Precipitation Analysis using Extreme Value Theory

Matthew Self, ASA, MAAA

March 2023

Overview

This paper sets out to apply the core concepts of Extreme Value Theory to the precipitation history of Lambert Airport in St. Louis, Missouri (STL). In July 2022, 8.6 inches of rain was recorded in one day, breaking the previous record of 5.6 inches. Previously, the Society of Actuaries Research Institute (SOA) reported on this extreme precipitation.

The analysis described in this paper can easily be applied to other time series: either a different station or a different weather variable, such as temperature. By following the methodology below and working through the accompanying workbook, any actuary can begin analysis of extreme weather observations.

First, we'll describe the source of the precipitation observations and the other data available, as well as the initial summary statistics of the STL observations. Next, we'll provide an overview of the Generalized Extreme Value (GEV) distribution and the methods for fitting data to the distribution; additionally, we'll illustrate how well the fitted distribution compares to the observations. Then, we'll walk through two methodologies for confidence interval construction for the fitted parameters. Lastly, we discuss the sensitivity of estimations of probabilities and return periods to the parameters and the challenges that such sensitivity presents when communicating analysis of extreme events.

GHCN and Precipitation Data

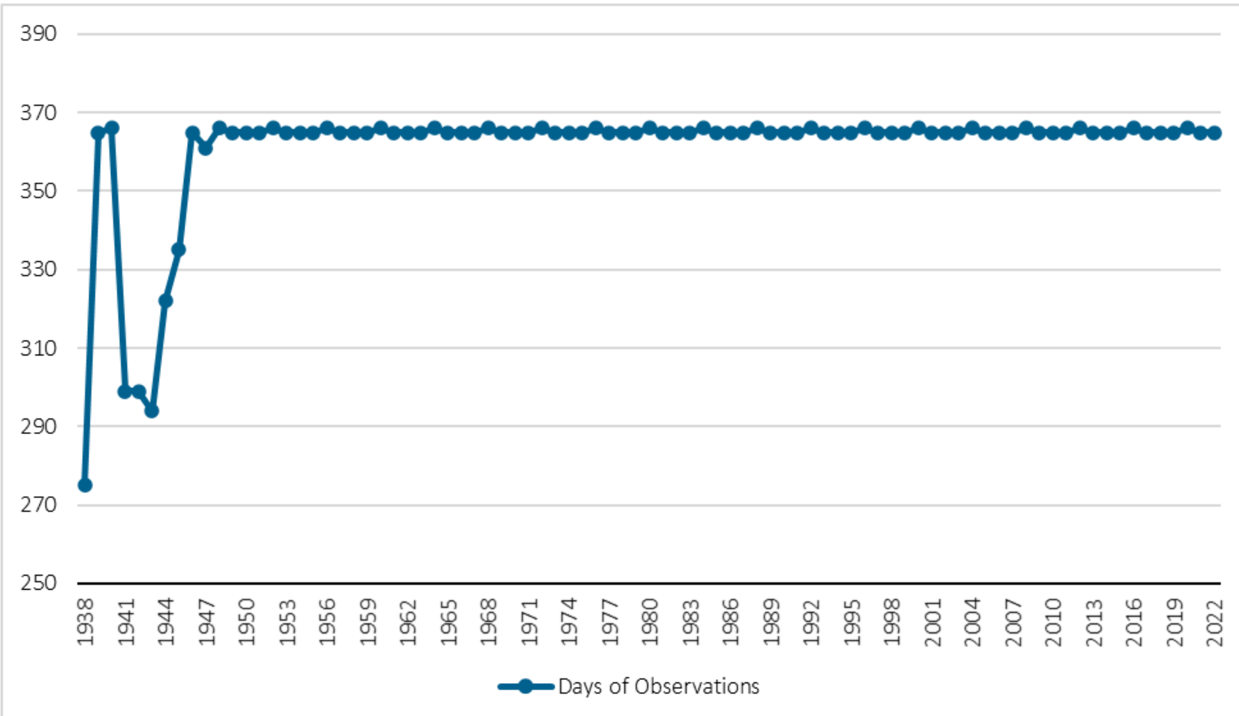
The Global Historical Climatology Network comprises over 100,000 stations around the world and reports daily variables such as maximum temperature, minimum temperature, total precipitation, and more. Data can be retrieved through FTP or HTTP requests. For this analysis, we retrieved all available precipitation observations at the

St. Louis Airport station, beginning April 1, 1938. Figure 1 below shows the complete precipitation history of the station.

The observation data set has exceptional completeness starting in 1946, with no year having fewer than 361 days of observations. The period 1938-1945 has more missing days, with 1943 only having 294 observations. Despite this incompleteness in earlier years, the dataset still provides a good basis for analysis.

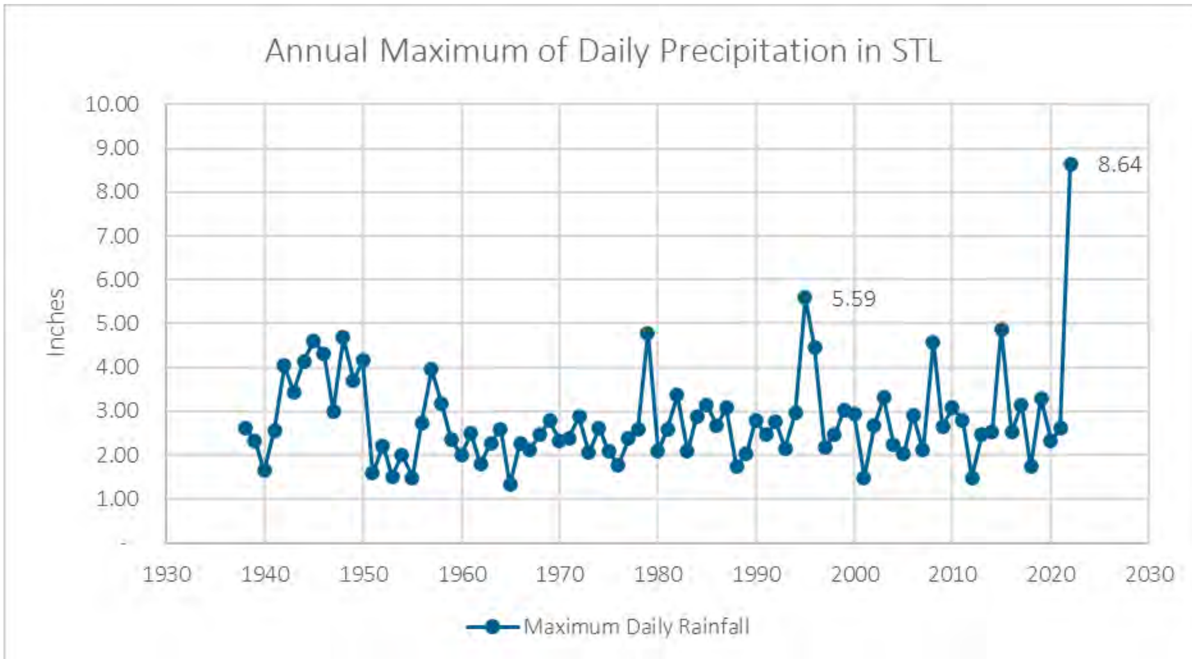
Figure 1

DAYS OF PRECIPITATION OBSERVATIONS BY YEAR AT STL GHCN STATION



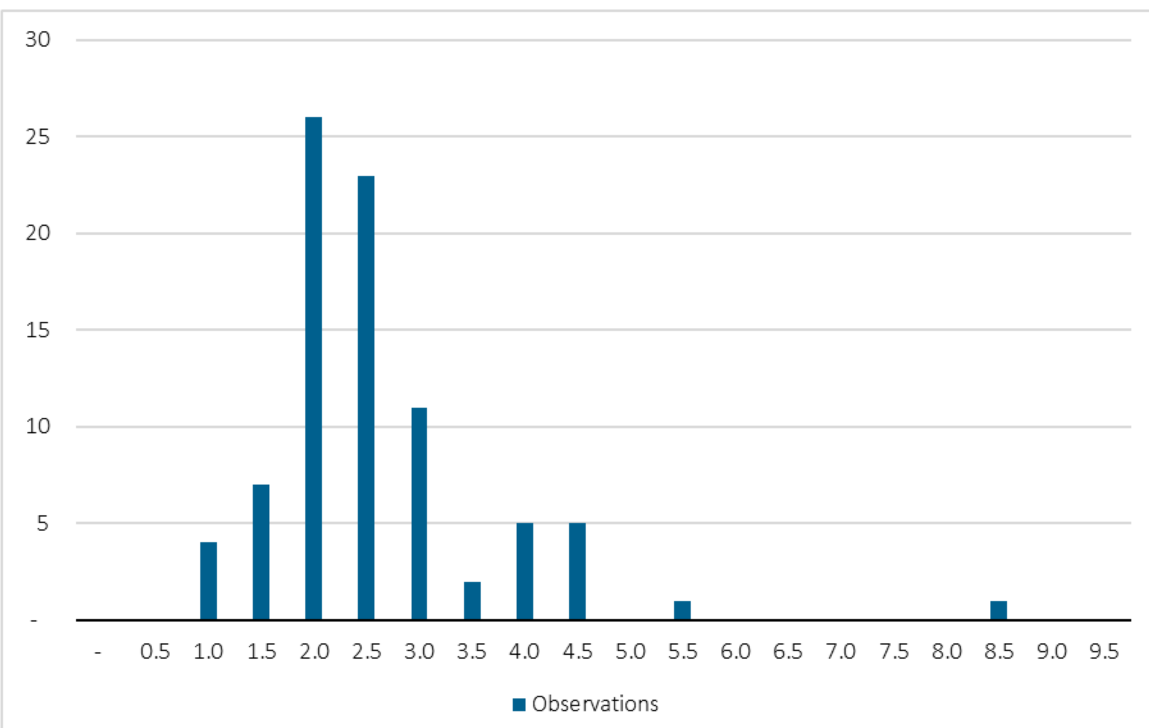
From the daily observations, we calculated the maximum daily rainfall for each year; this statistic is referred to as a "block maxima", where one block is equal to one year. The selection of one year as our block length is natural and provides high confidence in the convergence of our sample distribution to the GEV distribution.

Figure 2
ANNUAL MAXIMUM OF DAILY PRECIPITATION AT STL GHCN STATION



After aggregating the maxima observations into half-inch buckets, we visually inspect the frequency graph of the observations. Clearly, these observations would not fit a Normal distribution and are better-suited to a heavy-tailed distribution.

Figure 3
HISTOGRAM OF ANNUAL MAXIMA OF DAILY PRECIPITATION AT STL GHCN STATION



The Generalized Extreme Value Distribution

The Generalized Extreme Value (GEV) distribution is a family of distributions which are the limit distributions of the maxima of a sequence of random variables, e.g., the maximum daily rainfall in each year. The distribution takes three parameters: location, scale, and shape; these parameters roughly approximate the median, standard deviation, and "fatness" of the tail, respectively. As the concern of this analysis is an extreme event, we will put more focus on evaluation of the shape parameter.

The Cumulative Distribution Function (CDF) of the GEV is given by:

$$G(x;\mu,\sigma,\xi) \; = \; \exp\left\{-\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]_{+}^{-1/\xi}\right\},$$

where *mu* is the location parameter and can be any real number, *sigma* is the scale parameter and must be greater than zero, and *xi* is the shape parameter and can be any real number.

To fit the observations to the GEV distribution, we used the Maximum Likelihood Method to maximize the log-likelihood function, shown below. The log-likelihood function is defined as a logarithmic transformation of the probability density function and is used in parameter estimation because logarithms are strictly increasing functions.

Rather than a time-consuming brute force method to find optimal solutions, we implemented a gradient descent algorithm in Visual Basic for Applications (VBA). For a given set of parameters, we calculate the log-likelihood function if each parameter is 1% greater or lesser than its current value, and we then select the altered parameters that produce the greatest likelihood. This algorithm relies on the fact that the log-likelihood function is monotonically increasing and selects the parameters with the greatest gradient, i.e., slope.

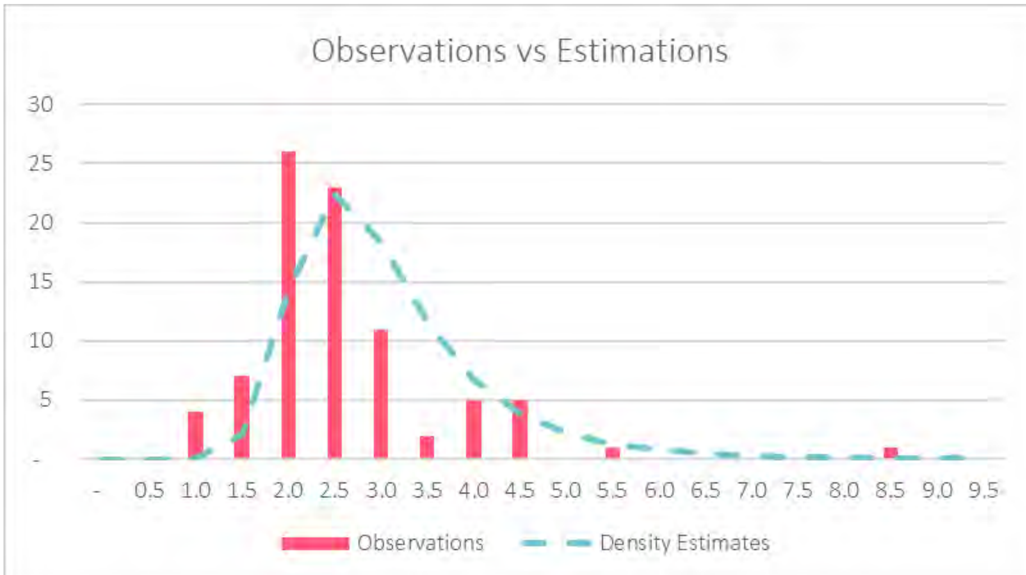
$$LL = -n\ln\sigma - \left(1+\frac{1}{\xi}\right)\sum_{i=1}^{n}\ln y_i - \sum_{i=1}^{n} y_i^{-1/\xi}$$

$$z_i = (x_i - \mu)/\sigma \text{ and } y_i = 1 + \xi z_i.$$

The fitted parameters for our precipitation data and the resulting estimates were:

- Location: 2.319
- Scale: 0.681
- Shape: 0.122
- Log Likelihood: -107.5475

Figure 4
COMPARISON OF OBSERVED ANNUAL MAXIMA AND EXPECTED ANNUAL MAXIMA



As shown in Figure 4, the distribution appears to fit the data well, particularly for observations greater than or equal to 4 inches: the area of extremes and of our concern. The median and standard deviation of our maxima data set are 2.59 and 1.09, respectively. Therefore, the fitted distribution is set further to the left and has a higher peak than we might have expected, if we had used the median and standard deviation as our initial estimation.

With the fitted parameters, we estimate the probability of observing a value greater than the 2022 maximum of 8.64 inches to be 0.2019% with approximately a 500-year return period (the reciprocal of the probability). These inferences confirm the extreme nature of the observed precipitation event.

Confidence Interval Construction

Given the nature and frequency of extreme events, it is essential to construct and report confidence intervals for the estimated parameters. For this analysis, we carried out two methodologies for estimating the standard error of each parameter estimate: Monte Carlo simulations and Bootstrap Resampling.

MONTE CARLO SIMULATIONS

The Monte Carlo method produces 85 observations from the fitted GEV parameters for 1000 scenarios, representing 85 years of Annual Maxima observations. For each scenario, we run the GEV parameter estimation algorithm maximizing the log-likelihood function. After 1000 scenarios, we calculate the Mean and Standard Deviation of each parameter's results and use the Standard Deviation to construct 95% confidence intervals for each parameter.

The traditional method of constructing confidence intervals with the standard error and Z-score tables is suitable for the Location and Scale parameters, for which the parameter estimates appear normally distributed, shown in Figures 5 and 6, respectively. However, the Shape parameter does not fit a Normal distribution and is itself heavy-tailed, shown in Figure7. Therefore, a more direct method is appropriate for constructing the confidence interval of the Shape parameter: using the 2.5th and 97.5th percentiles from the 1000 estimated shape parameters. This direct method produces confidence intervals for the Location and Shape parameters that differ from the traditional confidence interval by 0.02 and 0.01, respectively, thereby confirming the normality assumption.

Figure 5
FREQUENCY DISTRIBUTION OF LOCATION PARAMETER ESTIMATES

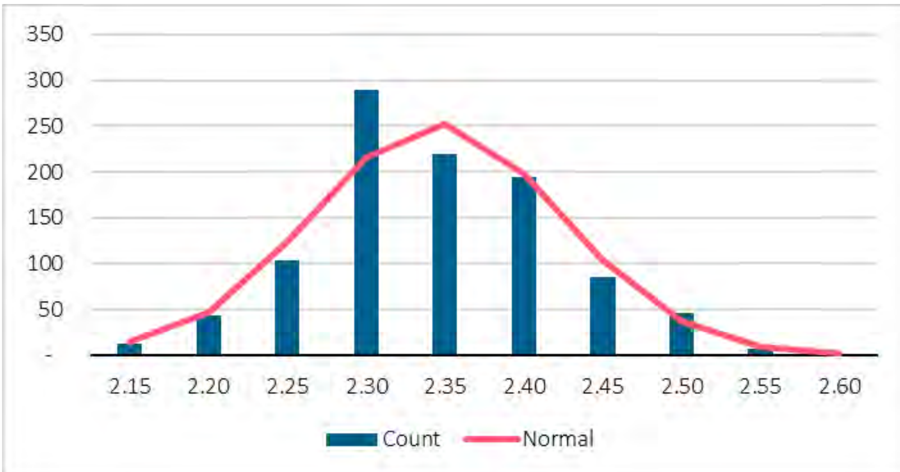


Figure 6
FREQUENCY DISTRIBUTION OF SCALE PARAMETER ESTIMATES

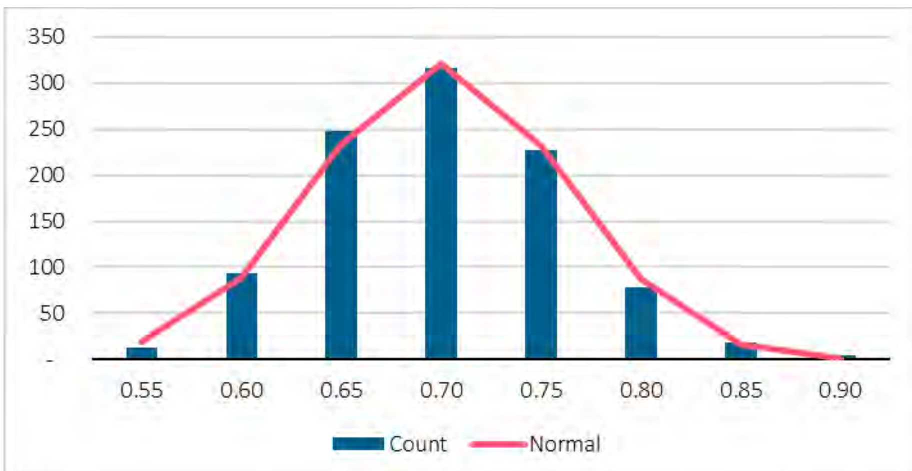
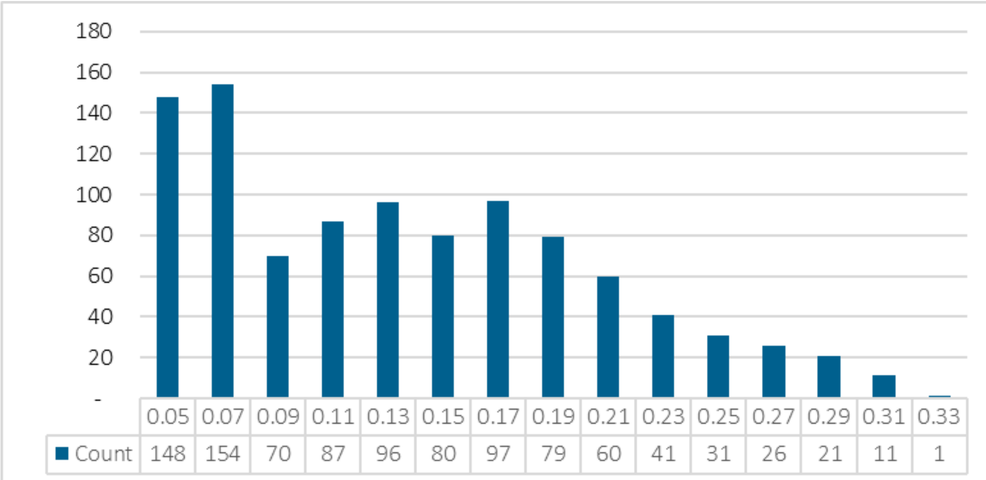


Figure 7

FREQUENCY DISTRIBUTION OF SHAPE PARAMETER ESTIMATES



From the simulations, we have the following parameter estimates:

Table 1

CONFIDENCE INTERVALS FOR GEV PARAMETERS FROM MONTE CARLO SIMULATIONS

	Best Estimate	Mean	Standard Error	CI Bottom	CI Top	2.5th Percentile	97.5th Percentile
Location	2.319	2.320	0.078	2.168	2.472	2.1835	2.4867
Scale	0.681	0.675	0.060	0.557	0.793	0.5626	0.7987
Shape	0.122	0.127	0.068	(0.006)	0.260	0.0470	0.2791

Visually, we can compare the resulting estimated observations to the actual maxima observations. As expected, the 2.5th percentile has a “smaller” tail and assigns smaller probabilities to extreme events, shown in Figure 9. Conversely, the larger Shape parameter of the 97.5th percentile assigns higher probabilities to the extreme events.

Figure 8

DISTRIBUTION ESTIMATES FROM MONTE CARLO SIMULATIONS

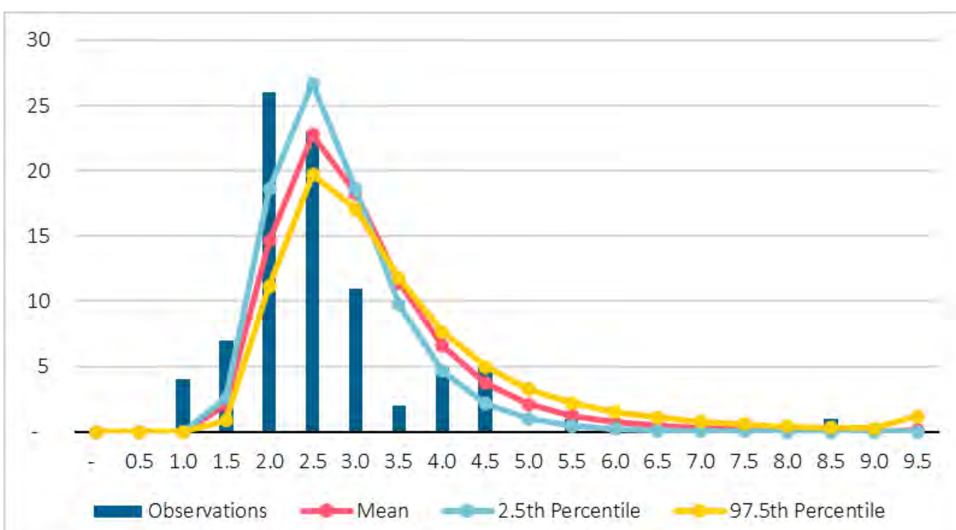
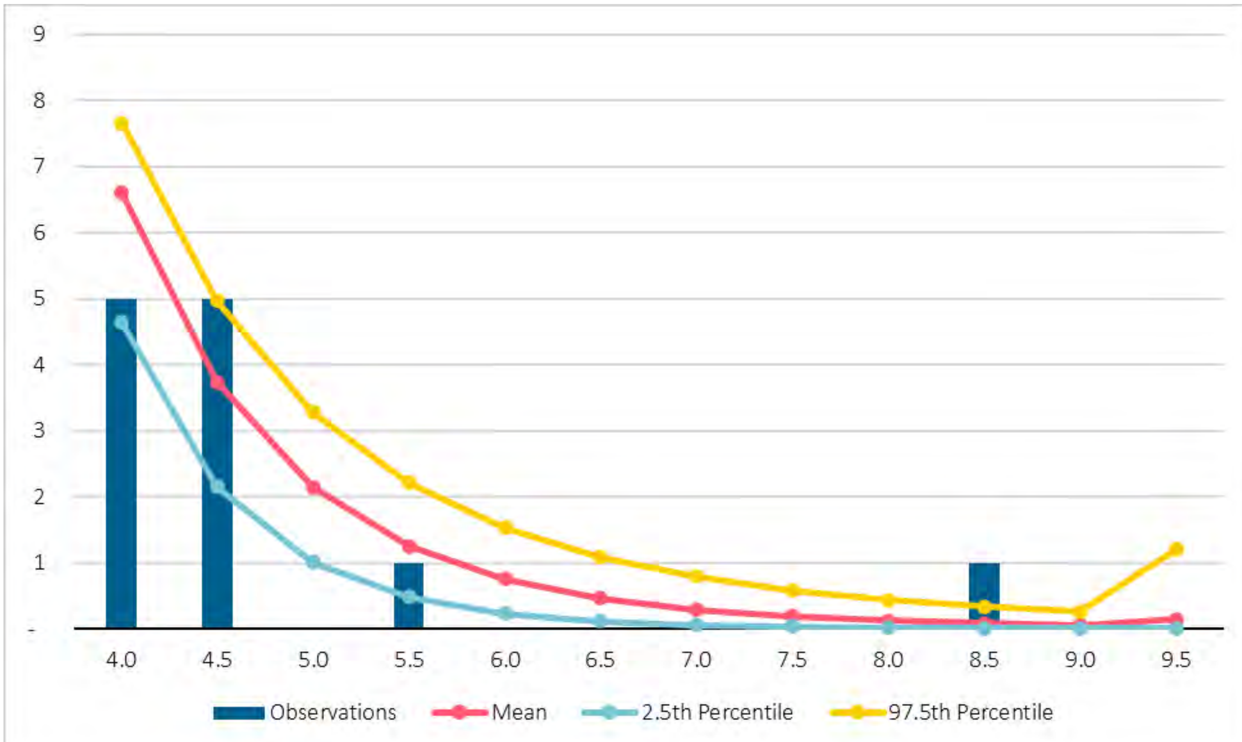


Figure 9
TAIL DISTRIBUTION ESTIMATES FROM MONTE CARLO SIMULATIONS



For annual maxima greater than 5 inches, Table 2 below shows the expected frequency of such an event in 85 years. While the expected frequency of an 8.5-inch maximum is still small under the 97.5th percentile, it is nevertheless 400% of the Mean expected frequency. This relative difference highlights the challenge of estimating the probability of extreme events.

Table 2
ESTIMATED OBSERVATIONS VS ACTUALS

Inches	Actuals	Mean	2.5th Percentile	97.5th Percentile
5.0	0	2.14	1.01	3.28
5.5	1	1.25	0.48	2.22
6.0	0	0.75	0.23	1.53
6.5	0	0.46	0.11	1.09
7.0	0	0.29	0.06	0.79
7.5	0	0.19	0.03	0.58
8.0	0	0.12	0.02	0.44
8.5	1	0.08	0.01	0.33
9.0	0	0.06	0.00	0.26
9.5	0	0.14	0.01	1.20

BOOTSTRAP RESAMPLING

The Bootstrap Resampling method follows a similar process, except that the scenario data is pulled from the original data set. For 1000 scenarios, we build a data set of 85 observations by selecting with replacement from the 85 observed annual maxima. We then find the parameter set that maximizes the log-likelihood for each scenario. From the resulting 1000 parameter estimates, we calculate standard errors and construct confidence intervals.

Table 3 below illustrates the resulting parameter estimates for the Location and Scale parameters from Bootstrap Sampling; notably these estimates are close to the estimates from the Monte Carlo simulations shown in Table 1 above.

Table 3
CONFIDENCE INTERVALS FOR GEV PARAMETERS FROM BOOTSTRAP RESAMPLING

	Best Estimate	Mean	Standard Error	CI Bottom	CI Top	2.5th Percentile	97.5th Percentile
Location	2.330	2.334	0.076	2.185	2.484	2.1837	2.4627
Scale	0.685	0.678	0.058	0.564	0.792	0.5683	0.7985
Shape	0.119	0.119	0.063	(0.004)	0.243	0.0466	0.2527

Figure 11 shows the resulting distribution for the Scale parameter estimates, with the distribution being similar to the Monte Carlo results in Figure 7.

Figure 10
FREQUENCY DISTRIBUTION OF SCALE PARAMETER ESTIMATES FROM BOOTSTRAP RESAMPLING

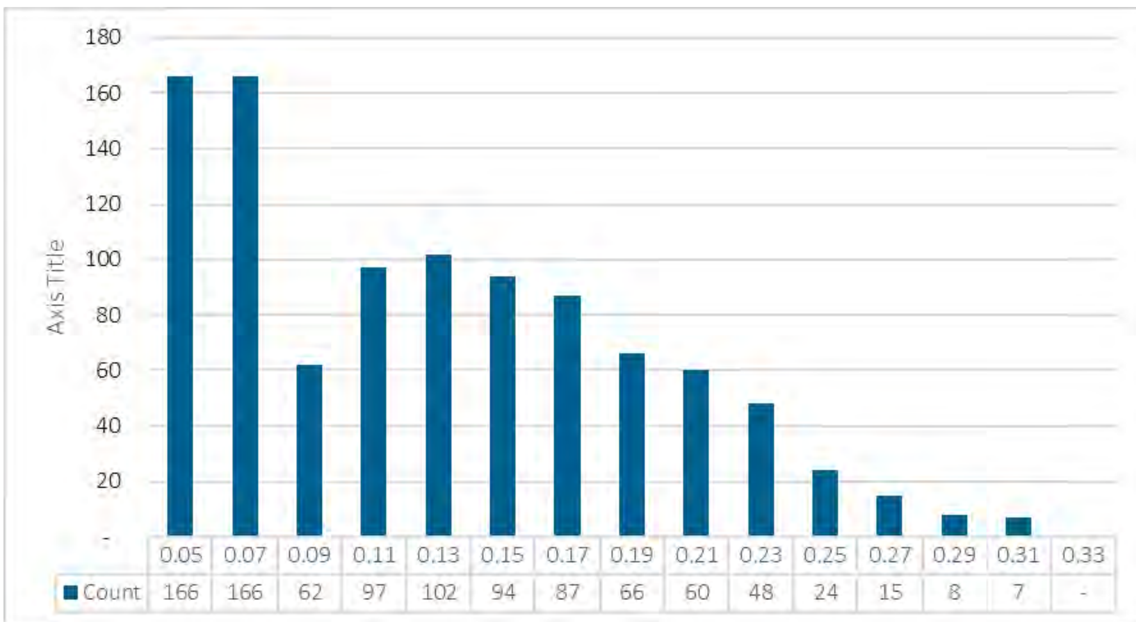


Figure 11
DISTRIBUTION ESTIMATES FROM BOOTSTRAP RESAMPLING

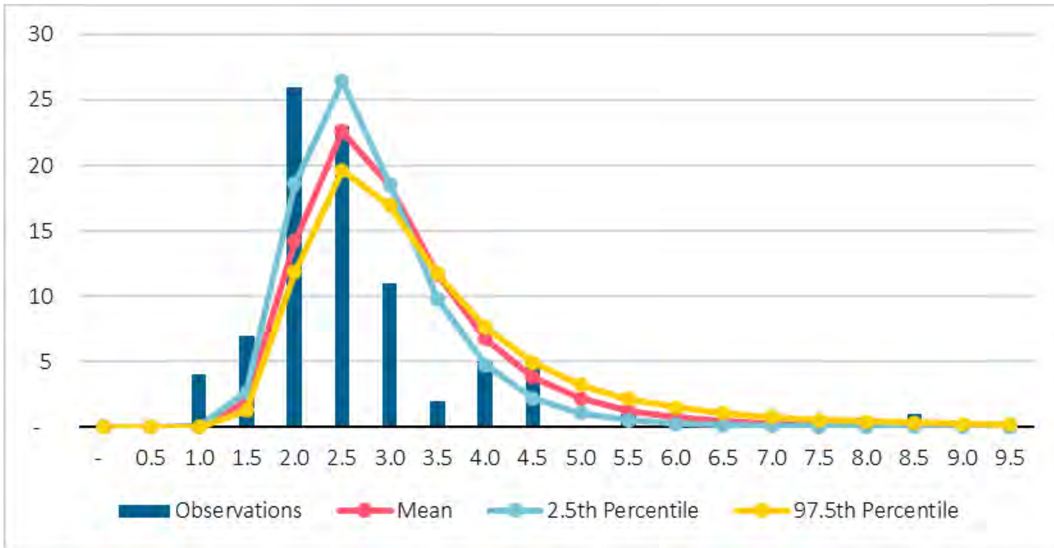
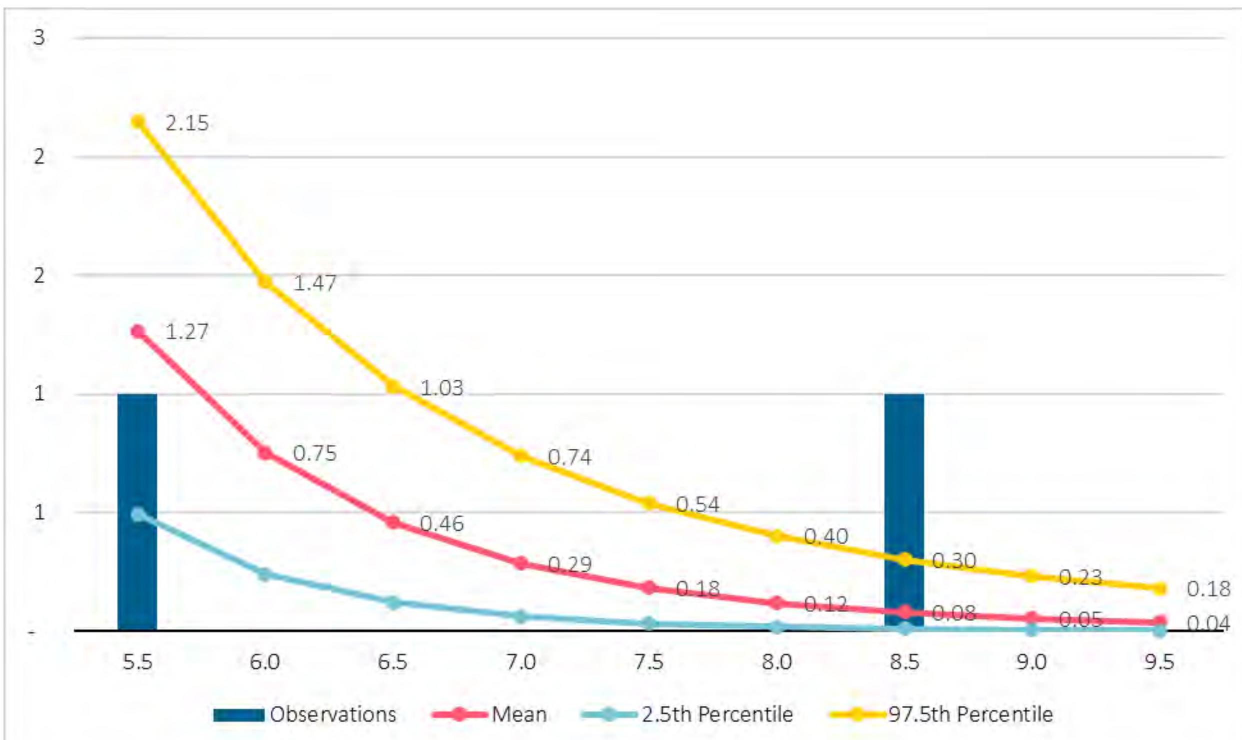


Figure 12
TAIL DISTRIBUTION ESTIMATES FROM BOOTSTRAP RESAMPLING



As expected, the distribution of the 2.5th percentile parameters shows a significantly lower expectation of extreme events. The expected frequency for an annual maximum of 8.5 inches aligns closely between the Bootstrap Resampling method and the Monte Carlo method.

Comparison of Estimation Methods

In order to validate the parameter estimations, we've compared the results of four different methods: 1) Incremental value testing within a specified range (i.e., brute force), 2) Gradient Descent with a set learning rate, 3) Excel's Solver functionality, and 4) Python's SciPy Stats library. From Table 4 below, we can see high alignment for the parameter estimates.

Table 4
PARAMETER ESTIMATES BY METHOD

Method	Location	Scale	Shape	Log Likelihood	Probability of > 2022 Max
Brute Force	2.2500	0.7100	0.1100	(108.466)	0.1915%
Gradient Descent	2.3194	0.6810	0.1222	(107.5475)	0.2019%
Excel Solver	2.3303	0.6847	0.1188	(107.5388)	0.1978%
Python	2.3303	0.6847	0.1188	(107.5388)	0.1977%

The differences between The Brute Force and Gradient Descent methods are due to the specified increments of the search algorithms. For the Brute Force method, we used increments of 0.25, 0.1, and 0.1 for the Location, Scale, and Shape parameters, respectively. For the Gradient Descent method, we used 1% changes in each parameter. To achieve higher precision with these methods, the discrete increments could be set smaller or updated dynamically based on the number of iterations and the iterative changes in the log-likelihood calculation.

For any Python users, it's important to note that Python's parameterization of the GEV distribution produces shape values with a sign that is opposite of the traditional parameterization, i.e., Python produces a shape value of -0.1188.

In the table below, we show the expected number of observations in an 85-year period by the annual maxima and method of parameter fitting. As expected from the table above, all of the expected values closely align.

Table 5
EXPECTED OBSERVATIONS BY ESTIMATION METHOD

Inches	Observations	Brute Force	Gradient Descent	Excel Solver	Python
5.0	0	2.1174	2.1626	2.1992	2.1993
5.5	1	1.2475	1.2639	1.2832	1.2832
6.0	0	0.7497	0.7560	0.7658	0.7658
6.5	0	0.4600	0.4630	0.4677	0.4677
7.0	0	0.2881	0.2902	0.2921	0.2921
7.5	0	0.1840	0.1858	0.1864	0.1864
8.0	0	0.1197	0.1213	0.1213	0.1213
8.5	1	0.0792	0.0807	0.0804	0.0804
9.0	0	0.0532	0.0546	0.0542	0.0542
9.5	0	0.0363	0.0376	0.0371	0.0371

Sensitivity of Results to the Parameter Estimations

Naturally, we would like to compare different parameter combinations to gauge the sensitivity of the log-likelihood function and of the estimated probabilities of extreme events. The table below shows the log-likelihood and the probability of an annual maxima greater than the 2022 maximum for 10 similar parameter combinations.

Table 6

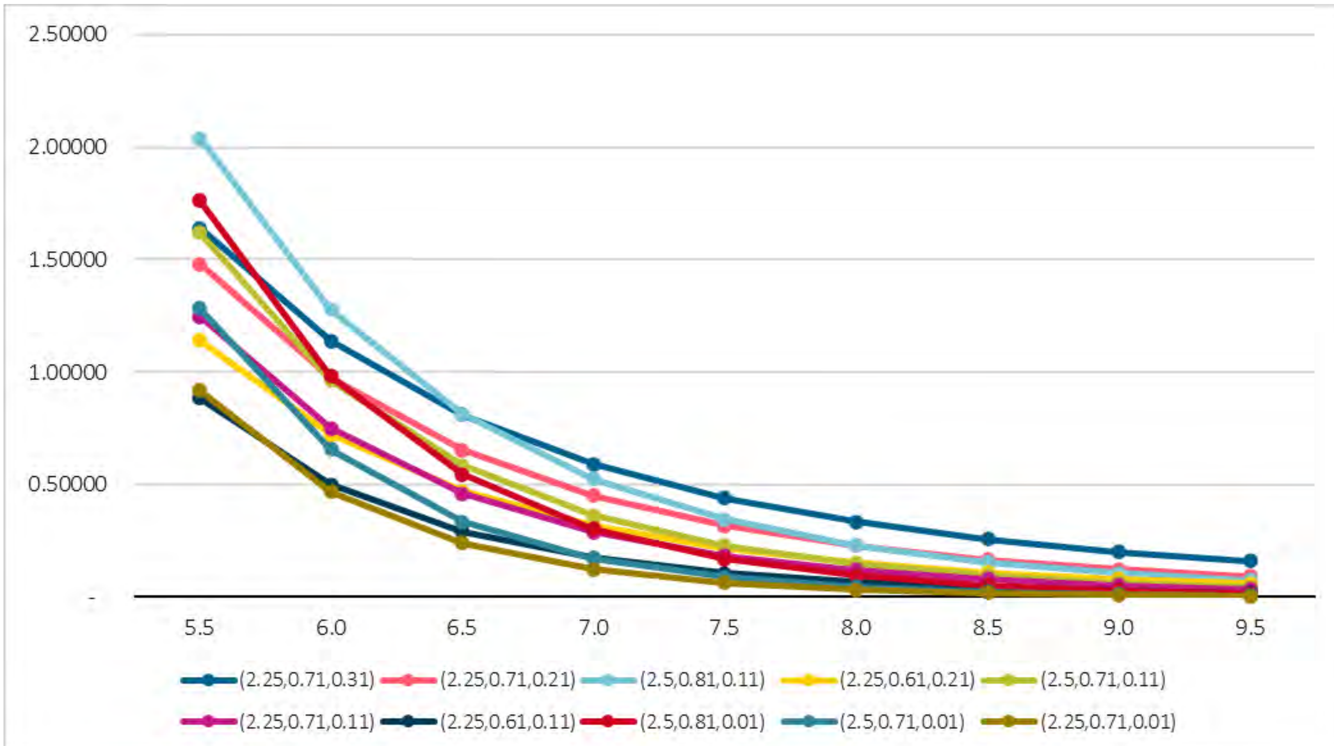
PROBABILITY OF EXCEEDING THE 2022 MAXMIUM FOR SELECTED GEV PARAMETERS

	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	Set 7	Set 8	Set 9	Set 10
Location	2.250	2.250	2.250	2.250	2.500	2.250	2.250	2.500	2.500	2.500
Scale	0.710	0.610	0.710	0.610	0.810	0.710	0.710	0.710	0.710	0.810
Shape	0.110	0.110	0.210	0.210	0.110	0.010	0.310	0.110	0.010	0.010
Log Likelihood	(108.5)	(108.6)	(108.7)	(108.9)	(109.9)	(110.0)	(110.0)	(110.1)	(110.3)	(110.4)
Probability of > 2022 Max	0.19%	0.09%	0.64%	0.39%	0.40%	0.02%	1.35%	0.23%	0.03%	0.07%

The log-likelihood values for all combinations are within the range of (-111,-108), indicating similar goodness of fit. Despite this similarity, we see in the last row of Table 5 that the probability of an annual maximum exceeding the 2022 maximum varies significantly. The probabilities increase as the Shape parameter increases, as the Shape parameter defines the "heaviness" of the tail.

The graph below shows the tail behavior for each of these parameter combinations, with the expected occurrence rate of an annual maximum of 8.5 inches labeled for three sets. All distributions show similar tail behavior, with the larger Shape parameters corresponding to greater expectations for more extreme annual maxima. While the estimated rates for 8.5 inches appear similar and are all below 0.50, there are significant differences in relativity of the rates that could translate to meaningful differences in the use or application of the rates, e.g., return-period estimation or premium for a parametric insurance product.

Figure 13
TRUNCATED OCCURRENCE RATES FOR 10 PARAMETER SETS



Probabilities of Extremes and Return Periods

Typically, the question of greatest concern is how often extreme events will occur. In the previous section, we've shown how the expected occurrence rates vary by parameter value. In this section, we'll explore how changes in the parameter values impact the percentiles and return periods. For clarity, the "return period" is an average estimated time between two events and is equal to the inverse of the annual probability of an event.

Table 6 below shows the implied return period of the 2022 observed maximum for five parameter sets. Because the return periods are equal to the inverse of the annual probability, the return period estimations are very sensitive to small changes in the probability. Additionally, extreme events are by definition low-probability events, thereby compounding the sensitivity of return period calculations.

Table 7
RETURN PERIODS FOR SELECT GEV PARAMETERS

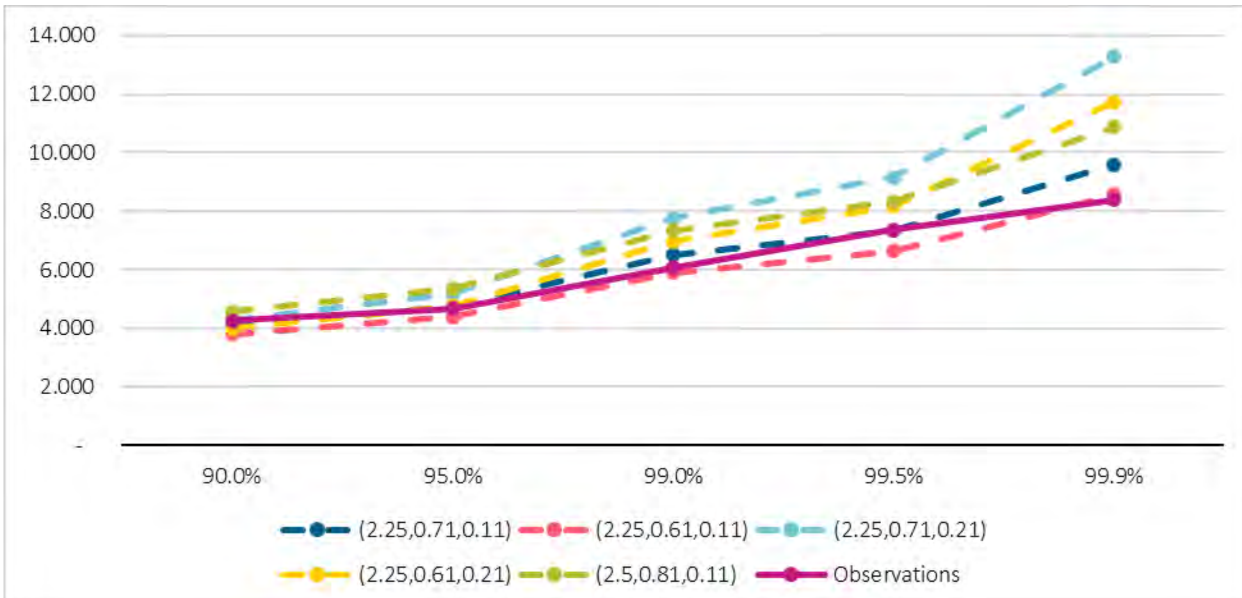
	Set 1	Set 2	Set 3	Set 4	Set 5
Location	2.2500	2.250	2.250	2.250	2.250
Scale	0.710	0.610	0.710	0.610	0.810
Shape	0.110	1.110	0.210	0.210	0.110
Probability of > 2022 Max	0.1915%	0.0939%	0.6361%	0.3921%	0.4022%
Implied Return Period	522	1,065	157	255	249

The difference between the probabilities becomes more notable when translated into differences between the return periods. While the smallest return is 157 years (corresponding to the highest probability), it's notable that this period is nearly twice the observation period of 85 years. Such a result emphasizes the extreme nature of the 2022 precipitation event.

Next, we can quantify the extreme percentiles for each distribution. The graph below shows how the percentiles align closely through the 90th percentile but increasingly diverge for larger percentiles. For the 99.9th percentile–corresponding to a 1,000-year return period–the smallest estimation is 8.56 inches, and the largest estimation is 13.29 inches. Such a wide range can present significant challenges for infrastructure planners that need to design projects to withstand an event with a prescribed return period, e.g., a storm drain can safely manage a 200-year precipitation event.

Figure 14

ANNUAL MAXIMA PERCENTILES BY PARAMETER SET



Summary and Conclusion

This analysis of St. Louis's unprecedented precipitation event in 2022 has sought to provide a clear, replicable framework for actuaries to use on other weather stations or weather variables. We have also tried to emphasize the need to quantify and communicate the uncertainty of any estimation about extreme weather events.

For those who wish to explore this framework in more detail, I recommend working through the Excel spreadsheet published with this report and following the results of this analysis in each tab. The workbook also provides resources for more detail about Extreme Value Theory and the relevant statistical techniques.

To apply this framework, I recommend identifying an extreme event report in 2022, such as river discharge rates or gauge heights in Kentucky; In July 2022, 37 people were killed in one of the worst flood events in the state's history.[1] Once an event of interest has been identified, the next step is finding the appropriate data set to quantify the event's severity. Then, applying the framework should simply entail updating the workbook to appropriately read the new data.

[1] Kentucky flooding death toll rises to 37 as governor says hundreds remain unaccounted for. Elizabeth Wolfe and Dakin Andone. CNN. August 1, 2022. [Kentucky flooding: Death toll rises to 37 as governor says hundreds remain unaccounted for | CNN]

About The Society of Actuaries Research Institute

Serving as the research arm of the Society of Actuaries (SOA), the SOA Research Institute provides objective, data-driven research bringing together tried and true practices and future-focused approaches to address societal challenges and your business needs. The Institute provides trusted knowledge, extensive experience and new technologies to help effectively identify, predict and manage risks.

Representing the thousands of actuaries who help conduct critical research, the SOA Research Institute provides clarity and solutions on risks and societal challenges. The Institute connects actuaries, academics, employers, the insurance industry, regulators, research partners, foundations and research institutions, sponsors and non-governmental organizations, building an effective network which provides support, knowledge and expertise regarding the management of risk to benefit the industry and the public.

Managed by experienced actuaries and research experts from a broad range of industries, the SOA Research Institute creates, funds, develops and distributes research to elevate actuaries as leaders in measuring and managing risk. These efforts include studies, essay collections, webcasts, research papers, survey reports, and original research on topics impacting society.

Harnessing its peer-reviewed research, leading-edge technologies, new data tools and innovative practices, the Institute seeks to understand the underlying causes of risk and the possible outcomes. The Institute develops objective research spanning a variety of topics with its strategic research programs: aging and retirement; actuarial innovation and technology; mortality and longevity; diversity, equity and inclusion; health care cost trends; and catastrophe and climate risk. The Institute has a large volume of topical research available, including an expanding collection of international and market-specific research, experience studies, models and timely research.

Society of Actuaries Research Institute
475 N. Martingale Road, Suite 600
Schaumburg, Illinois 60173
www.SOA.org