Estimation of Precipitation Rate Using Computational and Statistical Simulations

Abstract

It is not easy to use a single variable linear regression or simple exponential smoothing to determine the effectiveness of weather forecasts such as temperature or flooding forecast especially when the data pattern is complicated. In order to accurately predict the effectiveness of such a trend, iterative and statistical methods that can determine the status of temperature or flooding in the United States were chosen in this paper. The task of modeling the pattern in a focused period and performing data analysis were performed.

For the analysis, the extreme value theory was used to assess extreme events within probability distributions by quantifying tail behavior. By analyzing the maximum values of samples, it was possible to determine probabilities for extreme events. A comparison was made with events previously observed and analyzed for authenticity. As evident in our observations, lower values of data have much shorter return periods. In other words, they are more likely to reoccur; however, as the values increase for higher precipitation values, the length of the return periods increase exponentially. Therefore, there is a tendency for precipitation values to remain in lower ranges.

Introduction

Injuries and deaths caused by natural disasters take up a significantly high proportion of news each day. Natural disasters involve the interaction between hazardous natural factors that are extraneous to man and any population. These hazards are violent and calamitous phenomena, which include floods, tornadoes, tsunamis, earthquakes, fires, volcano eruptions, and hurricanes. Research shows that the number of natural disasters around the world in the past decades has significantly risen. The impact of these natural phenomena is catastrophic: buildings collapse, and many deaths follow. There are additional, long-lasting effects that exist after the disaster as well, such as famine, economic losses, several health risks, and emotional aftershocks.

Floods are the world's most common natural disaster, causing the majority of natural disaster fatalities. Floods are identified with a large, abnormal amount of water overflowing over a normally dry land. They follow a normal law of nature and happen accordingly with seasons.

The destructiveness of floods and the severe winds that accompany the floods are mainly due to the mechanical force of the water and the debris it carries, along with the contamination and wetness of the flood water. Floods happen regularly in the U.S. following heavy precipitation in the spring or melting of snow and ice in the mountains and cause a high number of deaths and property damage. Most of the deaths caused by floods are related to the effects of flash floods, which happen when there is a fast runoff of water from heavy precipitation in a short amount of time [1].

To minimize the damage caused by these natural disasters, it is important to prepare for these events. The predictability of catastrophic natural disasters is crucial but this task of modeling natural disasters in a focused area and performing data analysis is difficult especially when the data pattern is complex. Extracting patterns from the analytical process to characterize the natural disaster is a difficult job for the researchers.

Using the Gumbel distribution method, this research focuses on the predictability of floods, for they account for a large amount of casualties out of all the natural disasters. By using statistical and computational simulations including a regression analysis using least square method, we investigated the exceedance probability, which is the probability that the event will exceed some critical value. Also the return period corresponding to this exceedance probability was found.

Objectives

1. Study hydrologic extremes for design and assessing the impacts of rare climatic events.
2. Introduce a framework for estimating stationary and non-stationary return levels, return periods, and risks of climatic extremes using Bayesian inference.
3. Find return levels and return periods framework implemented in the non-stationary extreme value analysis , explicitly designed to facilitate analysis of extremes in the geosciences.
4. Manage the risks of extreme events and disasters figuring out how the global warming and precipitation would change temporal and spatial pattern of climatic extremes.

Wetland and Dryland

This study examined space-time historical flood and drought variability spanning for a century for the:

1. Wetland(Allegheny river, NY) and
2. Dryland(Great Plains)

EVT - Extreme value theory

Extreme value theory is used to model the risk of extreme, rare events, such as the 1755 Lisbon earthquake. Extreme value theory or extreme value analysis (EVA) is a branch of statistics dealing with the extreme deviations from the median of probability distributions. It seeks to assess, from a given ordered sample of a given random variable, the probability of events that are more extreme than any previously observed.

Extreme value analysis is widely used in many disciplines, such as structural engineering, finance, earth sciences, traffic prediction, and geological engineering. For example, EVA might be used in the field of hydrology to estimate the probability of an unusually large flooding event, such as the 100-year flood. Similarly, for the design of a breakwater, a coastal engineer would seek to estimate the 50-year wave and design the structure accordingly.

3 distribution parameters $\theta=(\mu,\sigma,\xi)$:

(1) the location parameter (μ) specifies the center of the distribution

(2) the scale parameter (σ) determines the size of deviations around the location parameter

(3) the shape parameter (ξ) governs the tail behavior of the distribution.

The limiting case of $\xi \rightarrow 0$ gives the Gumbel distribution, $\xi < 0$ the Weibull distribution and $\xi > 0$ the Fréchet distribution.

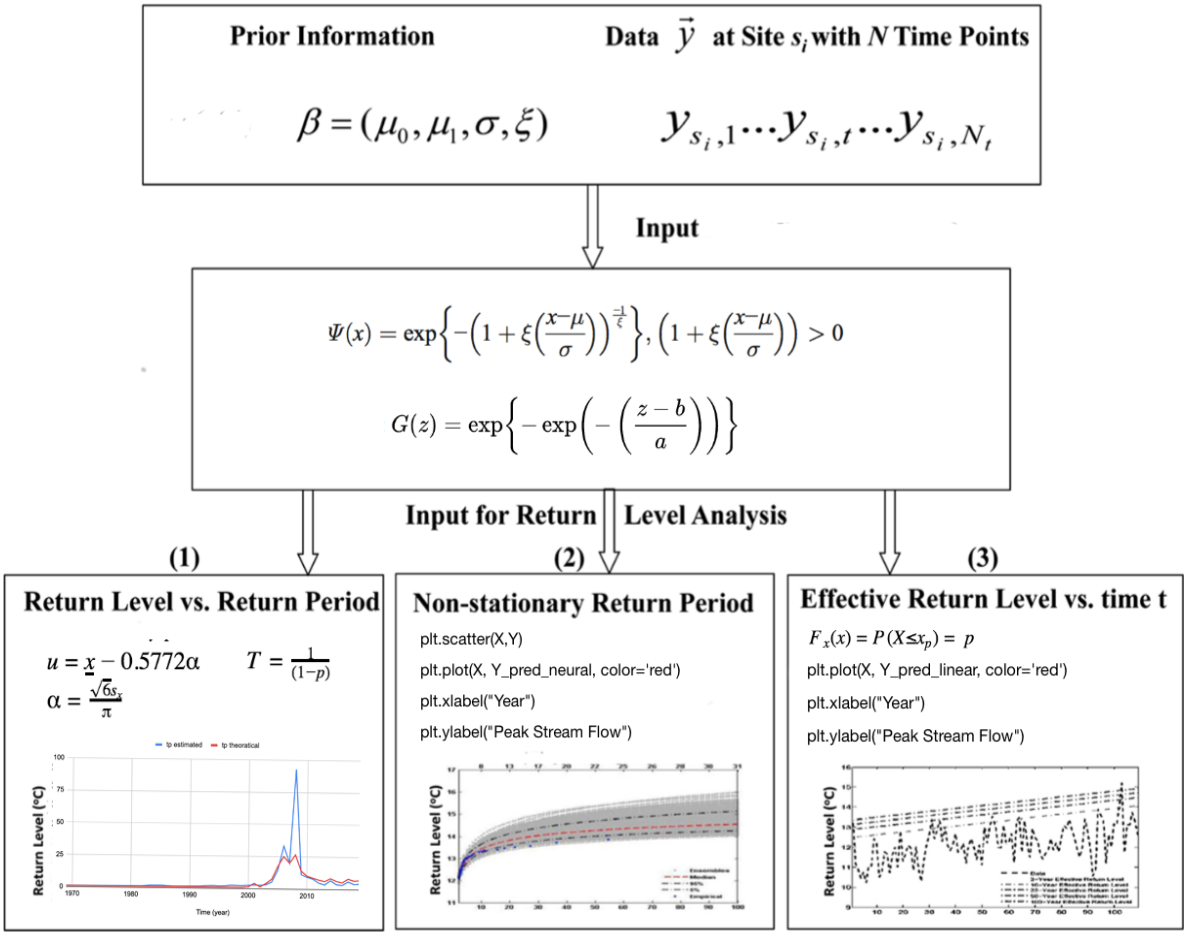


Fig. A Flow chart of the analysis

Research Tools

MATLAB and Python Programming

Microsoft Excel

Backgrounds

A. Time Series Analysis and Forecasting

There have been numerous different kinds of data such as stock prices and interest rates observed and gathered in the past. The sequential nature of these data requires us to account for the dynamic nature using special statistical skill and techniques. Time series analysis provides the appropriate methods necessary in order to analyze sequential data.

B. Smoothing

It may be problematic to picture the essential, underlying trend of the data if the time series has a lot of noise. To distinguish the signal and the noise from each other, various linear and nonlinear smoothers must be applied.

C. Curve fitting

In MATLAB®, best-fit line are available using the least-sum-of-squares line from the data. Also known as linear least squares regression or least squares regression line (LSRL), this type of linear modeling minimizes the sum of the squares of the deviations between the model and the actual data. [6] The deviations are squared in order to reduce the influence of negative or positive signs when added.

In general terms, curve fitting involves either interpolation - in which the fitting model exactly matches the data, but often in piecewise manners - or smoothing - in which the noise of the data is reduced and a function approximates the overall trend of the data. [7] [8] Curve fitting can be used to not only map out the data and render it computable in general terms, but also extrapolate other data points based on the trend provided by the model. [8]

- **Conceptual Outcomes**
  - This research demonstrates the understanding of return periods and frequency analysis of the natural disaster. We demonstrate the statistical parameters used in the frequency analysis
- **Practical Outcomes**
  - We utilized Matlab and MS Excel to make a statistical and computational analysis.
  - We used statistical distribution methods and regressions in order to create a frequency curve.
- **Computing/Data Outputs**
  - Numerical values: Theoretical and estimated return periods
  - Graphical: graph based on the real data, regression, and forecast curves
- **Data Collection - Real-time/Daily Data**

- To download data, the National Water Information System (NWIS) web interface was used.

Plotting the Flood Frequency Curve using Gumbel Distribution

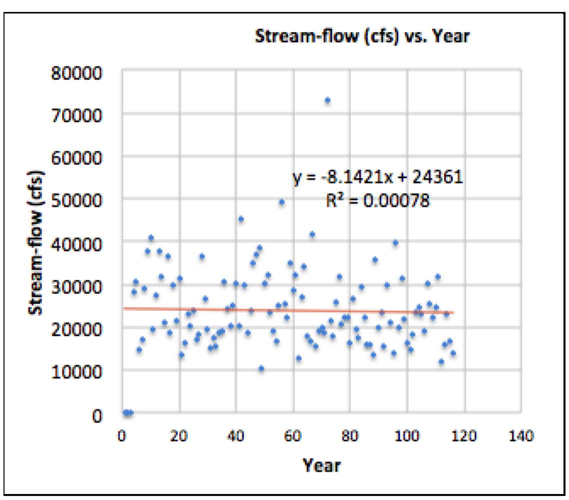
- Exceedance probability is the probability that the event will exceed some critical value (usually far from the mean).
- Return period is an estimate of the likelihood of an event to occur. A statistical measurement based on historic data denoting the average recurrence interval.
- Theoretical probability is the fraction of times we expect the event to occur if we repeat the same experiment over and over (i.e. flipping a coin and getting heads or tails is each 0.50).
- Estimated probability approaches the theoretical probability as the number of trials gets larger. It is an approximation of theoretical probability.

Investigated Area

This study examined space-time historical flood and drought variability spanning for a century for the:

1. Wetland(Allegheny river, NY)

USGS 03011020 ALLEGHENY RIVER AT SALAMANCA NY (Duration: 1904-2016)
Cattaraugus County, New York
Hydrologic Unit Code 05010001
Latitude 42°09'23", Longitude 78°42'55" NAD83
Drainage area 1,608 square miles
Contributing drainage area 1,608 square miles
Gage datum 1,358 feet above NGVD29



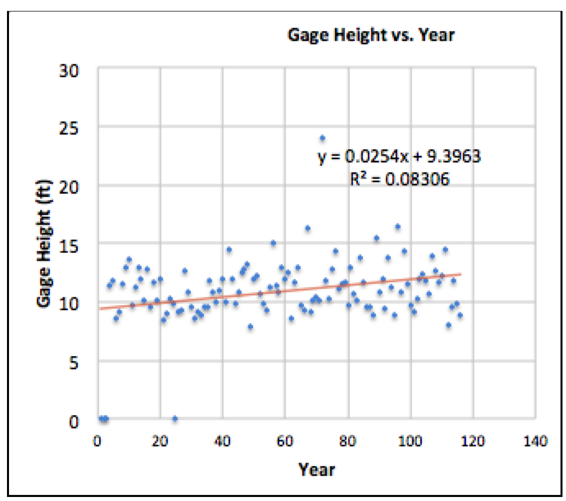


Fig. 1 Streamflow(cfs) of the Allegheny river in NY(1904-2016) - Used MATLAB for curve fit

Procedures

1) Label column A(first column) as 'time' and column B(second column) as the 'annual peak streamflow (cfs)' and enter in appropriate values in each cell.
2) Select the streamflow values in column B then sort the values from smallest to largest by clicking on the 'Sort and Filter' tool. Allow expanding the selection. Peak Streamflow for the Allegheny River in Salamanda, NY is as follows:

	A	B
1	Time	Peak Streamflow(cfs)
2	Jan. 06, 1949	10,300
3	Jan. 28, 2012	11,900
4	Apr. 08, 1962	12,800
5	Mar. 07, 1921	13,500
6	Apr. 04, 1988	13,700
7	Nov. 02, 1994	14,100
8	Apr. 12, 2016	14,100
9	Apr. 08, 2001	14,700
10	Jan. 24, 1906	14,800

Fig. 2 Time-Streamflow of the Allegheny River in Salamanda, NY

3) Label column C as 'Rank (i)' and rank the data in decreasing order (from N to 1).
4) Create a fourth column called q_i. Gringorten's plotting position formula will be used to calculate the estimated exceedance probabilities relevant to past observations. An example can be seen in Table C below.

$$q_i = \frac{i-a}{N+1-2a} \qquad (1)$$

q_i = exceedance probability associated with a specific observation

$N =$ number of annual maxima observations

$i =$ Rank of specific observation (i=1 is the largest and i=N is the smallest)

$a =$ constant for estimation (0.44)

	A	B	C	D
1	Time	Peak Streamflow(cfs)	Rank (i)	qi
2	Jan. 06, 1949	10,300	103	0.906647808
3	Jan. 28, 2012	11,900	102	0.897807638
4	Apr. 08, 1962	12,800	101	0.888967468
5	Mar. 07, 1921	13,500	100	0.880127298
6	Apr. 04, 1988	13,700	99	0.871287129
7	Nov. 02, 1994	14,100	98	0.862446959
8	Apr. 12, 2016	14,100	97	0.853606789
9	Apr. 08, 2001	14,700	96	0.84476662
10	Jan. 24, 1906	14,800	95	0.83592645
11	Apr. 11, 1931	15,000	94	0.82708628
12	Mar. 15, 1933	15,500	93	0.81824611
13	Oct. 20, 1967	15,500	92	0.809405941
14	Sep. 23, 1992	15,500	91	0.800565771
15	Jul. 03, 1987	15,800	90	0.791725601

Fig.3 Exceedance probability associated with a specific observation

5) Make another column and label it p_i. Then make it equal to $1 - q_i$. p_i refers to the non-exceedance probability.

A. Statistical Definition of Return Period

- If X is a random variable with a cumulative distribution function $F_x(x)$, the probability that X is less than equal (not exceeding) to a given event x_p is:

$$F_x(x) = P(X \le x_p) = p \qquad (2)$$

- The probability that this event will be exceeded is now $1 - p$, and the percent exceedance would be $100(1 - p)\%$.
- For an event x_p, the return period corresponding to this exceedance probability is denoted by T.

$$T = \frac{1}{(1-p)} \qquad (3)$$

- For example, a 100-year return period is an event with a probability of exceedance $1 - p = 0.01$ or a non-exceedance probability $p = 0.99$. There is a 99% chance that this event will not be exceeded within a given year.

6) Create one more column and label it 'T_p estimated' and evaluate the values in p_i using the equation for the return period. An example can be seen in Table D below.

6) We will follow the 'Gumbel' or 'Extreme Value Type 1' distribution. The CDF (Cumulative distribution of function) of the Gumbel distribution is the following:

$$F_x(x) = exp\,[-\,exp\,(-\,\frac{x-u}{\alpha})\,] = p \qquad (4)$$

where the x is observed data; u and α are the calculated parameters of the distribution. This distribution will allow us to calculate the theoretical estimate of p.

5) Create one more column and label it 'T_p estimated' and evaluate the values in p_i using the equation for the return period.

	A	B	C	D	E	F
1	Time	Peak Streamflow(cfs)	Rank (i)	qi	pi	Tp estimated
2	Jan. 06, 1949	10,300	103	0.9066478	0.0933522	1.10296412
3	Jan. 28, 2012	11,900	102	0.8978076	0.1021924	1.11382434
4	Apr. 08, 1962	12,800	101	0.8889675	0.1110325	1.12490056
5	Mar. 07, 1921	13,500	100	0.8801273	0.1198727	1.13619928
6	Apr. 04, 1988	13,700	99	0.8712871	0.1287129	1.14772727
7	Nov. 02, 1994	14,100	98	0.862447	0.137553	1.15949159
8	Apr. 12, 2016	14,100	97	0.8536068	0.1463932	1.17149959
9	Apr. 08, 2001	14,700	96	0.8447666	0.1552334	1.18375889
10	Jan. 24, 1906	14,800	95	0.8359264	0.1640736	1.1962775
11	Apr. 11, 1931	15,000	94	0.8270863	0.1729137	1.2090637
12	Mar. 15, 1933	15,500	93	0.8182461	0.1817539	1.22212619
13	Oct. 20, 1967	15,500	92	0.8094059	0.1905941	1.23547401
14	Sep. 23, 1992	15,500	91	0.8005658	0.1994342	1.24911661
15	Jul. 03, 1987	15,800	90	0.7917256	0.2082744	1.26306387

Fig.4 T_p estimated and the values in p_i

'T_p estimated' is the estimated distribution of 35 years of data. We assume that the data follows a specific distribution to estimate the parameters.

7) Create two columns labeled '$(x - u)/\alpha$' and 'p-theoretical. Using the following equations, calculate x̄, s_x, u and α. Table E (below) shows such values that result from the existing data.

$$\underline{x} = \sum_{i=1}^{n} (\frac{x_i}{n}) \qquad (5)$$

$$s_x^{\ 2} = \frac{1}{(n-1)} \sum_{i=1}^{n} [(x_i - \underline{x})^2] \qquad (6)$$

$$u = \underline{x} - 0.5772\alpha \qquad (7)$$

$$\alpha = \frac{\sqrt{6}s_x}{\pi} \qquad (8)$$

x-bar	22619.41748
Sx	6346.638113
u	19763.17015
alpha	4948.453436

6) Use the peak streamflow values (x) and calculate the column $(x - u)/\alpha$ as shown:

	A	B	C	D	E	F	G
1	**Time**	**Peak Streamflow(cfs)**	**Rank (i)**	**qi**	**pi**	**Tp estimated**	**(x-u)/a**
2	Jan. 06, 1949	10,300	103	0.906648	0.093352	1.1029641	-1.912349035
3	Jan. 28, 2012	11,900	102	0.897808	0.102192	1.1138243	-1.58901569
4	Apr. 08, 1962	12,800	101	0.888967	0.111033	1.1249006	-1.407140684
5	Mar. 07, 1921	13,500	100	0.880127	0.119873	1.1361993	-1.265682346
6	Apr. 04, 1988	13,700	99	0.871287	0.128713	1.1477273	-1.225265678
7	Nov. 02, 1994	14,100	98	0.862447	0.137553	1.1594916	-1.144432342
8	Apr. 12, 2016	14,100	97	0.853607	0.146393	1.1714996	-1.144432342
9	Apr. 08, 2001	14,700	96	0.844767	0.155233	1.1837589	-1.023182337
10	Jan. 24, 1906	14,800	95	0.835926	0.164074	1.1962775	-1.002974003
11	Apr. 11, 1931	15,000	94	0.827086	0.172914	1.2090637	-0.962557335
12	Mar. 15, 1933	15,500	93	0.818246	0.181754	1.2221262	-0.861515665
13	Oct. 20, 1967	15,500	92	0.809406	0.190594	1.235474	-0.861515665
14	Sep. 23, 1992	15,500	91	0.800566	0.199434	1.2491166	-0.861515665
15	Jul. 03, 1987	15,800	90	0.791726	0.208274	1.2630639	-0.800890663

Fig.5 Calculation of $(x - u)/\alpha$

7) Use the CDF equation from step 7 to calculate the value of p-theoretical.
8) Use the equation used to calculate 'T_p estimated' and use it to calculate 'T_p theoretical' using the p theoretical values.

	A	B	C	D	E	F	G	H	I
1	Time	Peak Streamflow(cfs)	Rank (i)	qi	pi	Tp estimated	(x-u)/a	p theoretical	Tp theoretical
2	Jan. 06, 1949	10,300	103	0.9066478	0.0933522	1.10296412	-1.912349	0.00114888	1.001150198
3	Jan. 28, 2012	11,900	102	0.8978076	0.1021924	1.11382434	-1.589016	0.0074546	1.007510585
4	Apr. 08, 1962	12,800	101	0.8889675	0.1110325	1.12490056	-1.407141	0.01683558	1.017123875
5	Mar. 07, 1921	13,500	100	0.8801273	0.1198727	1.13619928	-1.265682	0.02885387	1.029711151
6	Apr. 04, 1988	13,700	99	0.8712871	0.1287129	1.14772727	-1.225266	0.03320448	1.034344879
7	Nov. 02, 1994	14,100	98	0.862447	0.137553	1.15949159	-1.144432	0.04325433	1.045209847
8	Apr. 12, 2016	14,100	97	0.8536068	0.1463932	1.17149959	-1.144432	0.04325433	1.045209847
9	Apr. 08, 2001	14,700	96	0.8447666	0.1552334	1.18375889	-1.023182	0.06191245	1.065998579
10	Jan. 24, 1906	14,800	95	0.8359264	0.1640736	1.1962775	-1.002974	0.06545594	1.070040507
11	Apr. 11, 1931	15,000	94	0.8270863	0.1729137	1.2090637	-0.962557	0.07292061	1.07865627
12	Mar. 15, 1933	15,500	93	0.8182461	0.1817539	1.22212619	-0.861516	0.09378549	1.103491485
13	Oct. 20, 1967	15,500	92	0.8094059	0.1905941	1.23547401	-0.861516	0.09378549	1.103491485
14	Sep. 23, 1992	15,500	91	0.8005658	0.1994342	1.24911661	-0.861516	0.09378549	1.103491485
15	Jul. 03, 1987	15,800	90	0.7917256	0.2082744	1.26306387	-0.800891	0.107795	1.120818644

Fig.6 Calculation of T_p estimated and T_p theoretical using the p theoretical values.

Graphing the Flood Frequency Curve for wetland

Go to 'insert' tab and select charts. Plot 'T_p estimated' vs Annual Streamflow. On the same graph, also plot 'T_p theoretical' vs Annual Streamflow. Label the chart title and the axes on the obtained graph.

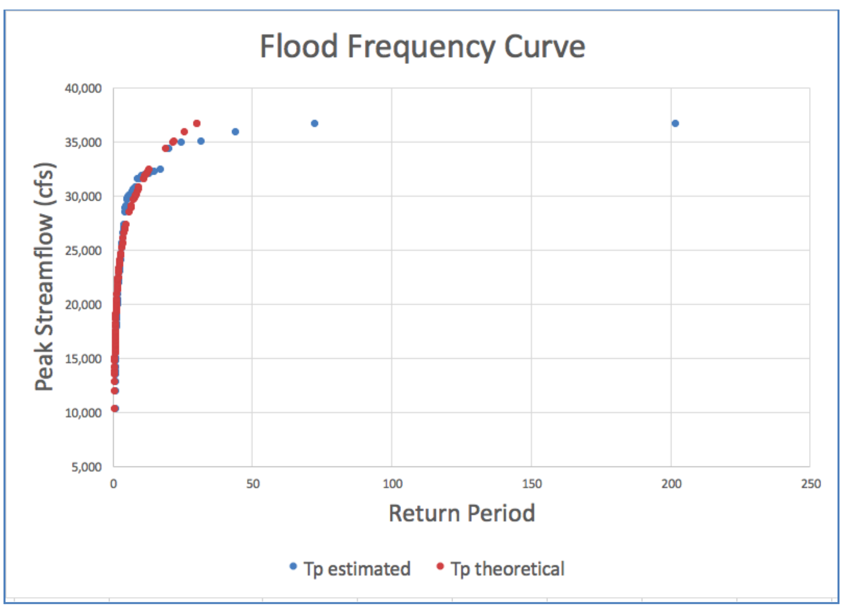


Fig.7 Flood frequency curve

Using the Matlab, interpolation method was used to determine the curve fitted line that most appropriately describes the pattern of the data.

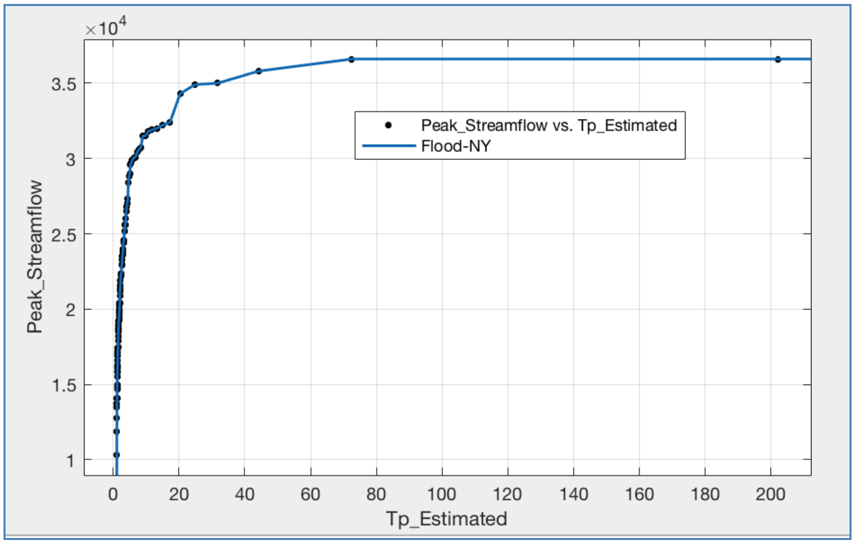
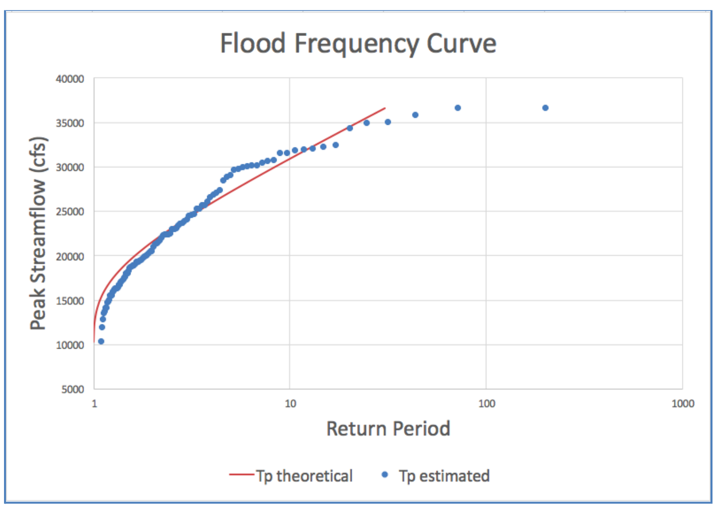
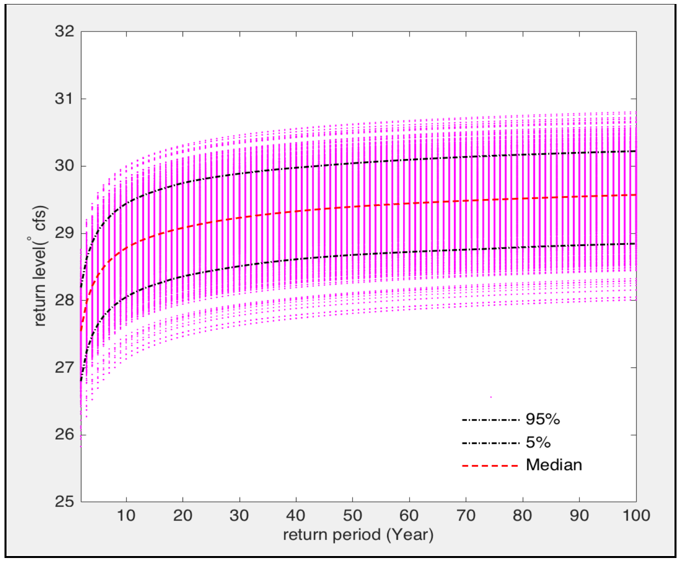


Fig.8 Flood frequency curve

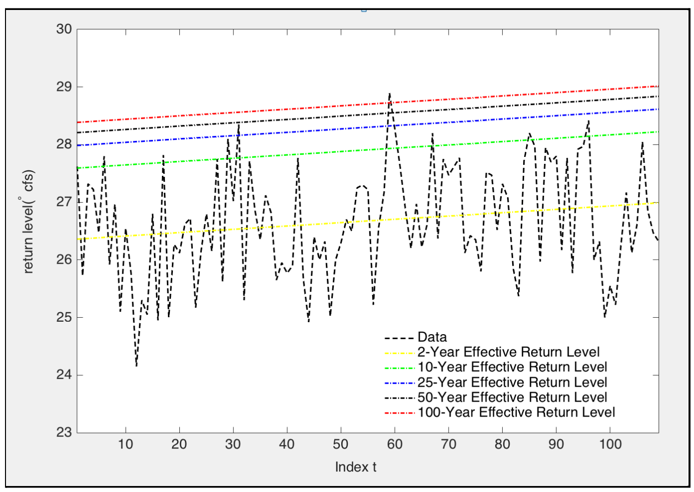
- Right click on the curve and select 'change chart type' and click on 'scatter with smooth lines' for theoretical and 'scatter' for estimated.
- Right click again and select 'format chart area' to use the axis options command for the X-axis and select logarithmic scale. It will then contain return periods displayed from 1 to 100 in log scale.



(a) Flood frequency curve(with Theoretical)



(b) Flood frequency curve(Median and 95%/5% tolerances)



(c) Flood frequency curve

Fig.9 Flood frequency curve(with Theoretical)

The orange line represents the theoretical distribution, while the blue dots stand for the fit of the annual peak streamflow data with respect to a Gumbel distribution. You can predict streamflow values corresponding to any return period from 1 to 100. The curve follows the distribution very well for low flows, but starts to drift away from the theoretical at higher flows.

Graphing the Flood Frequency Curve for Dryland (The Great plain)

Precipitation anomaly (inches) for annual (black), winter(blue), spring (green), summer (red), and fall (orange), for the northern (solid lines) and southern (dashed lines) in the U.S. Great Plains. Dashed lines indicate the best fit by minimizing the chi-square error statistic

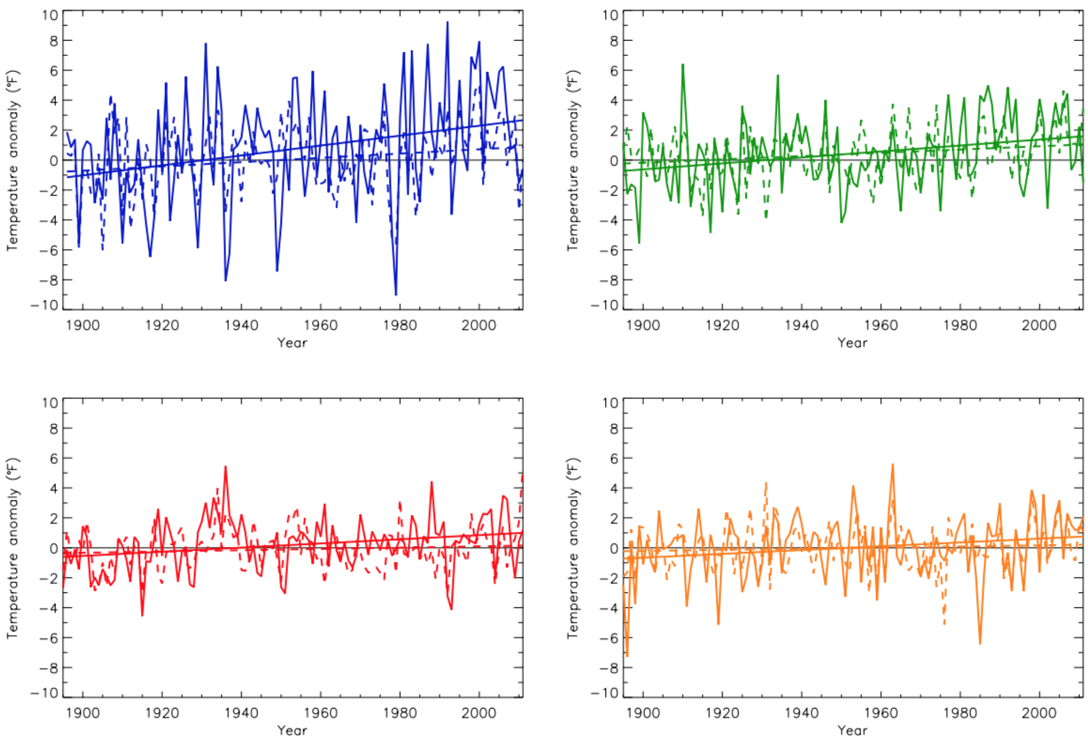


Fig.10 Precipitation anomaly (inches) for annual (black), winter(blue), spring (green), summer (red), and fall (orange), for the northern (solid lines) and southern (dashed lines) in the U.S. Great Plains

The same method is used as we did to plot the curve for wetland to graph the flood frequency curve for dryland.

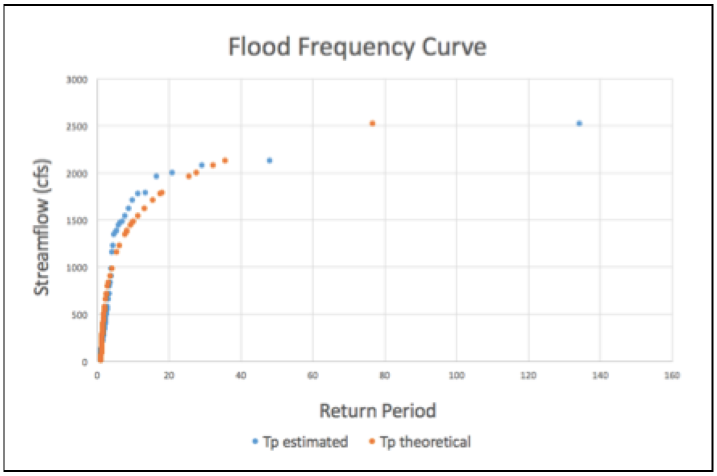


Fig.11 T_p estimated, T_p theoretical vs P/E Ratio Value

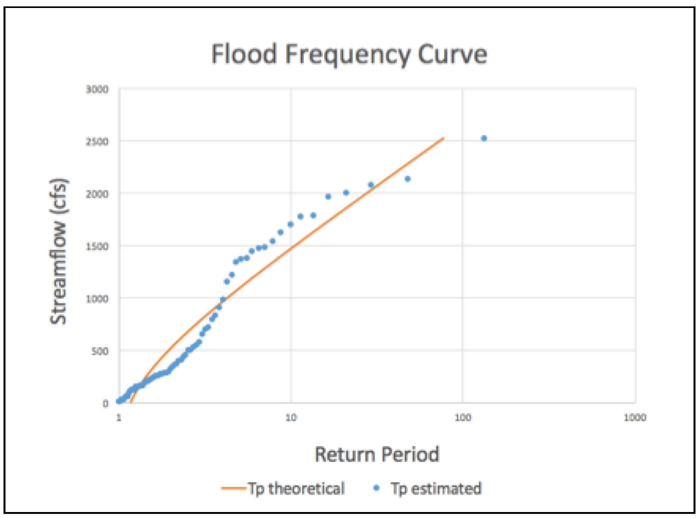


Fig.12 T_p estimated, T_p theoretical vs P/E Ratio Value(log scale)

The orange line represents the theoretical distribution, while the blue dots stand for the fit of the annual peak streamflow data with respect to a Gumbel distribution. You can predict streamflow values corresponding to any return period from 1 to 100. The curve follows the distribution very well for low flows, but starts to drift away from the theoretical at higher flows. (Right: log scale to predict it in 1 to 1000 years)

- Linear regression(left) vs. Neural Network(right) - NY Allegheny River

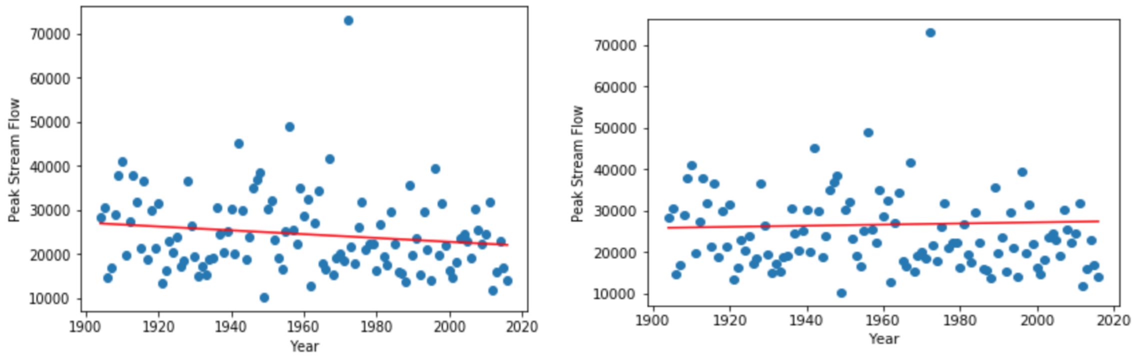


Fig.12 Linear regression(left) vs. Neural Network(right) - NY Allegheny River

Discussion

Flooding happens all throughout the United States, despite the climate differences. There are high numbers of flood events observed in the Northeast into the Midwest. Many floods occur in the spring in these regions from rainfall and the melting of snow. The area by the Red River, along the border of Minnesota and North Dakota, has noticeable numbers of river flooding. As mentioned previously, floods account for a large significant amount of damage done by all types of natural disasters. More than 2.5 billion people were affected by floods in the past few decades, with many people injured, homeless, or dead [3].

The prediction of increased flood frequency and damage causes for an increased focus on ways to prevent floods and minimize its damage. One main solution is to make sure people are aware of flood warnings as well as the appropriate responses of what to do when there is a flood. For example, they should be aware of risky behaviors such as entering flood waters, which may increase their chances of getting injured or drowning. In addition to the increased awareness, continuing to study and learn more about flood losses and patterns of natural disasters can help form future solutions and strategies to continue minimizing flood casualties.

Although the climate is dry, the Deserts in the Southwest are still prone to floods. Because the Southwestern areas are larger, they tend to have a higher concentration of flood events — a larger geographic footprint for floods. Flash floods also happen quite often in the summer from thunderstorms due to the Southwest monsoon. Because of the large amount of precipitation, streams, rivers, creeks, and other bodies of water are quickly filled up and lead to high water levels.

In this paper, the USGS geographical information and Gumbel distribution were used to find the return period corresponding to the exceedance probability. The Gumbel distribution is applied to San Jose in CA, Dryland (The Great plain), and Allegheny River NY.

References

1. Chang, Li-Chiu; Shen, Hung-Yu; Chang, Fi-John (2014-11-27). "Regional flood inundation nowcast using hybrid SOM and dynamic neural networks". Journal of Hydrology. 519 (Part A): 476–489. doi:10.1016/j.jhydrol.2014.07.036.
2. Application of self-organising maps and multi-layer perceptron-artificial neural networks for streamflow and water level forecasting in data-poor catchments: the case of the Lower Shire floodplain, Malawi
3. Hydrologic Ensemble Prediction EXperiment, an informal yet highly active group of researchers in the field of predictive hydrologic uncertainty
4. "AMS Glossary". allenpress.com. Archived from the original on 16 July 2012. Retrieved 9 July 2015.
5. https://nwis.waterdata.usgs.gov/nwis/peak?site_no=03011020&agency_cd=USGS&format=html
6. https://nwis.waterdata.usgs.gov/nwis/peak?site_no=03011020&agency_cd=USGS&format=html
7. Numerical Methods of Curve Fitting. By P. G. Guest, Philip George Guest. Page 349.
8. https://en.wikipedia.org/wiki/Gumbel_distribution
9. https://en.wikipedia.org/wiki/Return_period
10. https://www.usna.edu/Users/oceano/pguth/md_help/html/time6h9j.htm
11. https://www.itl.nist.gov/div898/handbook/pmd/section1/pmd141.html
12. http://web.iitd.ac.in/~pmvs/courses/mel705/curvefitting.pdf
13. https://en.wikipedia.org/wiki/Curve_fitting

Appendix - Data collection procedures:

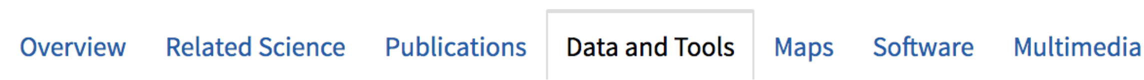
Reference links

https://nwis.waterdata.usgs.gov/nwis/peak/?site_no=06891500

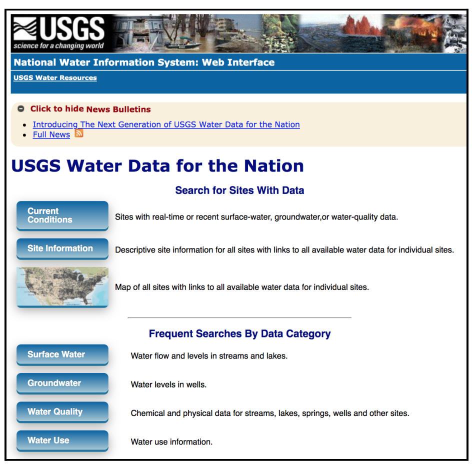
1. Go to

https://www.usgs.gov/mission-areas/water-resources/science/usgs-streamgaging-network

2. Click the Data and Tools from the munus below the slider image:



3. Click "Water Data for the Nation (NWIS): Automated Retrievals" to obtain USGS water data via automated retrievals



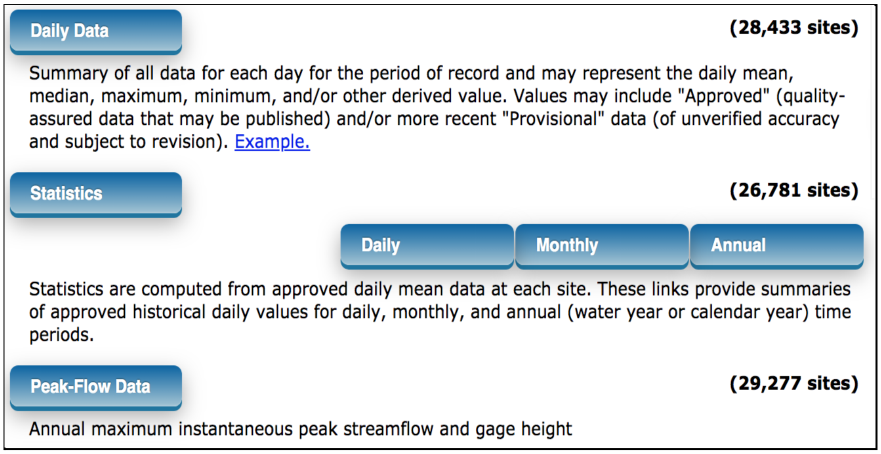
4. You are now directed to a new window: https://waterdata.usgs.gov/nwis?automated_retrieval_info
5. Select Surface Water and USA from the dropdown menu on top right corner:

6. The rest of the steps to extract data are shown as follows:
   - **Procedure**

- Click on Surface Water, to see water flow and levels in streams and lakes.

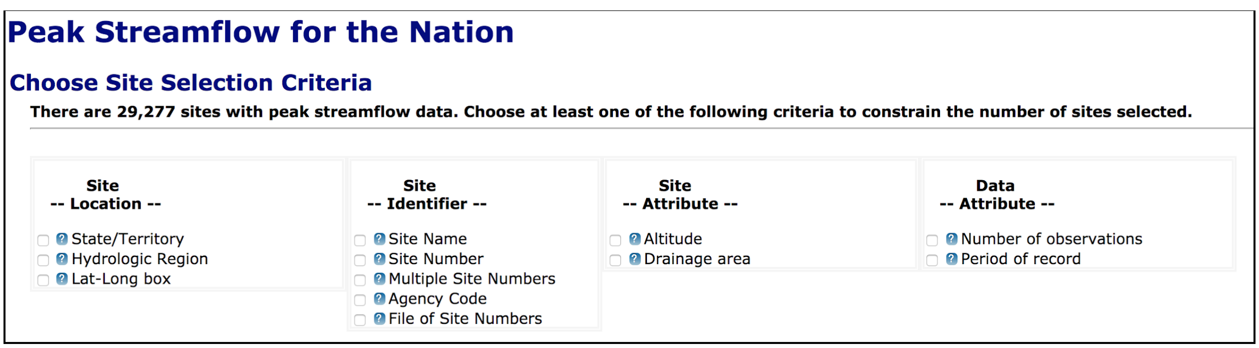
(USGS works with other federal, state and local agencies to manage approximately 7700 streamflow gauging stations in the United States.)

- Click Peak-Flow Data to see annual maximum instantaneous peak streamflow and gage height

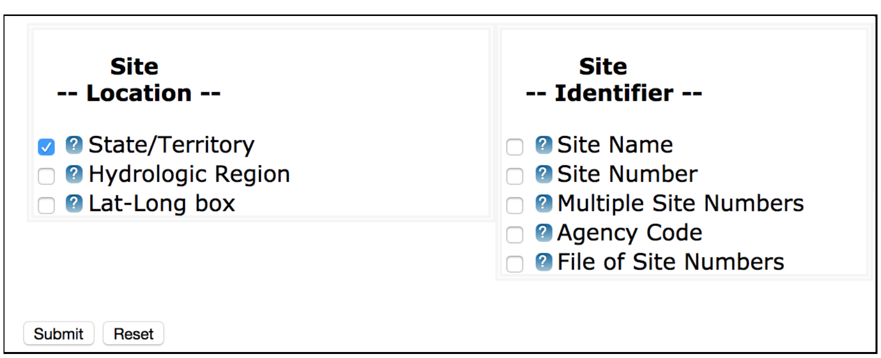


-
- Choose Site Selection Criteria: From the page, Peak Streamflow for the Nation, choose Site Selection Criteria,

There are 29,277 sites with peak streamflow data. Choose at least one of the following criteria to constrain the number of sites selected.



→ Select State/Territory and click Submit



Peak Streamflow for the Nation

Select sites which meet all of the following criteria:

Define one or more values for each of the following site-selection criteria: --- or select new criteria

State/Territory -- select one or more

New Hampshire
New Jersey
New Mexico
New York
North Carolina
North Dakota
Northern Mariana Islands

Choose Output Format

Display Summary of Selected Sites

Choose one of the following options for displaying descriptions of the sites meeting the criteria above:

- Show sites on a map NEW
- Table of sites grouped by
- Scroll list of sites -- allows selection of data for multiple sites
- Brief descriptions -- allows selection of data for multiple sites
- Site-description information displayed in XML format

  (Select fields to include in site-description output)

  Agency
  Site identification number
  Site name
  Site type

- Save file of selected sites to local disk for future upload

Retrieve Published peak streamflow data for Selected Sites

Choose one of the following options for displaying data for the sites meeting the criteria above

Retrieve data from: to: (YYYY-MM-DD -- **Blank = all data**)

Graphs of data — log scale

Table of data

Tab-separated data — YYYY-MM-DD — Save to file *

* Save compressed files with a .gz file extension.

peakfq (watstore) format — Save to file *

* Save compressed files with a .gz file extension.

Submit Reset Help