On the impact, identification and treatment of extraordinary floods in the systematic record

Richard M. Vogel

Tufts University, Medford, MA, USA

Email: richard.vogel@tufts.edu

July 13, 2020

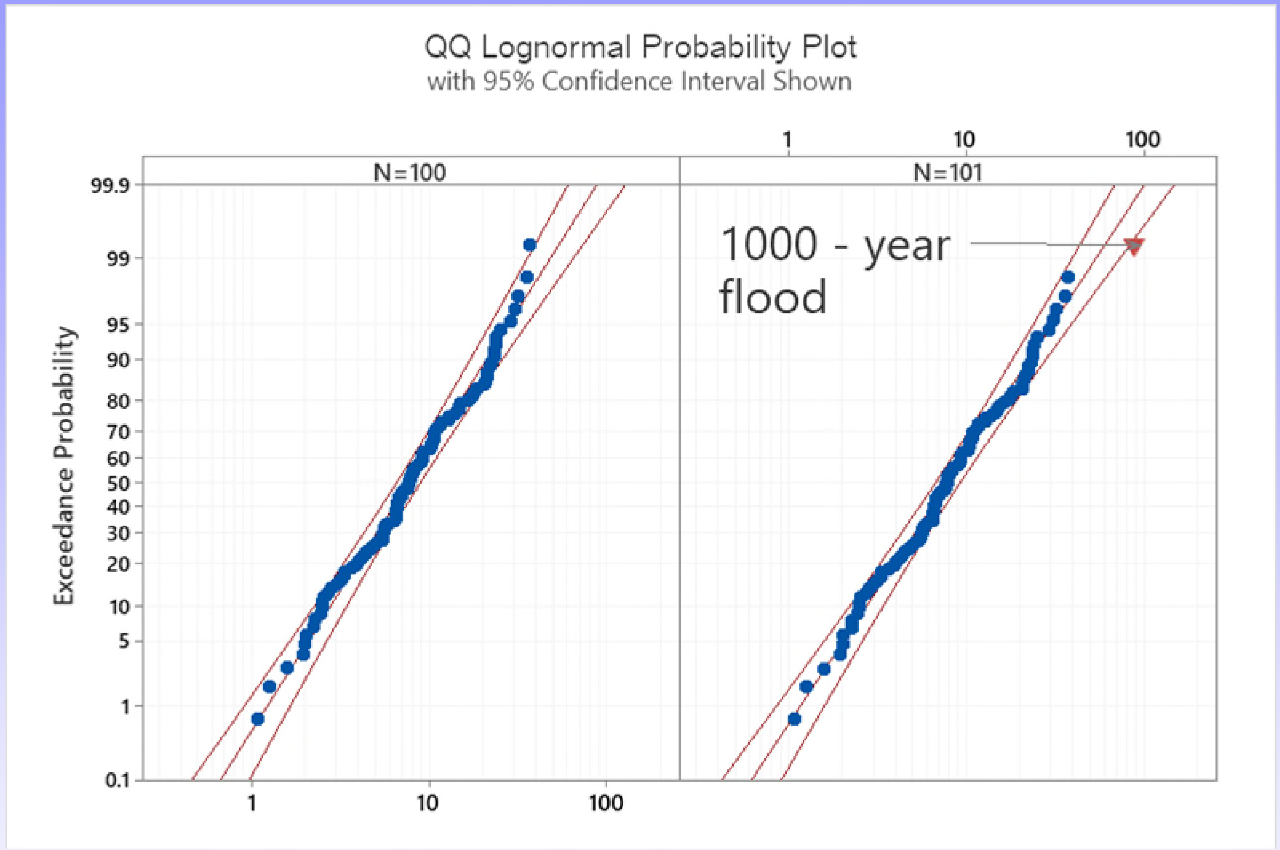
A taste of extremes in water science:

Preparing for July 2021 Summer School

Zoom Meeting

Tufts UNIVERSITY

A Synthetic Lognormal Flood Record, μ = 10, σ = 10 cms



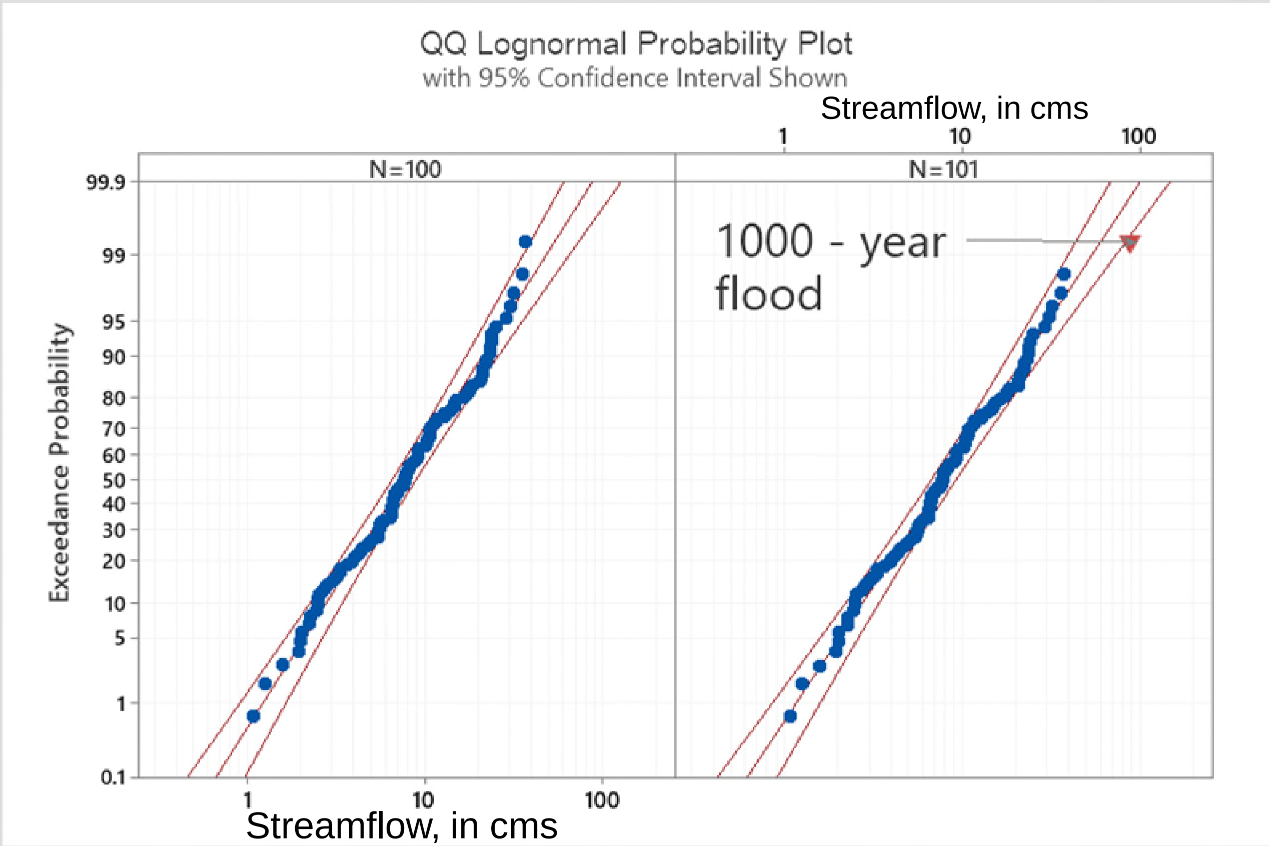
Lognormal n=100

Lognormal n=101 with 1000 yr flood

The 1000-year flood would be called the flood of record

It is the largest flood in n=101 years of systematic gaging

Background: Setting the Stage



Lognormal n=100 n=101 with 1000 year flood

How do we deal with situation on right?

Probability of at least one T=1000 year flood in 100 years = $1 - (1 - 0.001)^{100} = 0.095$

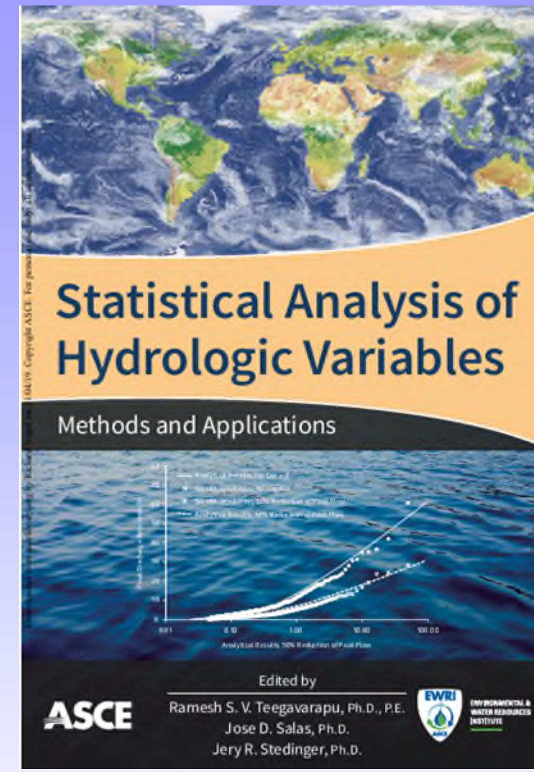
Probability of at least one T=100 year flood in 100 years = $1 - (1 - 0.001)^{100} = 0.634$

Theoretical Treatment of the Flood of Record

CHAPTER 12

Hydrologic Record Events

Richard M. Vogel
Attilio Castellarin
N. C. Matalas
John F. England, Jr.
Antigoni Zafirakou



Citation:
Vogel, R.M., N.C. Matalas, A. Castellarin and J.F. England, Hydrologic Record Events, Chapter 12 in Manual on Applications of Statistical Distributions in the Hydrologic Sciences, ASCE, 2019.

This chapter is a theoretical treatment of the behavior of the flood of record

As with all my papers, it can be downloaded from my website:

https://sites.tufts.edu/richardvogel/research/publications/

Tufts
UNIVERSITY

Nomenclature: Types of Flood Observations

- **Systematic Record**: Discharge and Stage (elevation) data collected at regular intervals, typically at gaging station.
- **Historic Record:** Flood events directly observed by nonhydrologists, in a nonsystematic manner. Usually occurred prior to the systematic record.

Figure 2. Photograph of historic flood high-water marks and flood of March 13–15, 2010, Potomac River at Great Falls Park, Virginia, upstream of U.S. Geological Survey streamflow-gaging station 01646500, Potomac River near Washington, D.C.

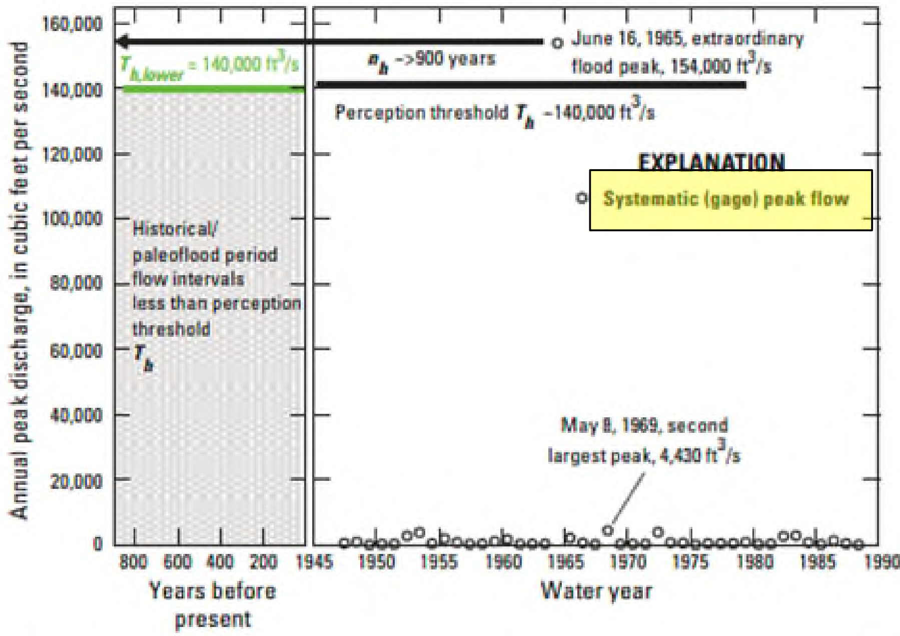


Figure 5. Graph showing an example site with an extraordinary flood peak that represents a longer time frame, Plum Creek near Louviers, Colorado, U.S. Geological Survey streamflow-gaging station 06709500. A scale break is used to separate the gaging station data from the longer historical/paleoflood period. Horizontal lines indicate the approximate historical period n_h and the perception threshold T_h.

Outliers and Spurious Observations

- **What is a Spurious Observation?**
  - Result from spurious error such as: measurement error, typographical error, etc.
- **What is a Flood Outlier?**
  - Outliers are potentially influential floods (PIF) that are exceedingly low or high compared to the distributional properties of the vast majority of the data (England et al. 2018)

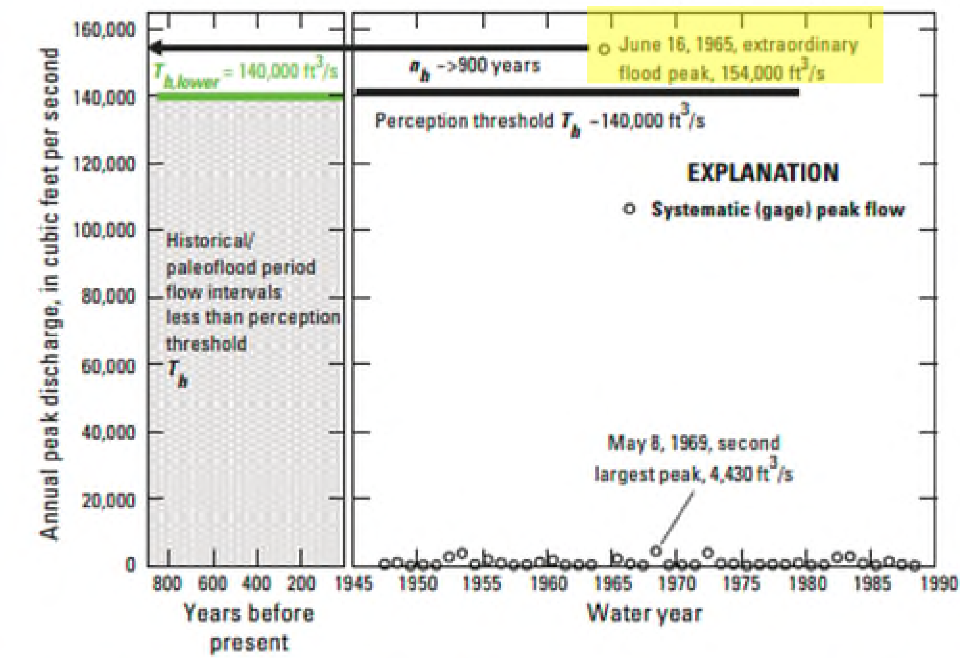


Figure 5. Graph showing an example site with an extraordinary flood peak that represents a longer time frame, Plum Creek near Louviers, Colorado, U.S. Geological Survey streamflow-gaging station 06709500. A scale break is used to separate the gaging station data from the longer historical/paleoflood period. Horizontal lines indicate the approximate historical period n_h and the perception threshold T_h.

Tufts UNIVERSITY

Example: Hamdi et al. 2018, NHESS.

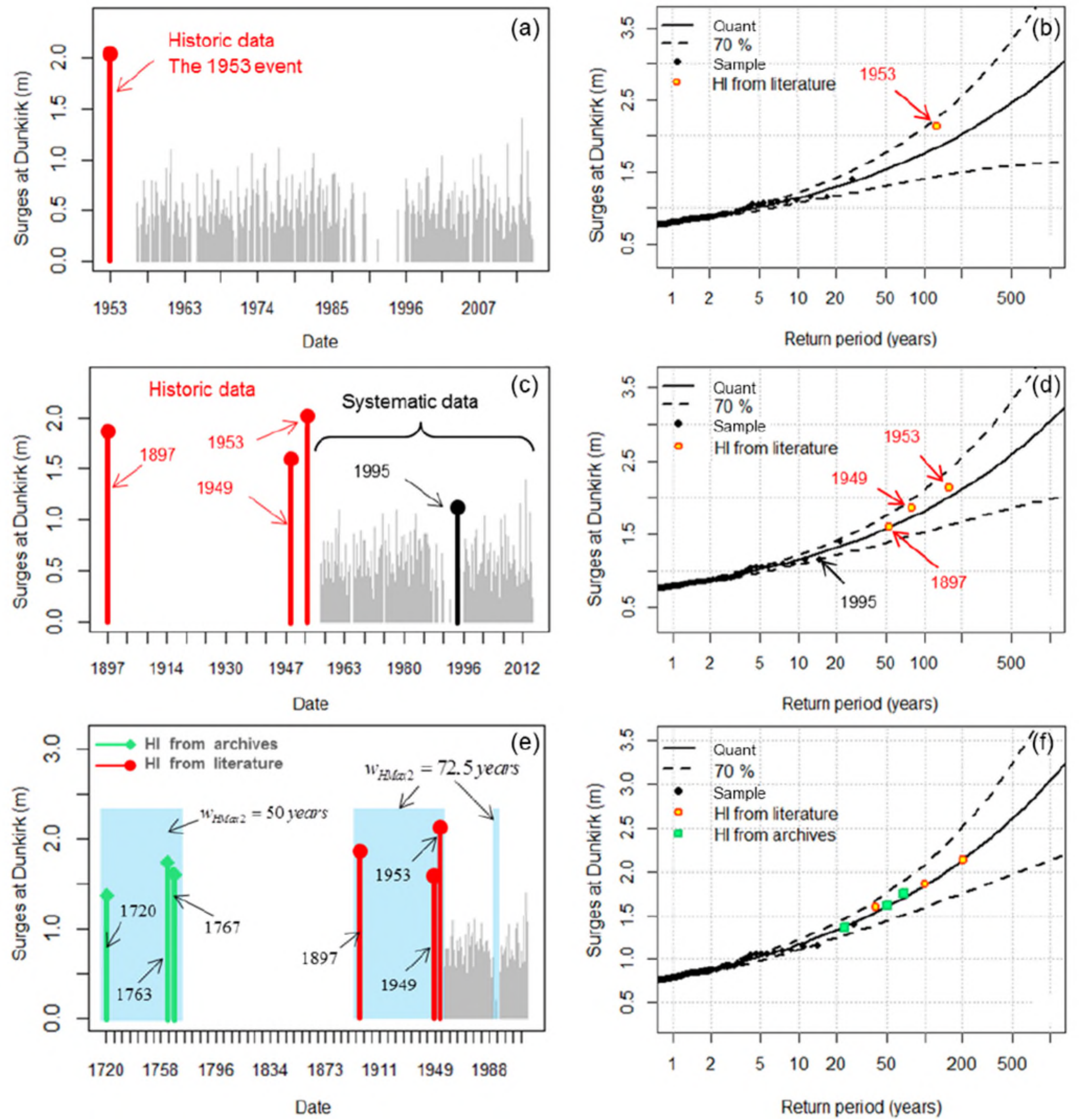
\

Generalized Pareto fit to POT series of storm surges, Dunkirk, France

(a, b) the 1953 event as historical data;

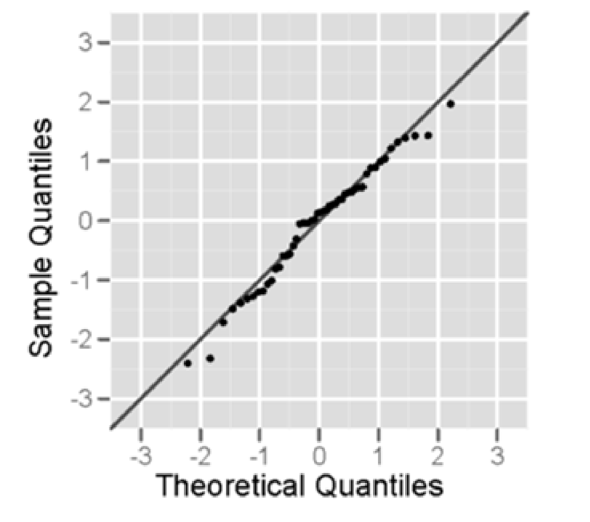
(c, d) historical data from literature;

(e, f) data from literature and archives.



Extraordinary Events Can Dominate Flood and Drought Frequency Analysis

- QQ Probability plots can assess goodness of fit of a theoretical probability distribution
- (QQ = Quantile-Quantile)
- Probability Plot Correlation Coefficient (PPCC) is a common goodness of fit statistic
- PPCC= correlation between ordered observations and estimate of ordered observations



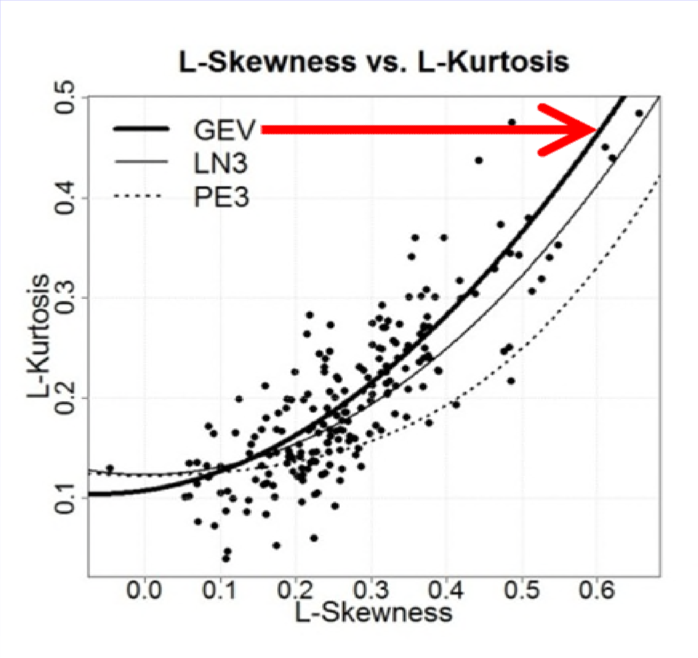
Consider 200 USGS sites with 85-127 years of annual maximum floods

Dataset used by Hirsch and Ryberg (HSJ, 2012)

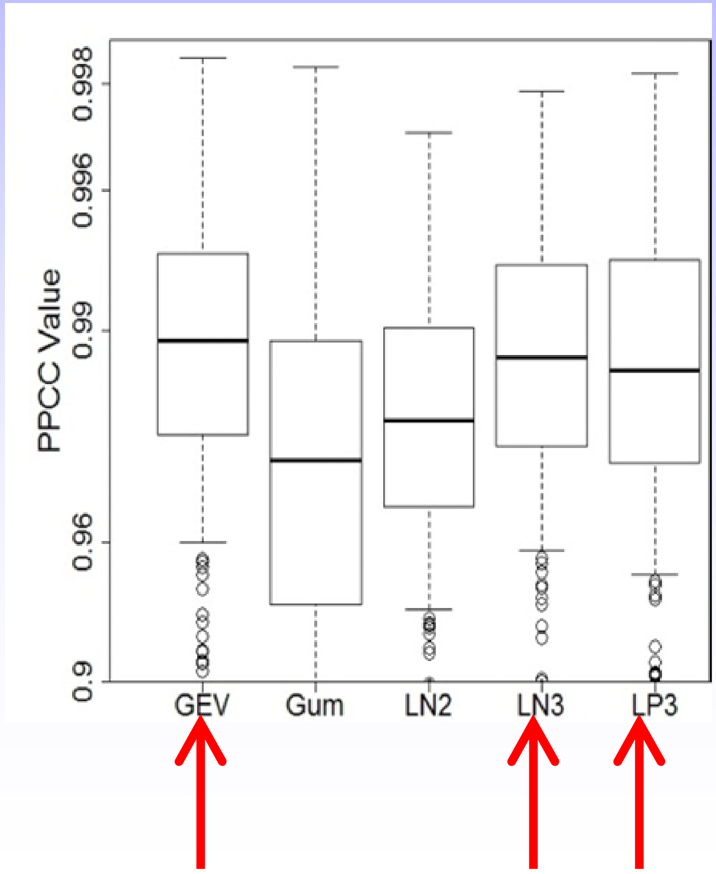
Tufts UNIVERSITY

Goodness of Fit of Various Probability Distributions

- **Lmoment Diagram indicates GEV provides a good fit**



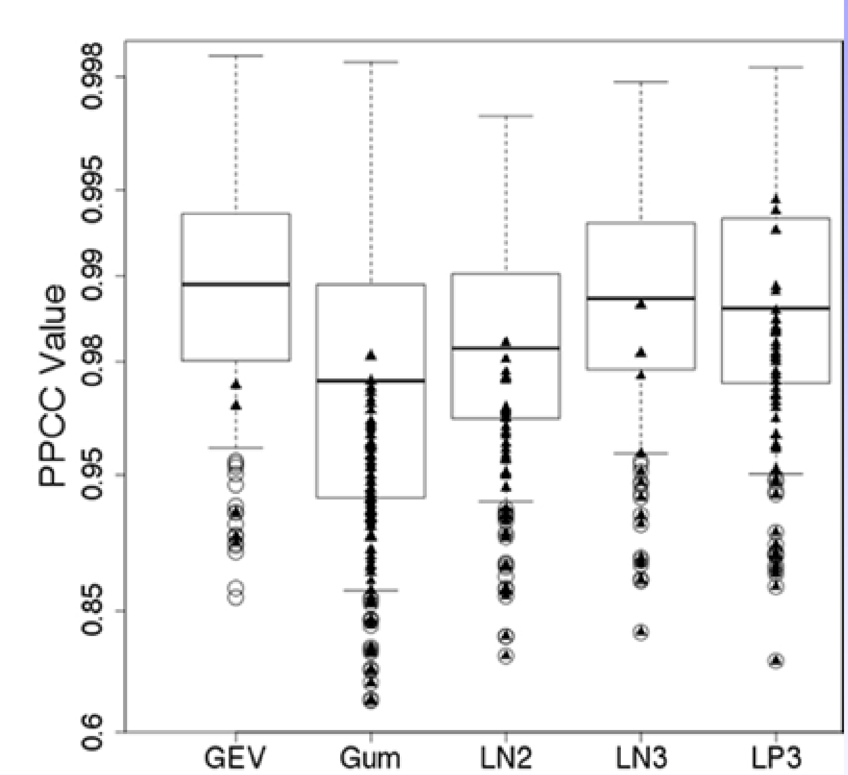
- **PPCC Values for GEV, LN3 and LP3 indicate excellent fits**



Tufts UNIVERSITY

Samples With Important Floods Always Lead to Poor Goodness of Fit!

- **Hirsch and Ryberg (2012)**
  - **200 Sites (85 < n < 127 years)**
- **Open Circles:**
  - **Outliers based on Grubbs (1969)**
- **Solid Triangles:**
  - **Observations > 1000 yr flood**
- **Regardless of the model considered, samples with observations larger than 1000-yr flood (solid triangles) had the LOWEST VALUES OF PPCC**



Example from: Boehlert, PhD Dissertation, Tufts Univ., 2015.

Tufts UNIVERSITY

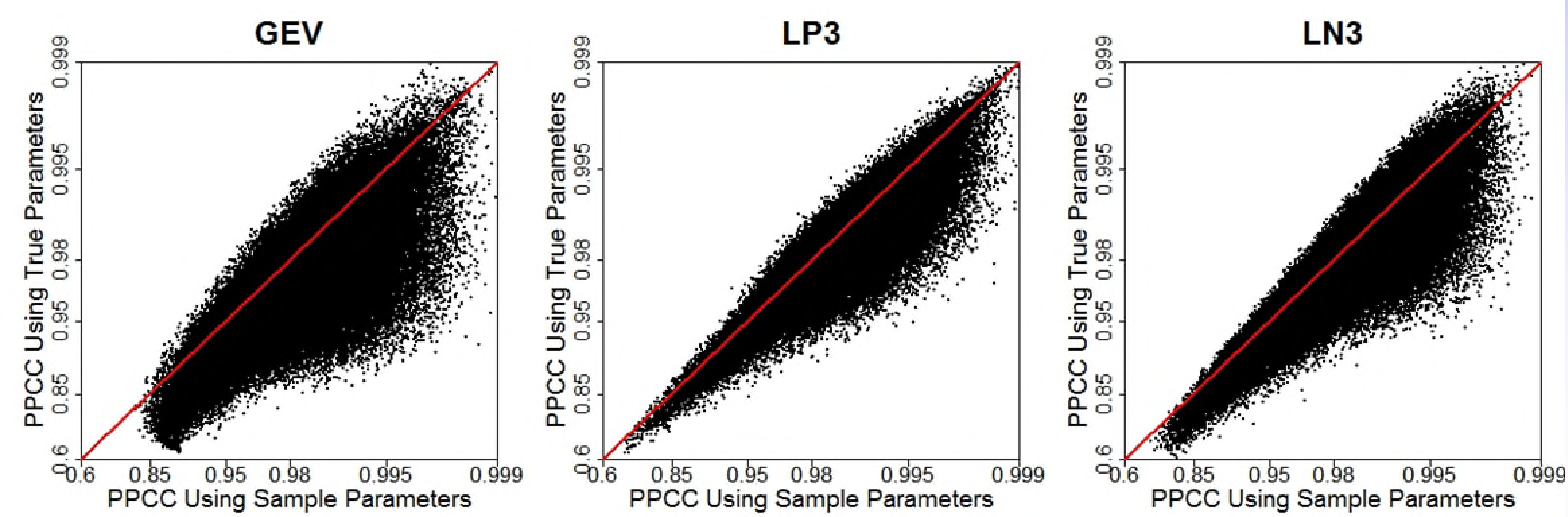
Goodness-of-Fit Can Be VERY misleading!

- **Experiment:**

  Generate 10,000 samples each of length 100 from GEV, LP3 and LN3 distributions.

  Compare goodness-of-fit using 'known' or 'true' parameters with goodness-of-fit using at-site sample estimates.

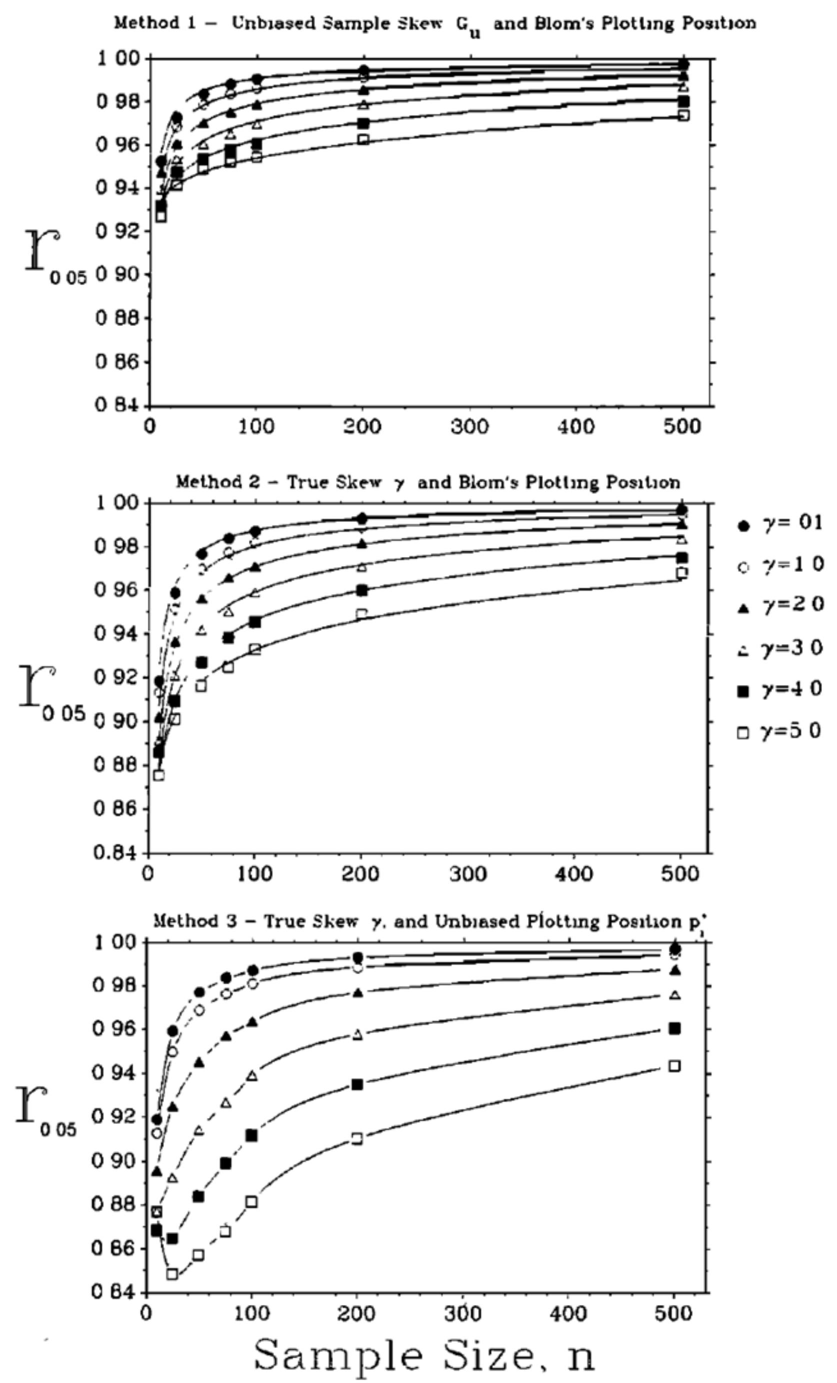
Result: Fitting true model degrades goodness-of-fit!



Example from: Boehlert, PhD Dissertation, Tufts Univ., 2015.

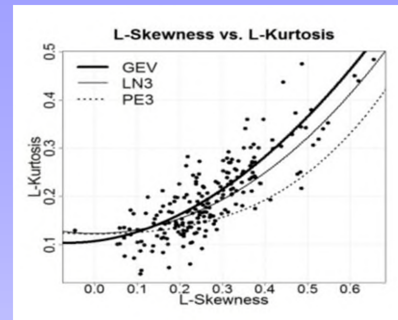
PPCC Test for Log Pearson type 3 (LP3)

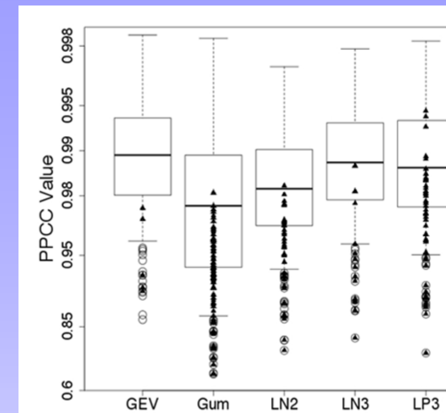
- $r_{0.05}$ is value of PPCC for LP3 at 5% significance level
- Note how goodness-of-fit is inflated when true skew must be estimated
- Goodness of fit can be misleading



From: Vogel, R. and McMartin, D. 1991. Probability plot goodness-of-fit and skewness estimation procedures for the Pearson type 3 distribution. *Water Resour. Res.* 27: 3149–3158.

Downward Bias in GEV Quantile Estimates





- Hirsch and Ryberg 200 Sites with (85 < n < 127)
  - In the **19,000** observations, we expect **19 - 1000-yr** floods (95% confidence interval: 12-26)
  - Yet only **6 – 1000-year** floods were obtained from GEV at-site models, (6/19) = 30%
  - Suggests that GEV exhibits downward bias
  - Why?

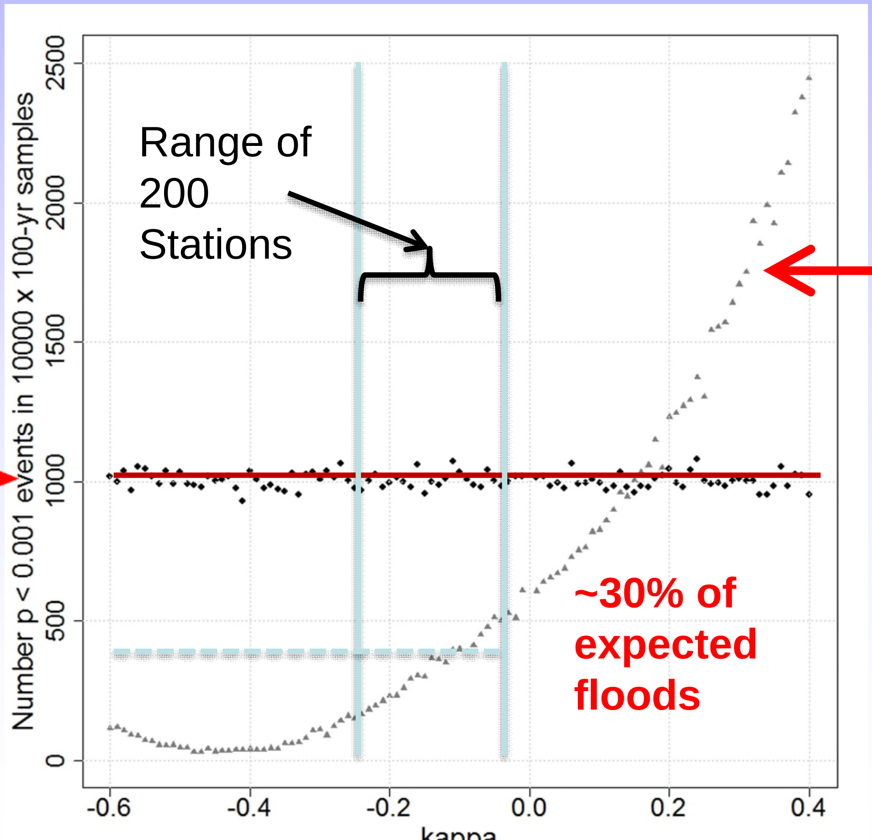
Tufts UNIVERSITY

Downward Bias Associated with Extraordinary Floods – WHY? – Its all about GEV shape parameter kappa κ

Experiment:

Generate 10,000 samples of length 100 years and count the number of 1000 year floods

Kappa κ, known (black– NO bias

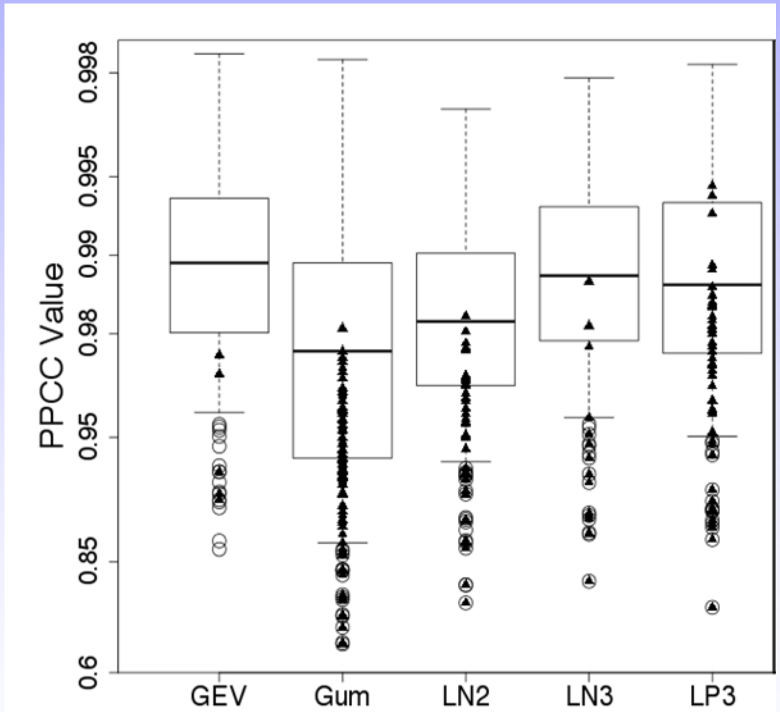


Kappa κ, unknown (gray) – downward bias

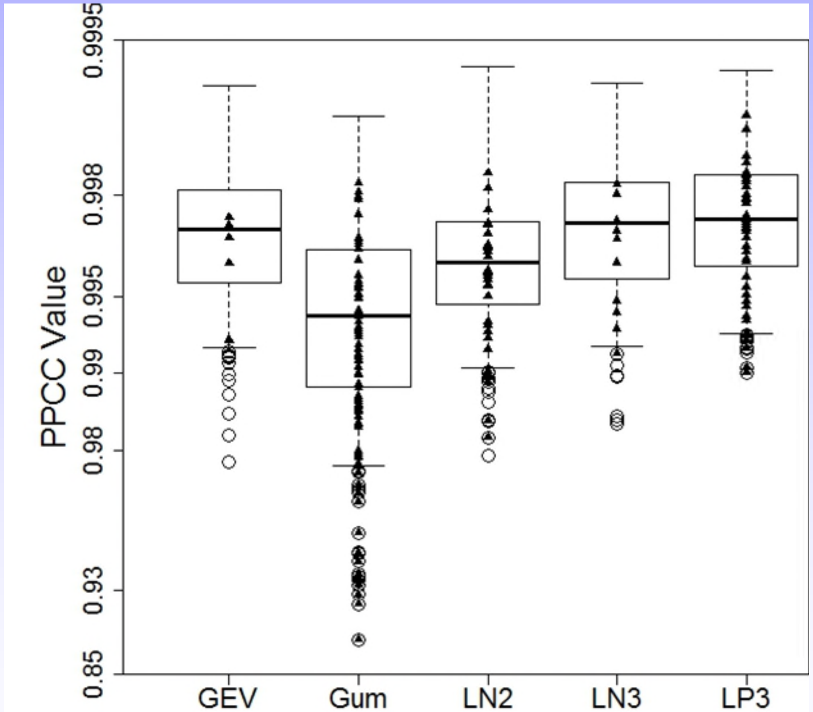
Example from: Boehlert, PhD Dissertation, Tufts Univ., 2015.

Recommendation use P-P plots rather than Q-Q plots

PPCC using Q-Q plot



PPCC using P-P plot



Example from: Boehlert, PhD Dissertation, Tufts Univ., 2015.

Summary

- **Goodness of fit can be misleading:**
  - Fitted models tend to ‘look’ better than correct/true model.
  - Do not base decisions on ‘goodness of fit’ alone unless P-P plots are used instead of Q-Q plot

- **Reliable shape parameter(s) are critical for design flood/drought estimation**

- **Treatment of high outliers is critical for design flood/drought estimation**

Possible Paper Titles

- On the impact, identification and treatment of extraordinary floods in the systematic record
- P-P Probability Plots and Hypothesis Tests
- Goodness-of-fit is Misleading
- Improvements in Estimation of Shape Parameter for Flood Frequency Analysis

Tufts UNIVERSITY

Other Possible Paper Titles

- Accounting for stochastic persistence and deterministic trends when updating a flood and drought frequency analysis

- On the recurrence interval of the ‘drought of record’

- Methods for generation of stochastic ensembles of extreme events using deterministic simulation models.

Tufts UNIVERSITY

References

England, J.F., Jr., Cohn, T.A., Faber, B.A., Stedinger, J.R., Thomas, W.O., Jr., Veilleux, A.G., Kiang, J.E., and Mason, R.R., Jr., 2018, Guidelines for determining floo flo frequency—Bulletin 17C: U.S. Geological Survey Techniques andMethods, book 4, chap. B5, 148 p., https://doi.org/10.3133/tm4B5 .

Gen, F. and K. Koehler. 1990. Goodness-of-Fit Tests Based on P-P Probability Plots. *Technometrics.* 32(3): 289-303.

Grubbs, F. 1969. Procedures for detecting outlying observations in samples, *Technometrics*, 11, 1–21.

Hirsch, R.M. and Ryberg, K.R., 2012, Has the magnitude of floods across the USA changed with global CO2 levels?, *Hydrological Sciences Journal,* 57(1).

Hosking, J. 1990. L-moments: Analysis and Estimation of Distributions using Linear Combinations of Order Statistics. *J. R. Statist. Soc. B*. 52(1): 105-124.

Hosking, J. and J. Wallis. 1997. Regional Frequency Analysis: An Approach Based on L-moments. Cambridge University Press.

Liao, F., P. Allamano, and P. Claps. 2010. Exploiting the information content of hydrological “outliers” for goodness-of-fit testing. *Hydrol. Earth Syst. Sci.* 14: 1909-1917. doi:10.5194/hess-14-1909-2010

Stedinger, J., Vogel, R., and Foufoula-Georgiou, E. 1992. Handbook of Hydrology, chap. 8: Frequency analysis of extreme events, McGraw-Hill, New York.

Vogel, R. and McMartin, D. 1991. Probability plot goodness-of-fit and skewness estimation procedures for the Pearson type 3 distribution. *Water Resour. Res.* 27: 3149–3158.

Vogel, R.M., N.C. Matalas, A. Castellarin and J.F. England, Hydrologic Record Events, Chapter 12 in Manual on Applications of Statistical Distributions in the Hydrologic Sciences, ASCE, 2019.