MatLab. The result can be seen in figure 26.
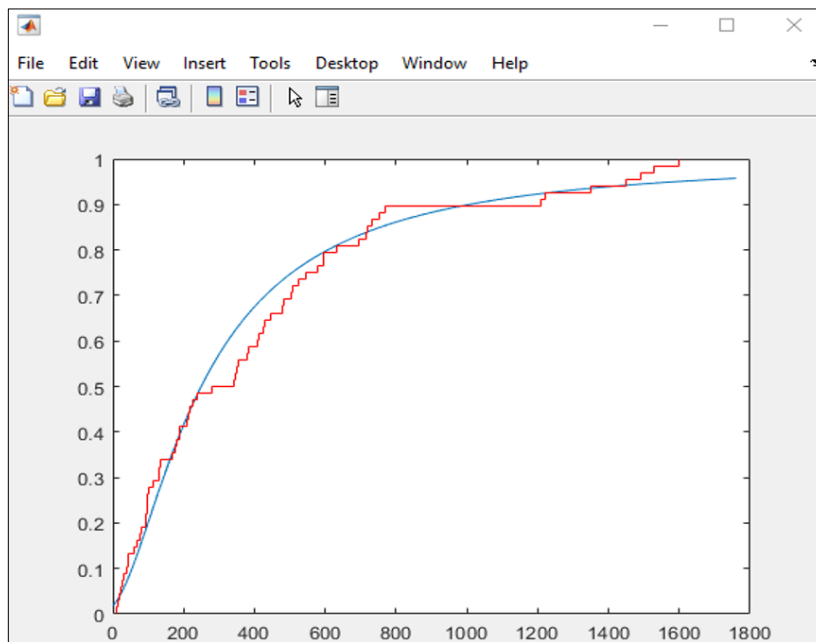

Figure 27 - Example of GEV CDF curve plotted together with historical data as a stair graph for gage 778.

It is possible to see that the GEV curve fits and represents the data records really well.

5. CONCLUSION

From the results obtained in this thesis, it can be inferred that for any of the data with short records, the confidence intervals (CIs) for either LP3 or GEV are wider (with very high upper values). The results obtained for GEV were directly affected by the length of the data. When the length of the record is short, it is not accurate to use a projection of 100-year return period to represent future projections but, instead, the lenght available.

By analizing USGS report and streamgages data, it was possible to see that many of the streamgages presented short-period data records. Also, some analyzed streamgages had missing peaks in some of the water years that could potentially affect in the analysis. For some gages, USGS provided a perception threshold discharge for missing peaks. For the streamgages analyzed in this thesis, no perception threshold was detected in PeakFQ. The report presented the expected moments algorithm (EMA) method as an alternative that could incorporate censored data and interval peak-discharge data into the analysis but the CIs obtained for the examples used in section 4.2, shows B17B can provide more accurate data when considering streamgages with no missing records, no perception threshold and no additional historical information. Therefore, the document comes up with alternatives for when data does not fit LP3 well because of low outliers or the previous commented reasons with the data itself.

Even with the results showing that GEV provides wider than LP3 95% CIs for the projected discharges, when considering the analyzes, regarding the CDFs for the majority of the streamgages analyzed in this project, <u>GEV proves to be the most conservative method</u>, with smaller return periods, between the two distributions.

Stakhiv (2011), in his work, shows a smaller return period for the same discharge when using GEV instead of LP3 and finishes his work by stating how necessary it is to consider the climate change's assumption to change the evaluation procedures, such as replacing LP3 and applying GEV probability distribution for flood frequency analysis (FFA) for water infrastructure.
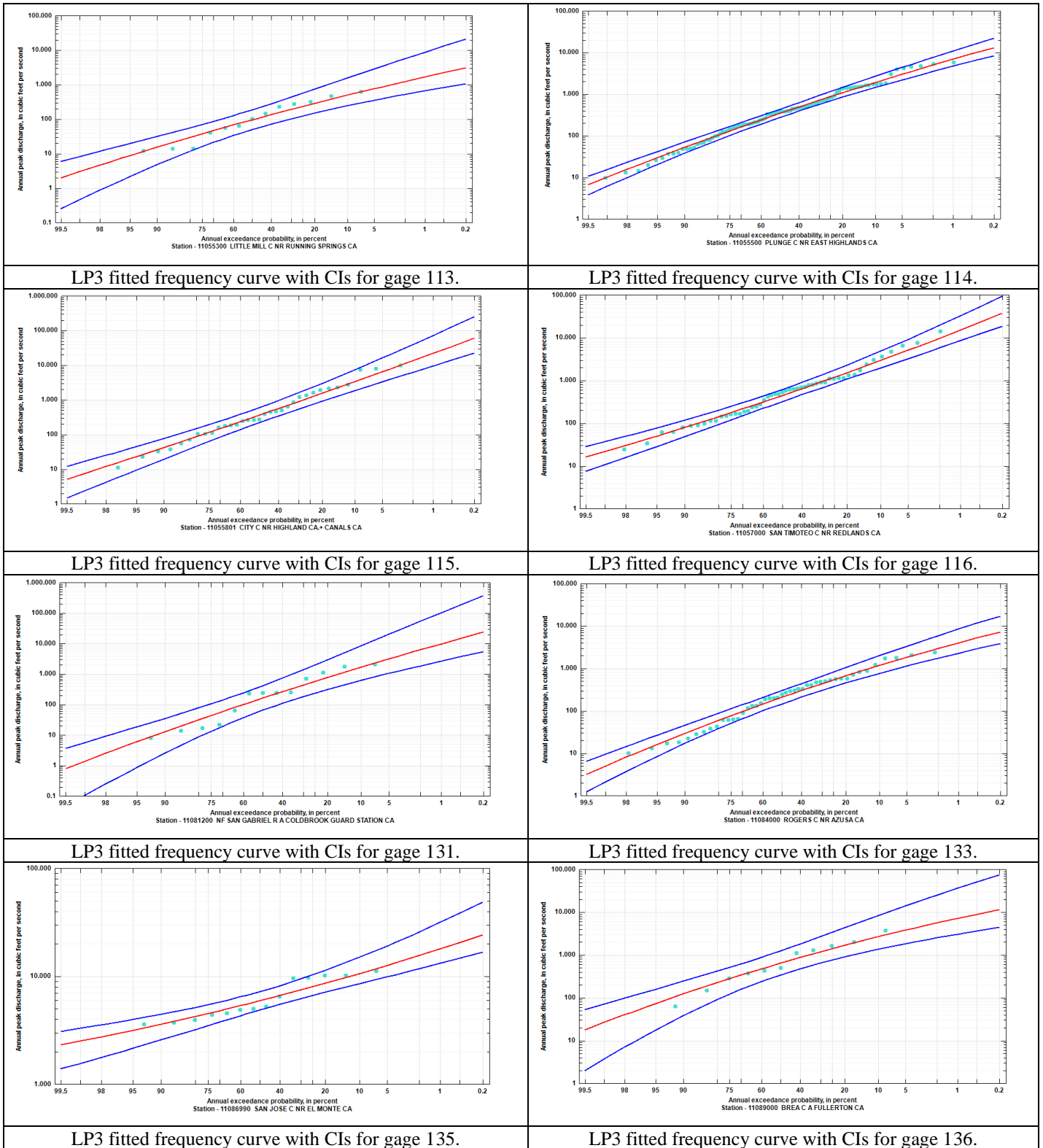
Considering a nonstationary climate, flooding issues happening more frequently and knowing the GEV itself gives more weight to extreme events, it is necessary to analyze, together with LP3 (the most conventional statistical method for FFA in US), which method seems to be the best for each scenario instead of using the exact same distribution for all cases.
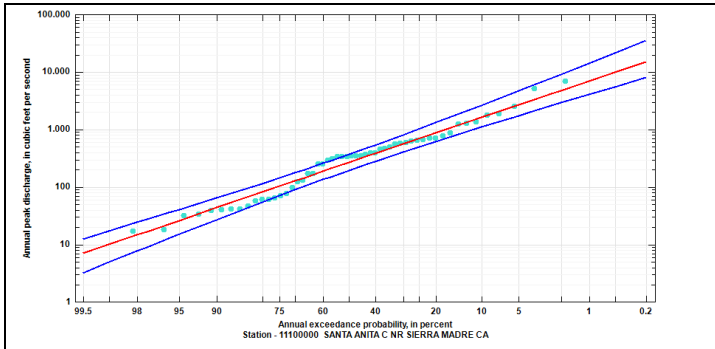
6. REFERENCES

1. Clarke, A. O. (1996). Estimating probable maximum floods in the Upper Santa Ana Basin, Southern California, from stream boulder size. Environmental & Engineering Geoscience, 2(2), 165-182.

2. England, Jr, J. F., & Cohn, T. A. (2007). Scientific and practical considerations related to revising Bulletin 17B: The case for improved treatment of historical information and low outliers. In World Environmental and Water Resources Congress 2007: Restoring Our Natural Habitat (pp. 1-9).

3. Farooq, M., Shafique, M., & Khattak, M. S. (2018). Flood frequency analysis of river swat using Log Pearson type 3, Generalized Extreme Value, Normal, and Gumbel Max distribution methods. Arabian Journal of Geosciences, 11(9), 1-10.

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

5. Guinn, J. M. (1890). Exceptional years: a history of California floods and drought. Historical Society of Southern California, Los Angeles (1890), 1(5), 33-39.

6. Hajani, E., & Rahman, A. (2018). Design rainfall estimation: comparison between GEV and LP3 distributions and at-site and regional estimates. Natural Hazards, 93(1), 67-88.

7. Mays, L. W. (2010). Water resources engineering. John Wiley & Sons.

8. Millington, N., Das, S., & Simonovic, S. P. (2011). The comparison of GEV, log-Pearson type 3 and Gumbel distributions in the Upper Thames River watershed under global climate models.

9. Orsi, J. (2004). *Hazardous metropolis: flooding and urban ecology in Los Angeles*. Univ of California Press.

10. PeakFQ. (2019). Available in: <https://water.usgs.gov/software/PeakFQ/>. Page last modified: Monday, 30-Apr-2014 17:34:00 EDT.

11. Parrett, Charles, Veilleux, Andrea, Stedinger, J.R., Barth, N.A., Knifong, D.L., and Ferris, J.C. (2011). Regional skew for California, and flood frequency for selected sites in the Sacramento–San Joaquin River Basin, based on data through water year 2006.

12. Rohmer, J., Gehl, P., Marcilhac-Fradin, M., Guigueno, Y., Rahni, N., & Clément, J. (2020). Non-stationary extreme value analysis applied to seismic fragility assessment for nuclear safety analysis. Natural Hazards and Earth System Sciences, 20(5), 1267-1285.

13. Stakhiv, Eugene Z. "Pragmatic approaches for water management under climate change uncertainty 1." *JAWRA Journal of the American Water Resources Association* 47, no. 6 (2011): 1183-1196.

14. Stedinger, J. R., Vogel, R. M., & Foufoula-Georgiou, E. (1993). Frequency analysis of extreme events. Pages 18.1–18.66 in DR Maidment, editor. Handbook of Hydrology.

15. The MathWorks, Inc. (2022). Modelling Data with the Generalized Extreme Value Distribution. Available in: <https://www.mathworks.com/help/stats/modelling-data-with-the-generalized-extreme-value-distribution.html>. Acessed on June 8, 2022.

16. Triola, M. F., Goodman, W. M., Law, R., & Labute, G. (2006). *Elementary statistics* (p. 794). 11th Ediction. Reading: Pearson/Addison-Wesley.

17. Triola, Mario F., et al. (2018). *Elementary statistics*. 13th Ediction. Reading: Pearson.

18. USGS Water Data for USA (2022). National Water Information System: Web Interface. Available in: <https://nwis.waterdata.usgs.gov/nwis?>. Page Last Modified: 2022/07/01.

19. Vogel, R. M., & Wilson, I. (1996). Probability distribution of annual maximum, mean, and minimum streamflows in the United States. *Journal of hydrologic Engineering*, *1*(2), 69-76.

APPENDIX I – Fitted frequency curve (in red) and CI curves (in blue) for annual peak discharge (cfs) by AEP (%) with gaged peak discharges plotted as well (in green).
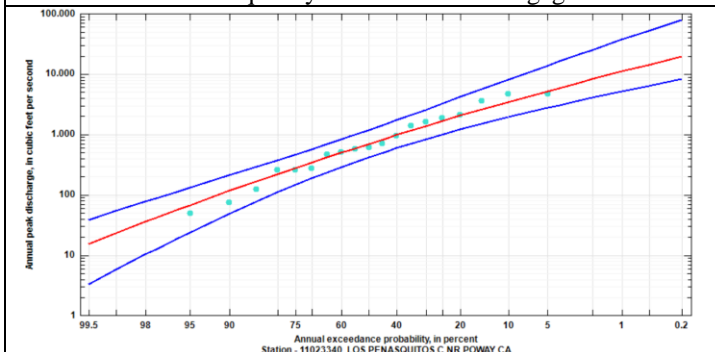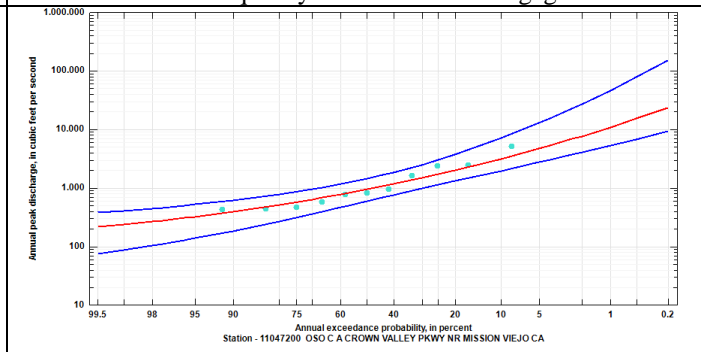

LP3 fitted frequency curve with CIs for gage 113.	LP3 fitted frequency curve with CIs for gage 114.
LP3 fitted frequency curve with CIs for gage 115.	LP3 fitted frequency curve with CIs for gage 116.
LP3 fitted frequency curve with CIs for gage 131.	LP3 fitted frequency curve with CIs for gage 133.
LP3 fitted frequency curve with CIs for gage 135.	LP3 fitted frequency curve with CIs for gage 136.
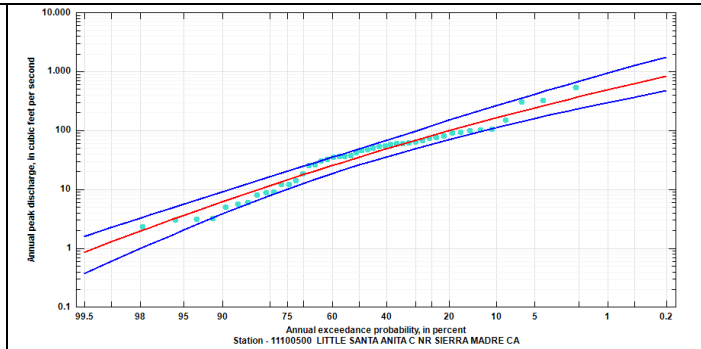
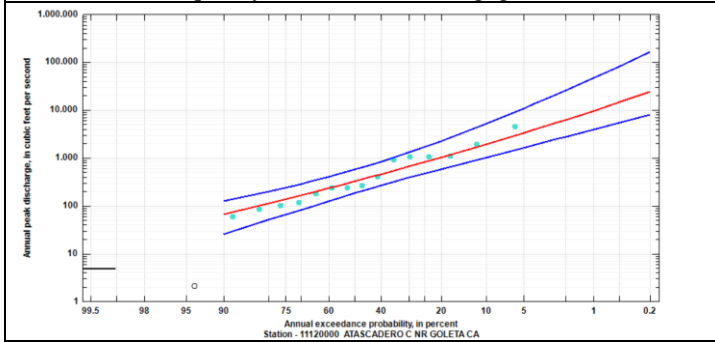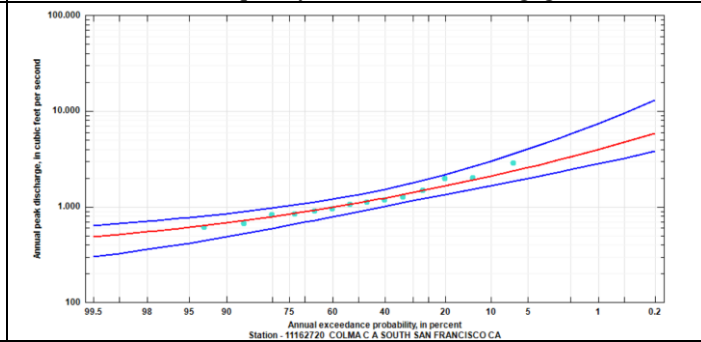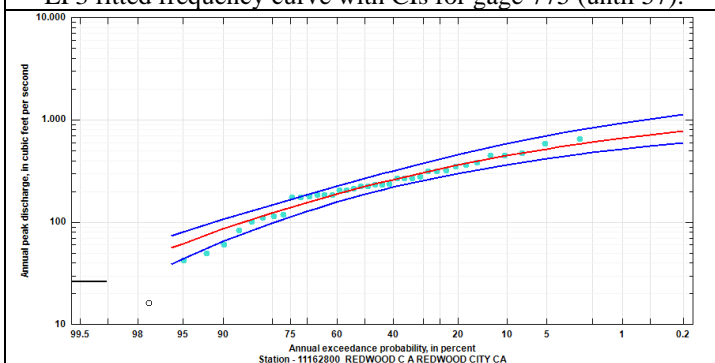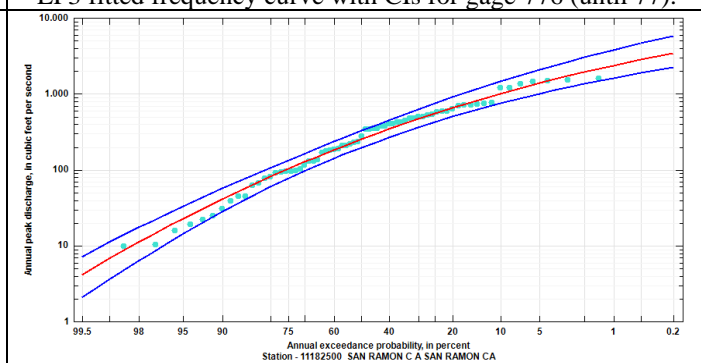|	
LP3 fitted frequency curve with CIs for gage 146.	LP3 fitted frequency curve with CIs for gage 147.
LP3 fitted frequency curve with CIs for gage 773 (until 83).	LP3 fitted frequency curve with CIs for gage 774.
LP3 fitted frequency curve with CIs for gage 775 (until 57).	LP3 fitted frequency curve with CIs for gage 776 (until 77).
LP3 fitted frequency curve with CIs for gage 777.	LP3 fitted frequency curve with CIs for gage 778.

APPENDIX II – Discharges for 100-year return period using GEV for the main streamgages.

| Gage Map ID | # W.Y. | Matlab (0.99) GEV | 95% Confidence Intervals (CIs): for 100-yr discharges | | Errors up and down the discharge obtained for 100-year GEV. | |
		100-yr discharge	Lower	Upper	Error Up	Error Down
gage 113	13	193,650.00	38.54	4,008,600,000.00	4,008,406,350.00	193,611.46
gage 114	103	17,156.00	4,809.10	66,890.00	49,734.00	12,346.90
gage 115	33	91,744.00	5,635.30	1,823,900.00	1,732,156.00	86,108.70
gage 116	53	38,358.00	5,841.90	291,430.00	253,072.00	32,516.10
gage 131	13	672,520.00	773.80	970,220,000.00	969,547,480.00	671,746.20
gage 133	45	11,506.00	1,335.70	226,060.00	214,554.00	10,170.30
gage 135	14	46,714.00	5,802.10	2,090,200.00	2,043,486.00	40,911.90
gage 136	38	13,612.00	1,548.00	126,400.00	112,788.00	12,064.00
gage 146	54	27,742.00	2,923.10	323,990.00	296,248.00	24,818.90
gage 147	46	772.97	141.40	5,741.10	4,968.13	631.57