Extreme-Value Analysis

Eric Gilleland
Research Applications Laboratory
National Center for Atmospheric Research

2023 Clemson Climate Extremes Workshop

16 – 18 May 2023

NCAR | RESEARCH APPLICATIONS LABORATORY

DTC Developmental Testbed Center

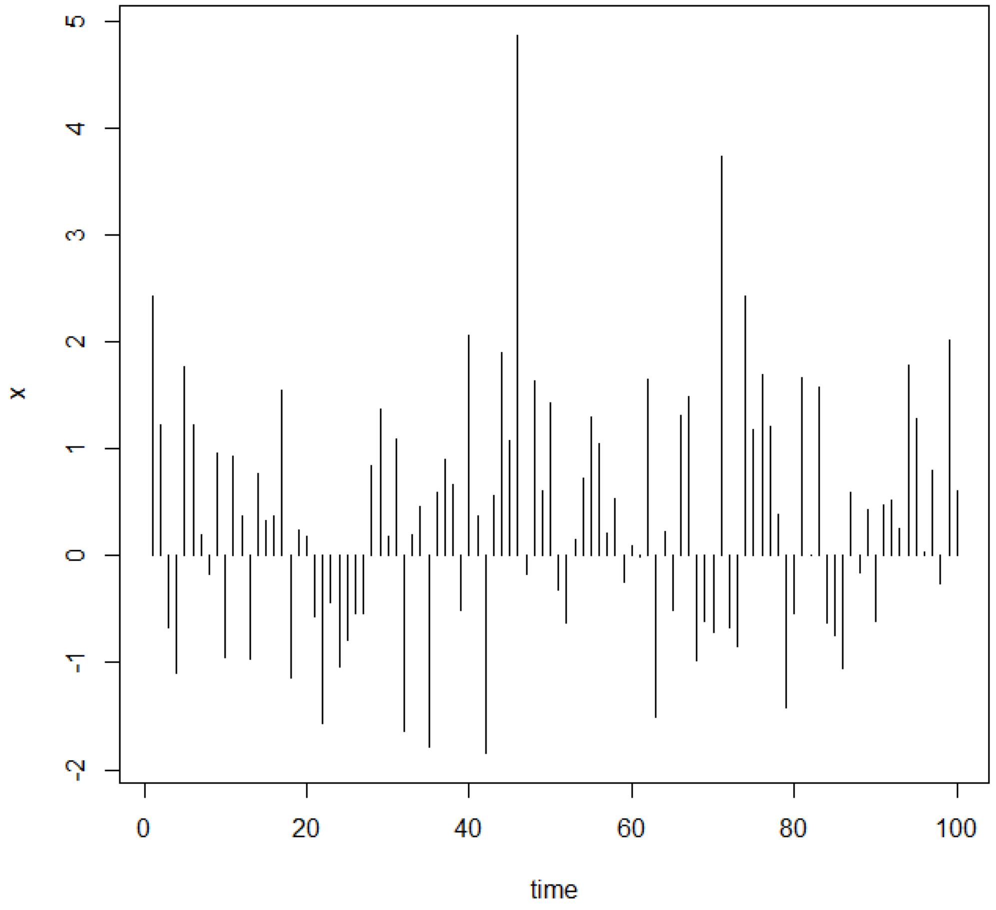
NSF

This material is based upon work supported by the National Center for Atmospheric Research, which is a major facility sponsored by the National Science Foundation under Cooperative Agreement No. 1852977.

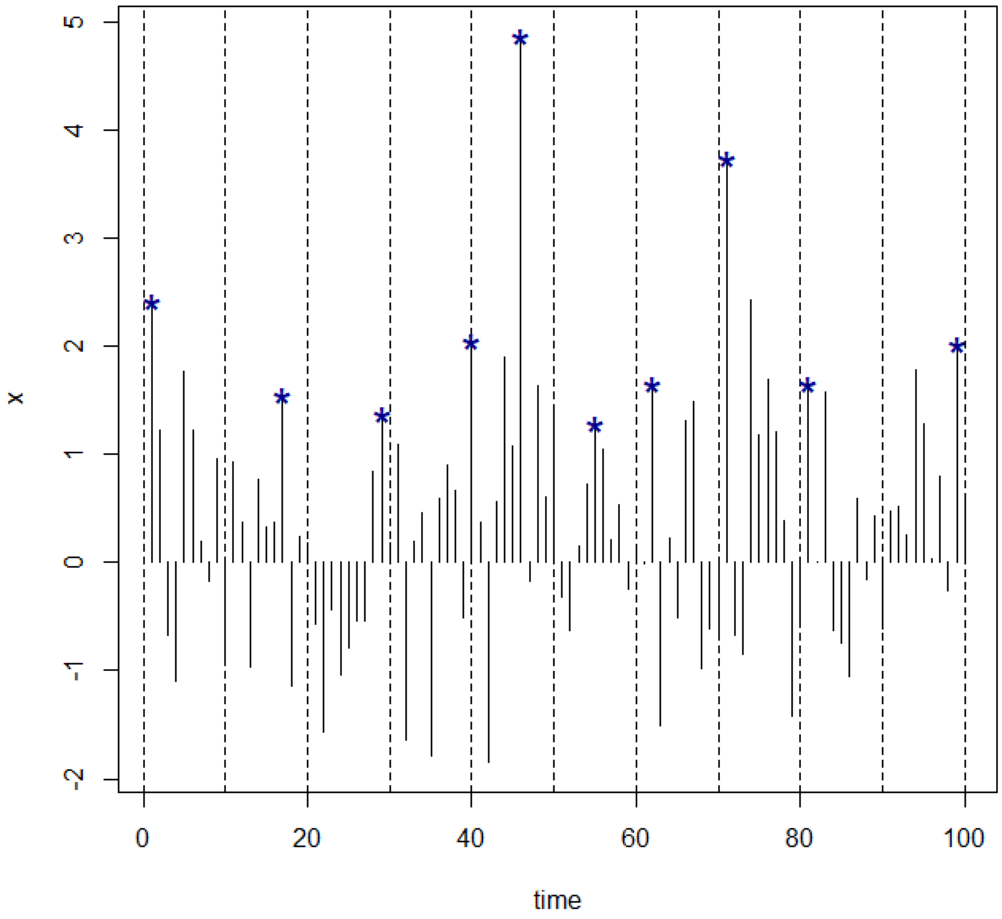
Extreme Values

Image citation: http://n2t.net/ark:/85065/d72v2d5b

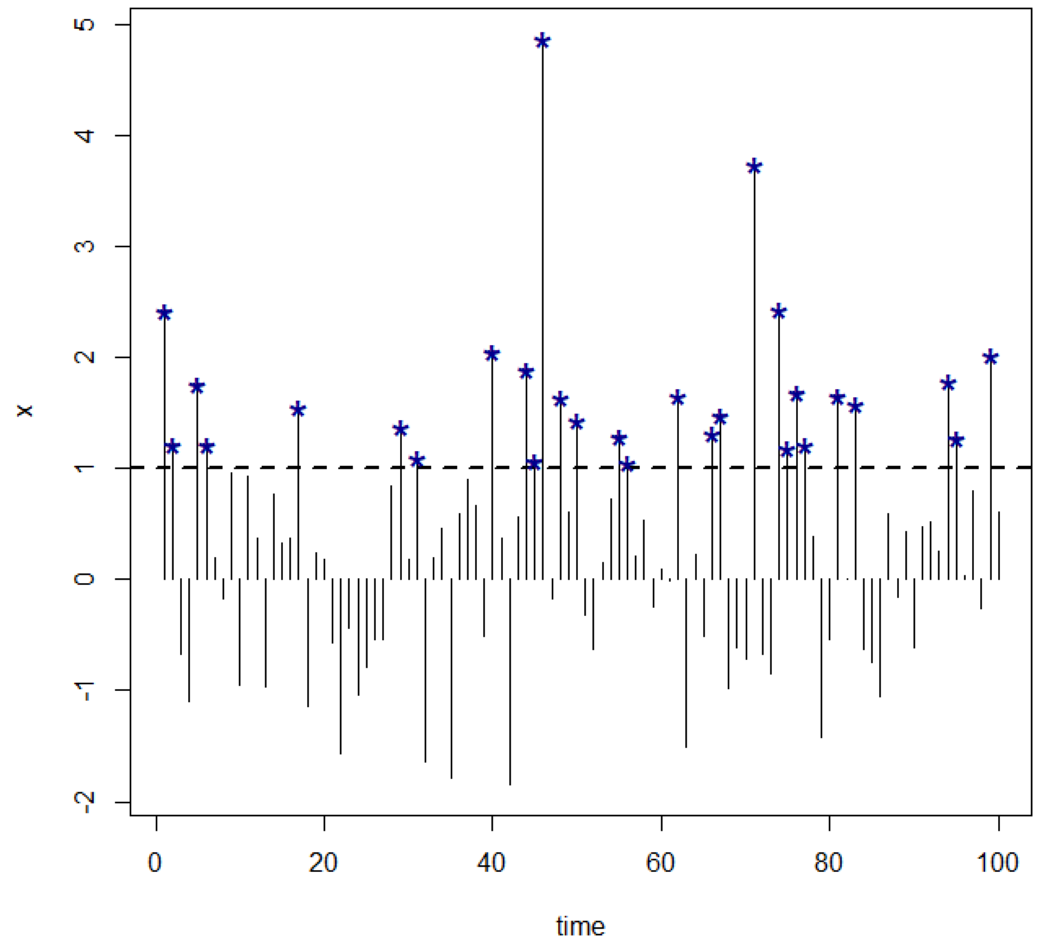
What is extreme?



What is extreme?

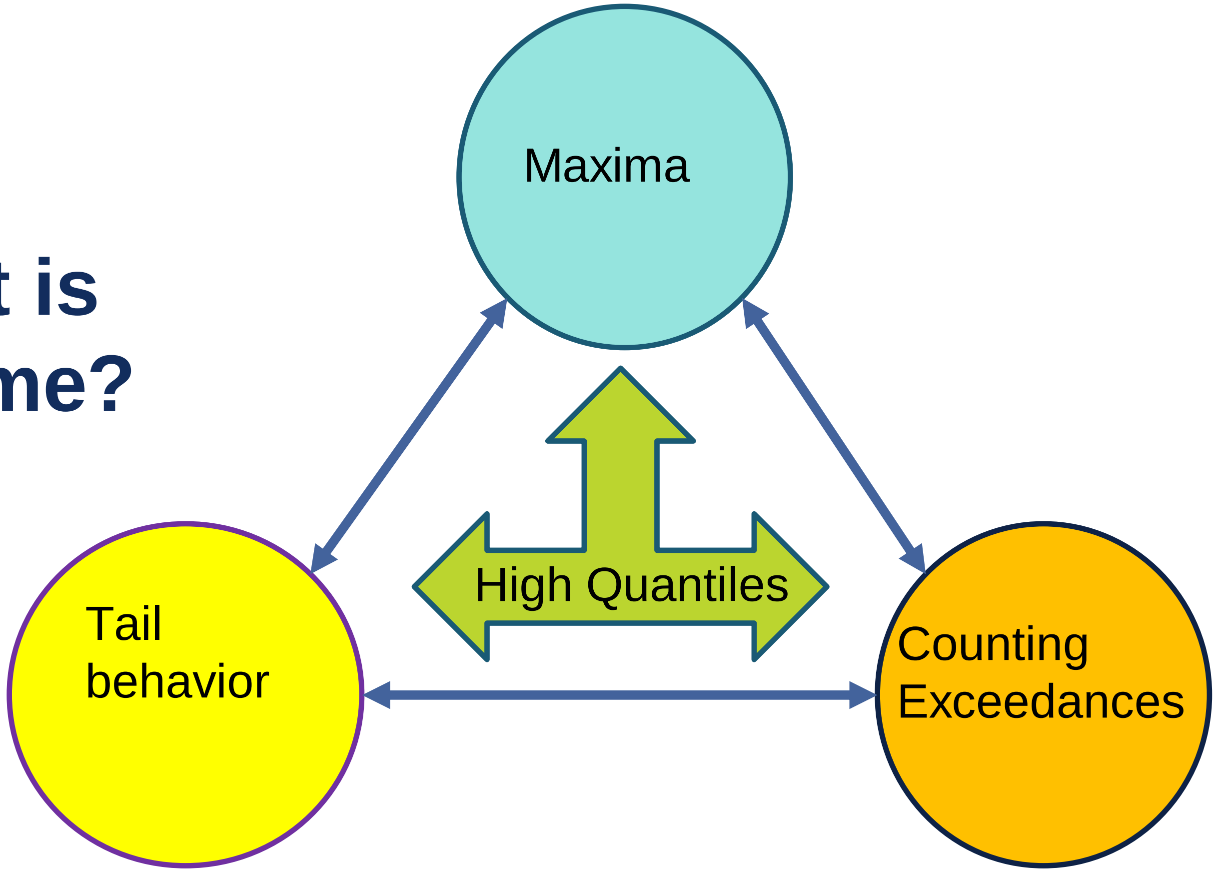


What is extreme?



Extreme-Value Analysis (EVA)

What is extreme?



Extreme-Value Analysis (EVA)

Suppose $X_1, X_2, \ldots, X_n$ are iid with distribution F, and we want to know $\mathbb{P}[M_n \leq z]$, where $M_n = \max\{X_1, X_2, \ldots, X_n\}$.

If $M_n \leq z$, then every $X_1, X_2, \ldots, X_n$ is $\leq z$. So,

$$\mathbb{P}[M_n \leq z] = \mathbb{P}[X_1 \leq z, X_2 \leq z, \ldots, X_n \leq z]$$

Then, because they are independent…

$$\mathbb{P}[X_1 \leq z, X_2 \leq z, \ldots, X_n \leq z] = \mathbb{P}[X_1 \leq z] \cdot \mathbb{P}[X_2 \leq z] \cdots \mathbb{P}[X_n \leq z]$$

And because they are identically distributed with distribution F…

$$\mathbb{P}[X_1 \leq z] \cdot \mathbb{P}[X_2 \leq z] \cdots \mathbb{P}[X_n \leq z] =$$

$$\mathbb{P}[X_1 \leq z] \cdot \mathbb{P}[X_1 \leq z] \cdots \mathbb{P}[X_1 \leq z] =$$

$$\mathbb{P}[X_1 \leq z]^n = F^n(z)$$

Extreme-Value Analysis (EVA)

But, there is a problem…

$0 \leq F \leq 1$, so that $F^n \rightarrow 0$ quickly and because we must estimate F, small errors in the estimation are exponentiated by the sample size!

Sum Stability

$$\frac{2+7+14+8+10+22}{6} = \frac{63}{6}$$

$$\frac{1}{2} \cdot \left(\frac{2+7+14}{3} + \frac{8+10+22}{3}\right) = \frac{1}{2} \cdot \left(\frac{23}{3} + \frac{40}{3}\right) = \frac{63}{6}$$

The two averages are the same. The first averages six numbers while the second averages only two.

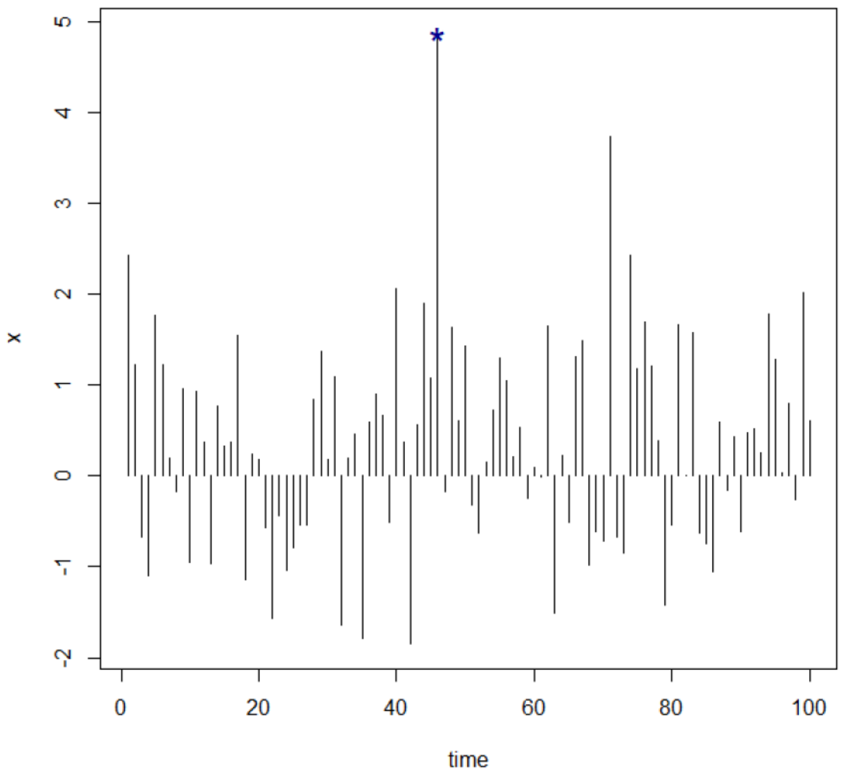
Sum Stability

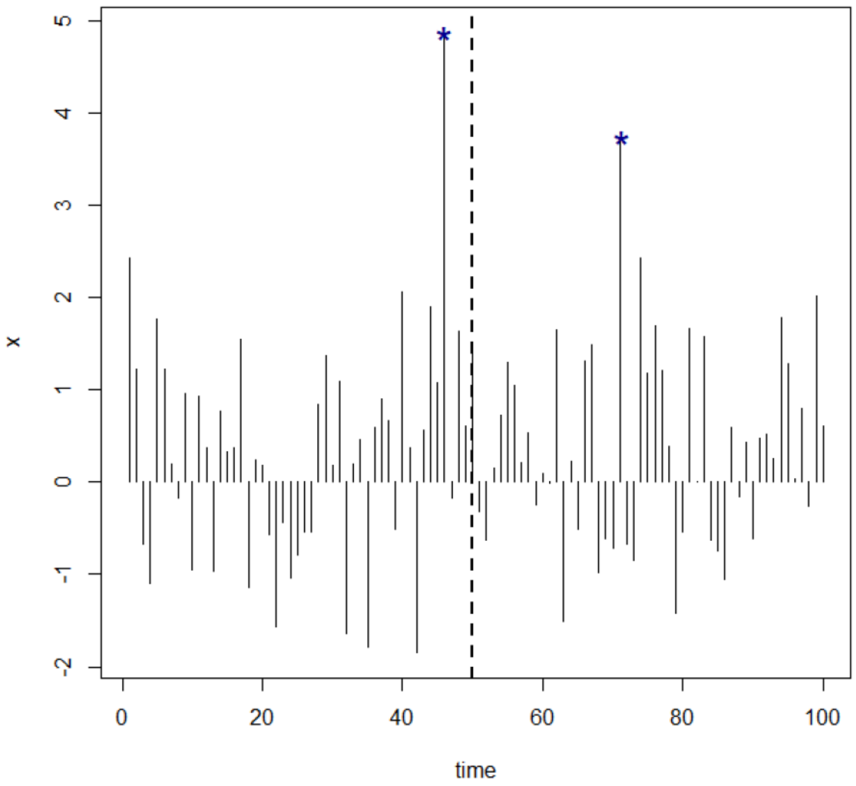
Let $X_1, X_2, \dots, X_n$ be iid random variables. Let $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ and let $\bar{Y}_m = \frac{1}{mk}\sum_{j=1}^{m} Y_j$, where

$$Y_1 = X_1 + X_2 + \cdots + X_k, Y_2 = X_{k+1} + X_{k+2} + \cdots + X_{2k}, \dots, Y_m = X_{n-k+1} + \cdots + X_n.$$

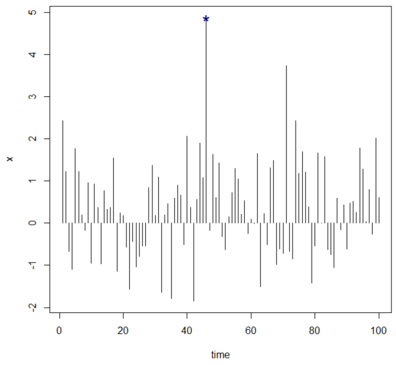
The distribution of $\bar{X}_n$ should be the same as that of $\bar{Y}_m$. And sum stability says that it is! In this case, the normal distribution is sum-stable (among others). For example, if $X_i \sim N(\mu, \sigma^2)$ for all $i = 1, \dots, n$, then $\sum_i^n X_i \sim N(n\mu, n\sigma^2)$. Note the distribution is the same but for a change in location and some re-scaling.

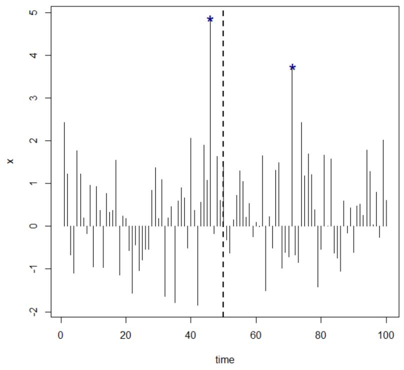
Max Stability





Max Stability





Similar to the sum-stable case,

$$\max\{X_1, \dots, X_n\} = \max\left\{\max\left\{X_1, \dots, X_{\lfloor \frac{n}{2} \rfloor}\right\}, \max\left\{X_{\lfloor \frac{n}{2} \rfloor + 1}, \dots, X_n\right\}\right\} =$$

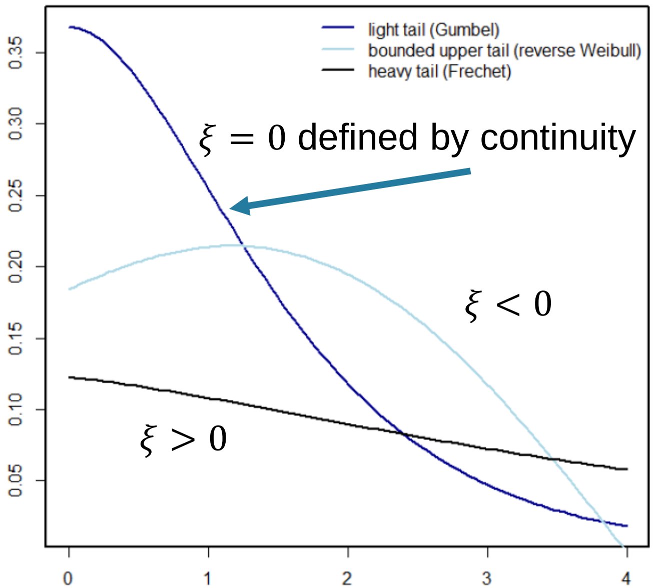
$$\max\{\max\{X_1, X_2, X_3\}, \max\{X_4, X_5, X_6\}, \dots, \max\{X_{n-2}, X_{n-1}, X_n\}\}$$

and so on.

Max Stability

So, apart from some re-scaling, we seek a distribution G such that $G(a_n z + b_n) = G^n(z)$ for some sequences of constants $a_n > 0$ and b_n.

Such a distribution is said to be max-stable, and there is only one (family) that fits the bill! The generalized extreme-value distribution (GEV).



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

$$E[Z] = \mu - \frac{\sigma\big(1 - \Gamma(1-\xi)\big)}{\xi}, \qquad \xi < 1$$

$$\mathbb{V}[Z] = \frac{\sigma^2\big(\Gamma(1-2\xi) - \Gamma^2(1-\xi)\big)}{\xi^2}, \qquad \xi < \frac{1}{2}$$

Extreme-Value Analysis (EVA)

$$G(z;\mu,\sigma,\xi) =$$

$$\exp\left\{-\left[1+\frac{\xi}{\sigma}(z-\mu)\right]^{-1/\xi}\right\}\cdot I_{(-\infty,\mu-\sigma/\xi]}(z)\cdot I_{(-\infty,0)}(\xi) +$$

$$\exp\left\{-\exp\left(\frac{z-\mu}{\sigma}\right)\right\}\cdot I_{\{0\}}(\xi) +$$

$$\exp\left\{-\left[1+\frac{\xi}{\sigma}(z-\mu)\right]^{-1/\xi}\right\}\cdot I_{[\mu-\sigma/\xi,\infty)}(z)\cdot I_{(0,\infty)}(\xi)$$

Extreme-Value Analysis (EVA)

$$G(z;\mu,\sigma,\xi) =$$

Upper bound
(reverse) Weibull distribution ($\xi < 0$)

$$\exp\left\{-\left[1+\frac{\xi}{\sigma}(z-\mu)\right]^{-1/\xi}\right\}\cdot I_{(-\infty,\mu-\sigma/\xi]}(z)\cdot I_{(-\infty,0)}(\xi) +$$

Light tail Gumbel distribution
defined by continuity as $\xi \to 0$

$$\exp\left\{-\exp\left(\frac{z-\mu}{\sigma}\right)\right\}\cdot I_{\{0\}}(\xi) +$$

$$\exp\left\{-\left[1+\frac{\xi}{\sigma}(z-\mu)\right]^{-1/\xi}\right\}\cdot I_{[\mu-\sigma/\xi,\infty)}(z)\cdot I_{(0,\infty)}(\xi)$$

Heavy-tail Fréchet distribution ($\xi > 0$)

GEV Return Levels

A useful, if often misunderstood, quantity that is readily obtained from the GEV distribution is the T-year return level.

A T-year return level is the value expected to be observed, on average, once every T years.

It is the $1 - \frac{1}{T}$ quantile of the GEV distribution.

GEV Return Levels

Let $p = \frac{1}{T}$, and let $y_p = -\frac{1}{\log(1-p)}$, then the associated return level, z_p, is given by

$$z_p = \begin{cases} \mu + \frac{\sigma}{\xi}\left[y_p^{\xi} - 1\right], \xi \neq 0 \\ \mu + \sigma \log y_p , \xi = 0 \end{cases}$$

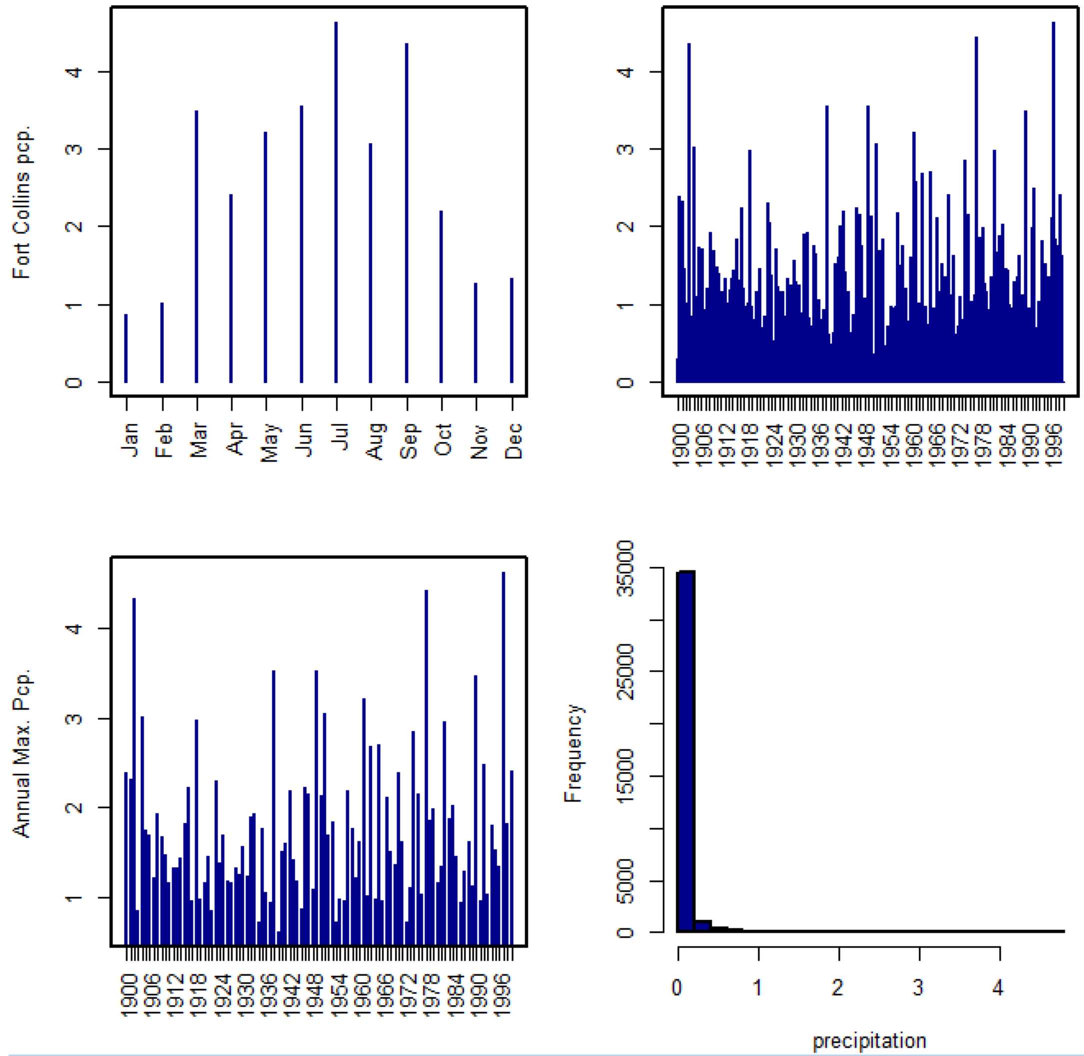
Plotting z_p against $\log y_p$ gives a plot that is:

- A straight line if $\xi = 0$ (and close to a straight line if $\xi \approx 0$).
- A concave curve with no finite upper bound if $\xi > 0$.
- A convex curve with an asymptote at the upper limit $\mu - \frac{\sigma}{\xi}$ as $p \to 0$ if $\xi < 0$.

GEV Return Levels

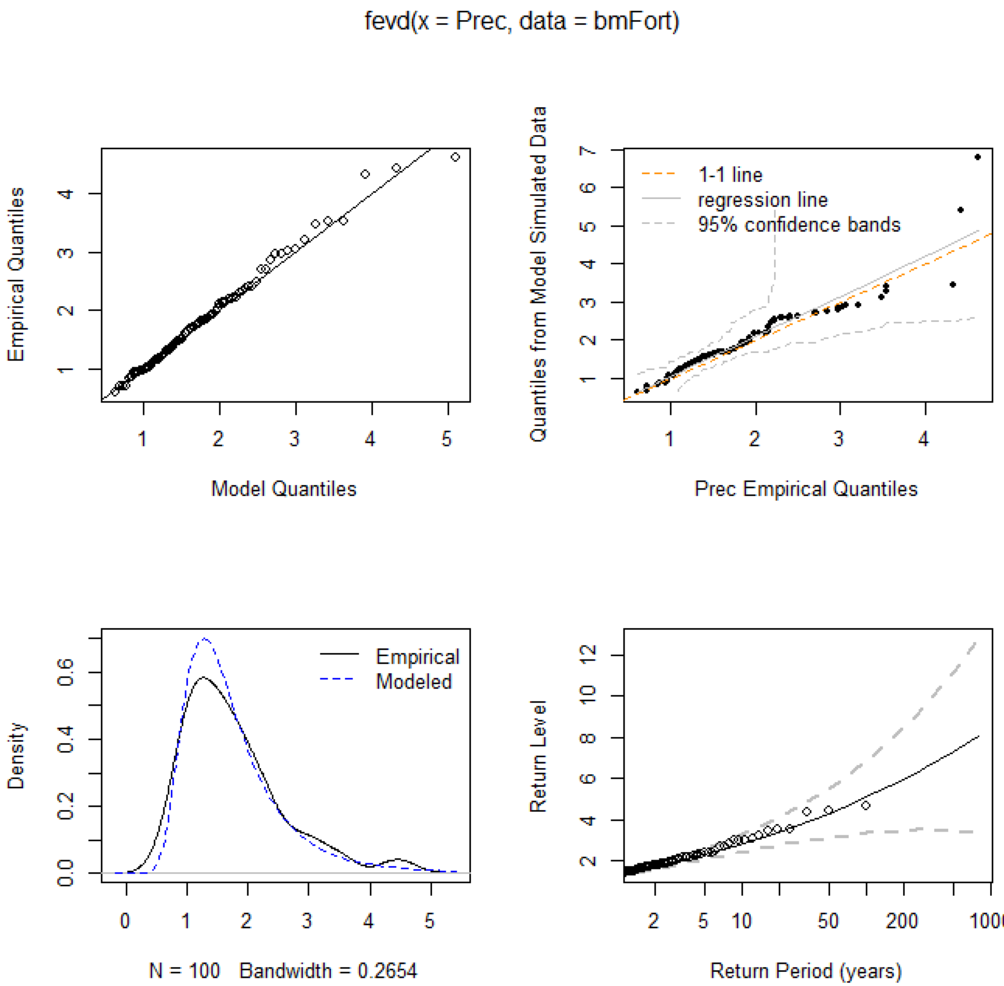
extRemes: An R software package

data("Fort")



Fit GEV to block maxima

```
data( “Fort” )
plot( Prec ~ obs, data = Fort, col = “darkblue”,
   type = “h”, lwd = 1.5 )
bmFort <- blockmaxxer(Fort, which = 6,
   blocks = Fort$year, units = “inches/100”  )

plot( Prec ~ year, data = bmFort,
   col = "darkblue", lwd = 1.5, type = "h" )

fit0 <- fevd( Prec, data = bmFort )
fit0
plot( fit0 )

ci( fit0 )
ci( fit0, type = "parameter" )
```



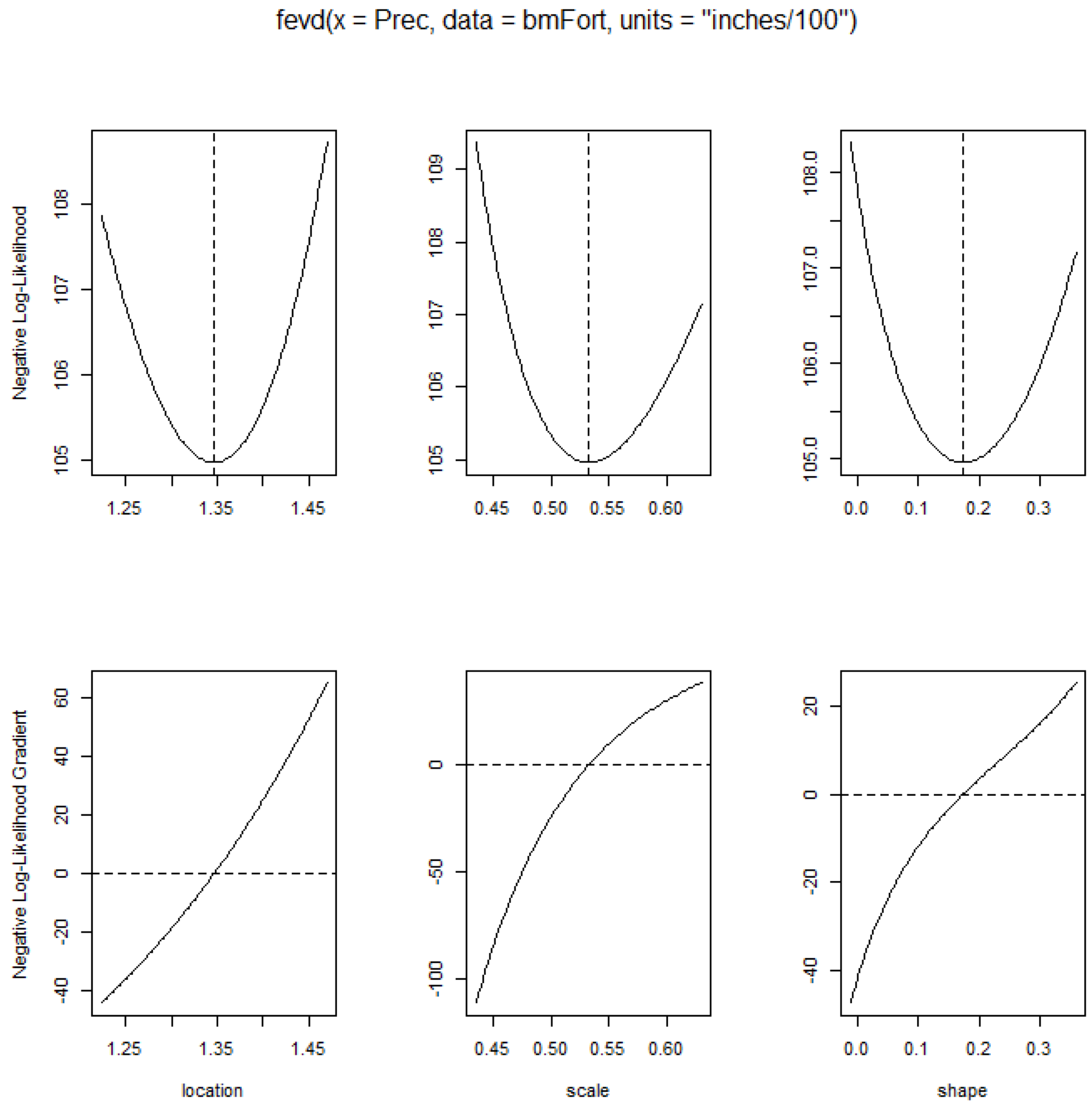
Fit GEV to block maxima

extRemes: An R software package

```
fevd(x = Prec, data = bmpcp)

[1] "Estimation Method used: MLE"


 Negative Log-Likelihood Value:  104.9645


 Estimated parameters:
 location     scale     shape
1.3466597 0.5328046 0.1736264

 Standard Error Estimates:
  location      scale      shape
0.06168793 0.04878843 0.09195458

 Estimated parameter covariance matrix.
            location        scale        shape
location  0.003805401  0.0017067043 -0.0020838301
scale     0.001706704  0.0023803113 -0.0008692638
shape    -0.002083830 -0.0008692638  0.0084556445

 AIC = 215.9291

 BIC = 223.7446
```

Fit GEV to block maxima

plot(fit0, type = “trace”)



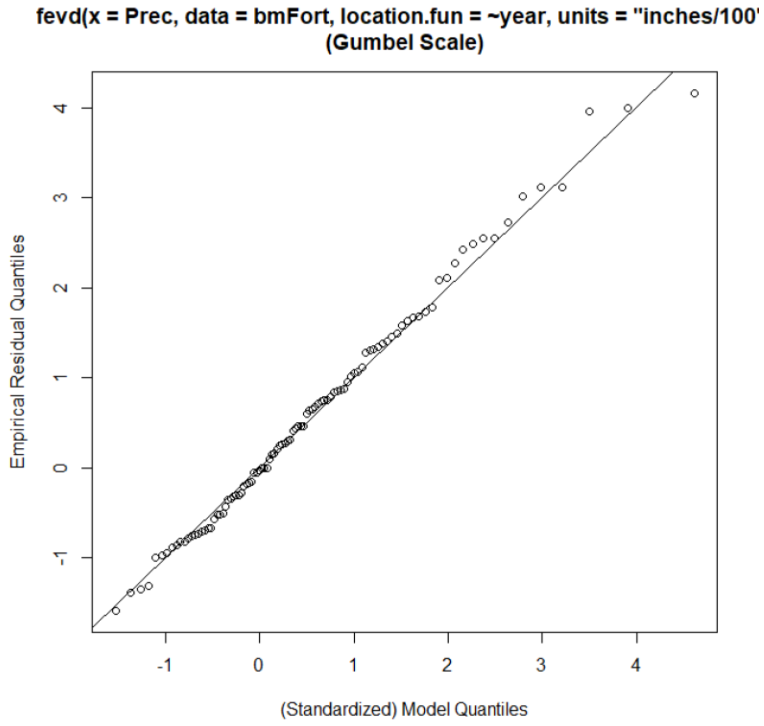
Fit GEV to block maxima

Test for a trend in the maxima by fitting a GEV whose location parameter has a linear trend with year (i.e., $\mu(t) = \mu_0 + \mu_1 t$, where t is the year):

```
fit1 <- fevd( Prec, data = bmFort,
   location.fun = ~year , units = "inches/100" )
fit1
plot( fit1, type = "qq" )

lr.test( fit0, fit1 )
```

Likelihood-ratio Test

data: PrecPrec
Likelihood-ratio = 0.13922, chi-square critical value = 3.8415, alpha = 0.0500, Degrees of Freedom = 1.0000, p-value = 0.7091
alternative hypothesis: greater



No significant trend

Modeling block maxima

GEV distribution

Fit the GEV distribution to block maxima where the blocks are long.

Advantages of modeling block maxima

- Usually do not need to worry about diurnal trends, or other cyclic behavior,
- Usually do not need to worry about temporal dependence,
- Quantiles are easy to find and are equivalent to the T-year return level.

Counting Exceedances

Suppose, again, $X_1, X_2, \ldots, X_n$ are iid with distribution F.

Let u represent a high value and suppose we are interested in $\mathbb{P}[X_i > u] = 1 - F(u), i = 1, \ldots, n$.

Then $1 - F(u)$ can be thought of as the probability of "success" for a binomial distribution.

If $1 - F(u) \to 0$ fast enough that the expected number of successes is constant, then the Poisson distribution is a good approximation.

Note that if $N \sim \text{Poisson}(\lambda)$, with N the number of events where $X_i > u$, for some large constant threshold u, then $\mathbb{P}[\max\{X_1, \ldots, X_n\} < u] = \mathbb{P}[N = 0] = e^{-\lambda}$.

Numbers of Hurricanes

```
data( "Rsum" )

fpois( Rsum$Ct )

 Test for Equality of (Poisson) Mean and Variance

data:  Rsum$Ct
Chi-square(n - 1) = 67.488, mean = 1.8169, variance = 1.7517,
degrees of freedom = 70.0000, p-value = 0.5629
alternative hypothesis: greater
```

Numbers of Hurricanes

Incorporate ENSO state as a covariate:

fit <- glm(Ct ~ EN, data = Rsum, family = poisson())
fit

Call: glm(formula = Ct ~ EN, family = poisson(), data = Rsum)

Coefficients:

(Intercept)	EN
0.5751	-0.2483

Degrees of Freedom: 70 Total (i.e. Null); 69 Residual
Null Deviance: 72.23
Residual Deviance: 67.49 AIC: 229.1

Model is found to be

$$\log \lambda \approx 0.58 - 0.25 \cdot \mathrm{ENSO}$$

Use summary(fit) to test for significance of the inclusion of ENSO.

Generalized Pareto Distribution (GPD)

Besides counting exceedances over a threshold, we might want to also infer about the magnitudes of the excesses, $X - u$, conditioned on $X > u$ and u a high threshold.

Analogous to the block maxima case, there is a limiting distribution family that encompasses three types that is appropriate for modeling excesses over a high threshold, the generalized Pareto (GP) family.

Three types:

- Beta distribution when $\xi < 0$ (upper bound),
- Exponential distribution $\xi = 0$ (light upper tail),
- Pareto distribution when $\xi > 0$ (heavy upper tail).

Excesses over a high threshold

Generalized Pareto Distribution (GPD)

Looks like the exponent of the GEV distribution

The GPD is given by:

$$\mathbb{P}[X > x | X > u] = H(x) = 1 - \left[1 + \xi\left(\frac{x-u}{\sigma_u}\right)\right]_+^{-1/\xi}$$

Threshold "replaces" the location parameter

If $\max\{X_1, \dots, X_n\} \sim \mathrm{GEV}(\mu, \sigma, \xi)$ then

$$Y = X - u | X > u \sim \mathrm{GP}(\sigma_u, \xi)$$

where $\sigma_u = \sigma + \xi(u - \mu)$.

Similarly, if we condition Y on a higher threshold, say v, then $Y \sim \mathrm{GP}(\sigma_v, \xi)$, where $\sigma_v = \sigma_u + \xi(v - u)$. Called POT-stability

$$H(x) = 1 - e^{\frac{-(x-u)}{\sigma}}, \text{ as } \xi \to 0$$

Scale parameter depends on the threshold

Inner part of "exponent" term must be positive. Otherwise, it is set to zero (same as for GEV distribution), so it, and its associated likelihood, also has support that depends on the parameter values!

Generalized Pareto Distribution (GPD)

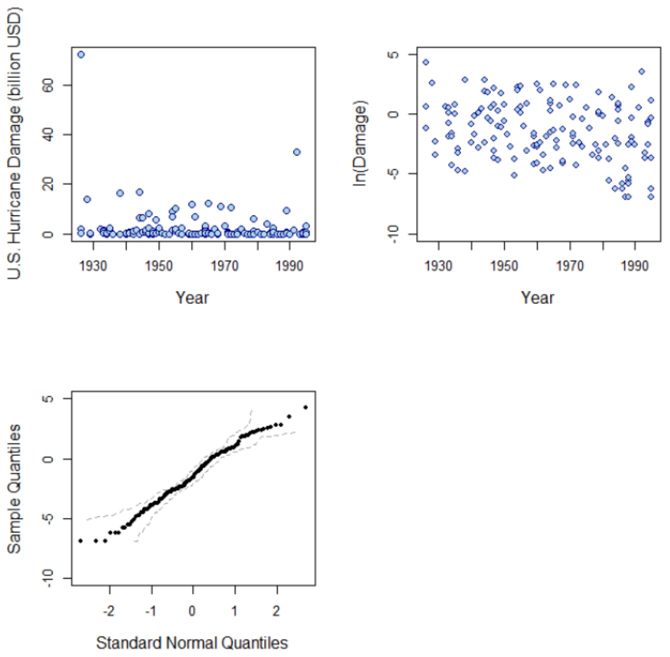
Quantiles are easy to find, as with the GEV distribution, but return levels require estimation of the probability of exceeding the threshold. They are given by:

$$x_m = \begin{cases} u + \frac{\sigma_u}{\xi}\left[(m\zeta_u)^{\xi} - 1\right], \xi \neq 0 \\ u + \sigma \log(m\zeta_u)\, , \xi = 0 \end{cases}$$

where $\zeta_u = \mathbb{P}[X > u]$ and x_m is the value that is exceeded, on average, once every m observations.

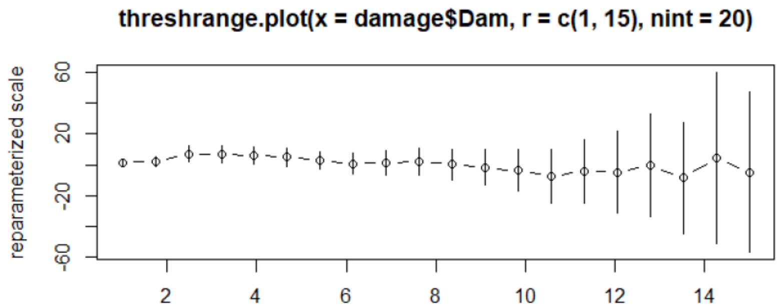
Generalized Pareto Distribution (GPD)

```
data("damage" )

par(mfrow = c(2, 2))

plot(damage$Year, damage$Dam, xlab = "Year",
   ylab = "U.S. Hurricane Damage (billion USD)",
   cex = 1.25, cex.lab = 1.25, col = "darkblue",
   bg = "lightblue", pch = 21)

plot(damage[, "Year"], log(damage[, "Dam"]), xlab = "Year",
   ylab = "ln(Damage)", ylim = c(-10, 5), cex.lab = 1.25,
   col = "darkblue", bg = "lightblue", pch = 21)

qqnorm(log(damage[, "Dam"]), ylim = c(-10, 5), cex.lab = 1.25)
```

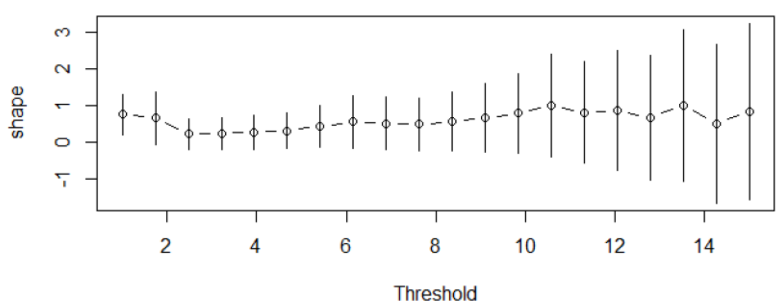


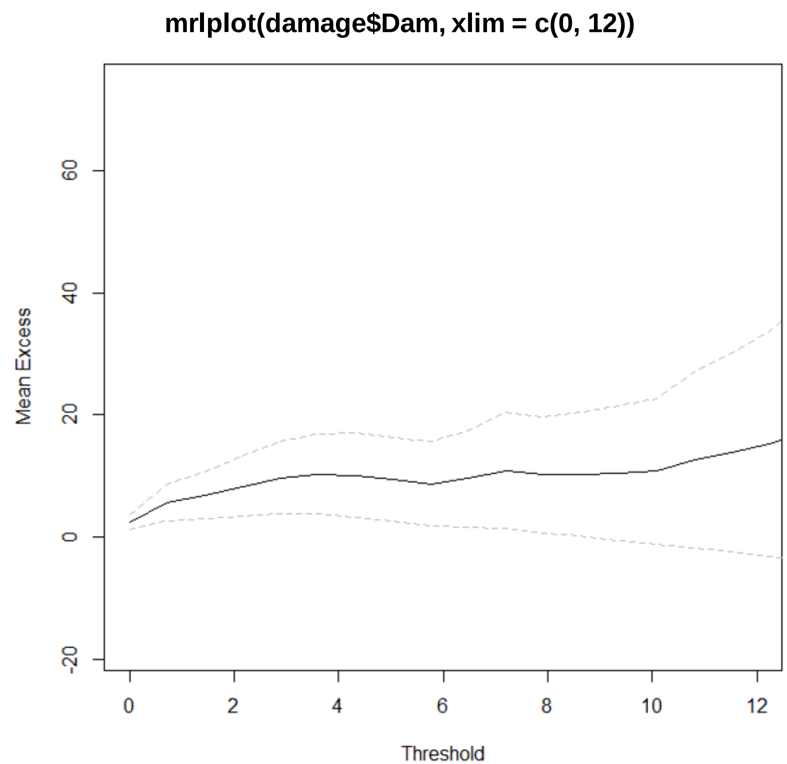
Generalized Pareto Distribution (GPD)

Before fitting the GPD to data, must choose a threshold, which is a bias-variance trade-off



Scale is transformed so that it is not a function of the threshold by $\sigma^* = \sigma_u - \xi u$

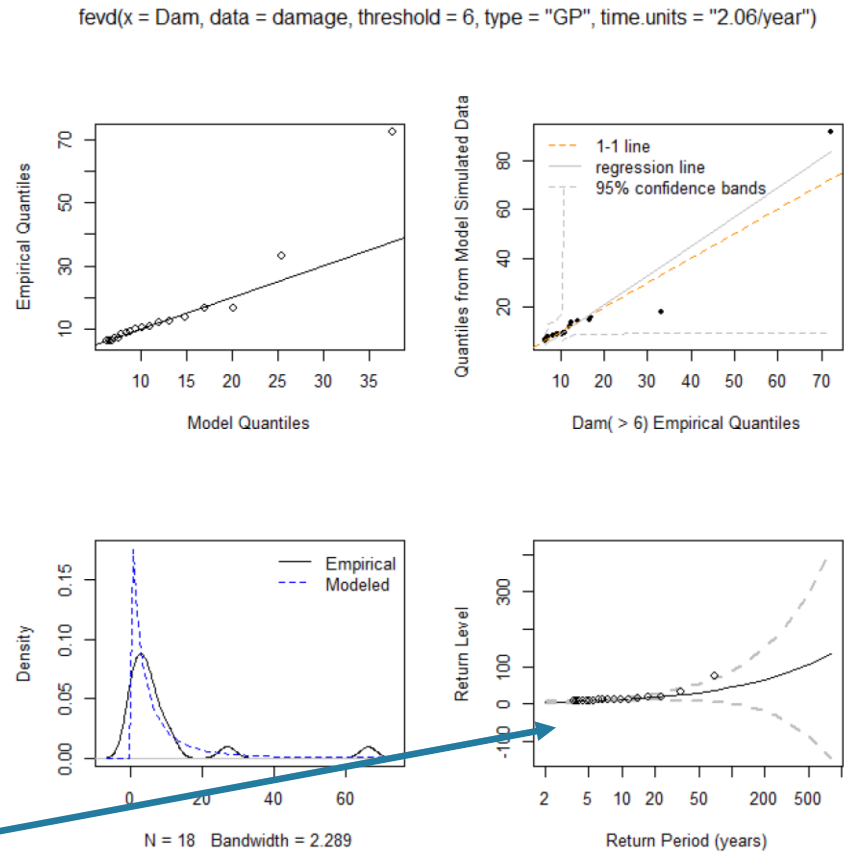




Excesses over a high threshold

Generalized Pareto Distribution (GPD)

```
range(damage$Year)
1995 - 1926 + 1

dim(damage)
144 / 70

fitD <- fevd(Dam, damage,
    threshold = 6,
    type = "GP",
    time.units = "2.06/year" )

fitD
plot(fitD)
```

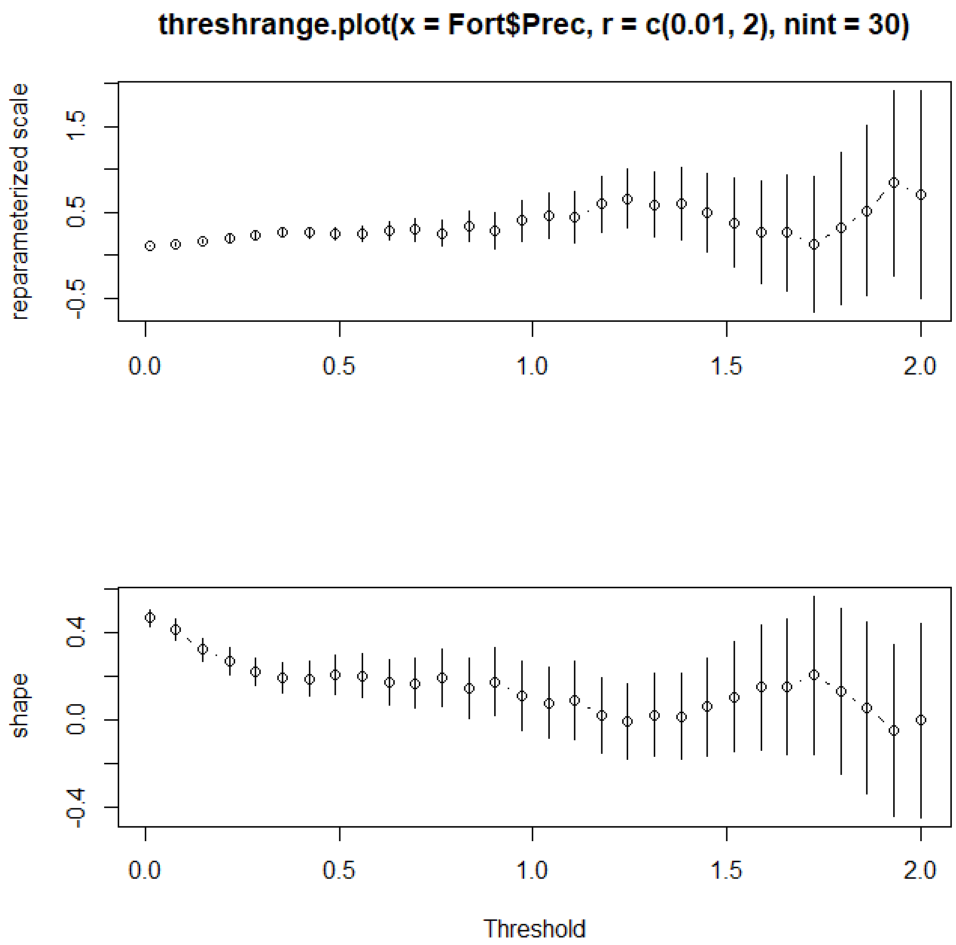


Estimate of the number of points per year (default is the usual 365.25 per year). Only important for getting the empirical return levels on the return-level plot.

Excesses over a high threshold

Generalized Pareto Distribution (GPD)

Fort Collins, Colorado
precipitation data (in/100)

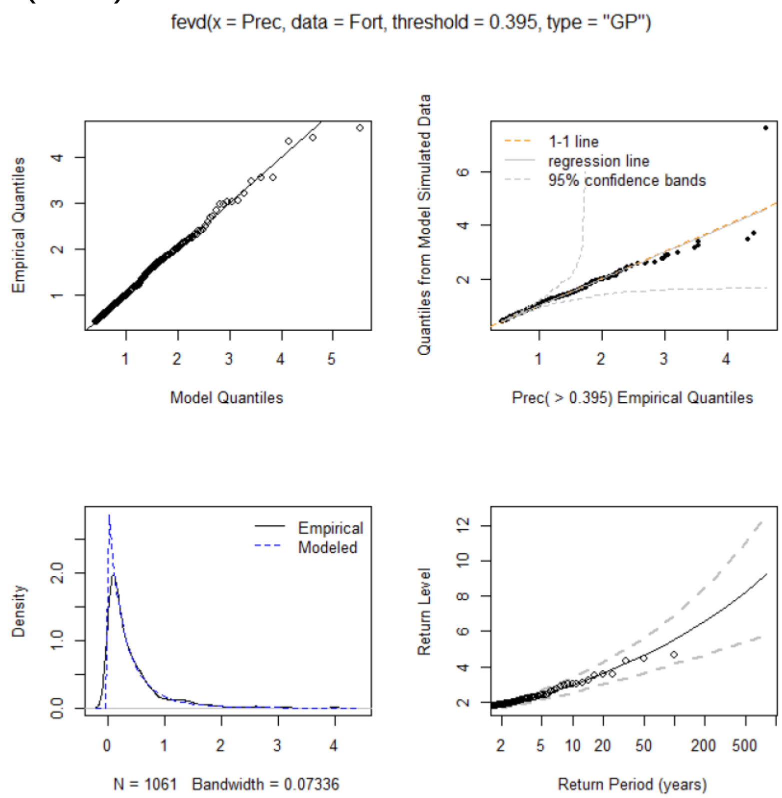


Generalized Pareto Distribution (GPD)

Fort Collins, Colorado precipitation data (in/100)

```
fitFC <- fevd(Prec, Fort, threshold = 0.395, type = "GP")
fitFC
plot(fitFC)
```

```
[1] "Estimation Method used: MLE"

 Negative Log-Likelihood Value:  85.07827

 Estimated parameters:
    scale     shape
0.3224764 0.2119121

 Standard Error Estimates:
     scale      shape
0.01571629 0.03840740

 Estimated parameter covariance matrix.
              scale         shape
scale  0.0002470018 -0.0003976316
shape -0.0003976316  0.0014751280

 AIC = 174.1565

 BIC = 184.0905
```



Generalized Pareto Distribution (GPD)

Fort Collins, Colorado precipitation data (in/100)

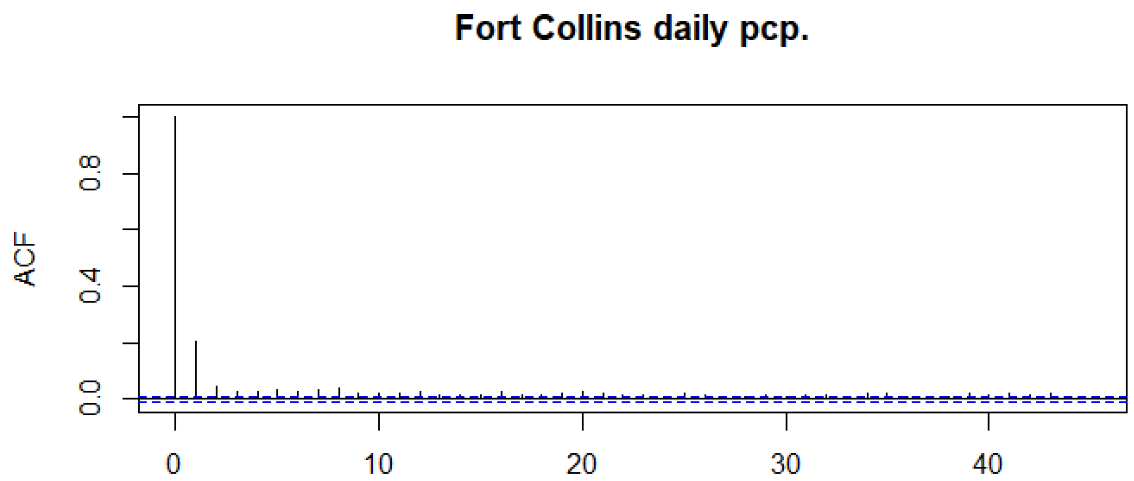
```
fitFC <- fevd(Prec, Fort, threshold = 0.395, type = "GP")
fitFC
plot(fitFC)

fitFC2 <- fevd(Prec, Fort, threshold = 0.395, scale.fun =
   ~ cos(2 * pi * tobs / 365.25) + sin(2 * pi * tobs / 365.25),
   type = "GP", use.phi = TRUE, units = "inches")

plot(fitFC2, type = "qq" )

lr.test(fitFC, fitFC2)
```

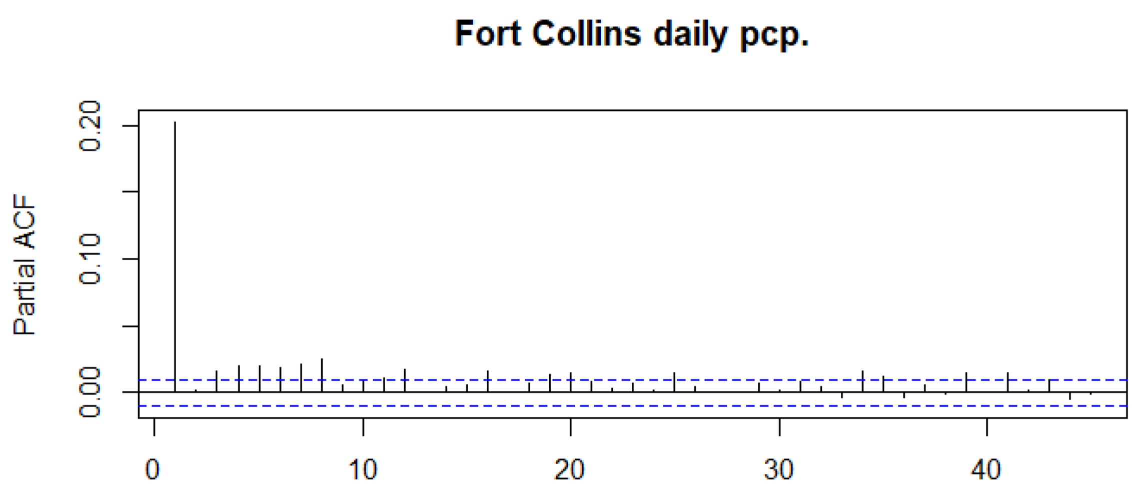
```
Likelihood-ratio Test

data:  Prec
Likelihood-ratio = 24.327, chi-square critical value = 5.9915,
alpha = 0.0500, Degrees of Freedom = 2.0000, p-value = 5.219e-06
alternative hypothesis: greater
```

Excesses over a high threshold

Dependence in the data (non-extremes)

extRemes: An R software package

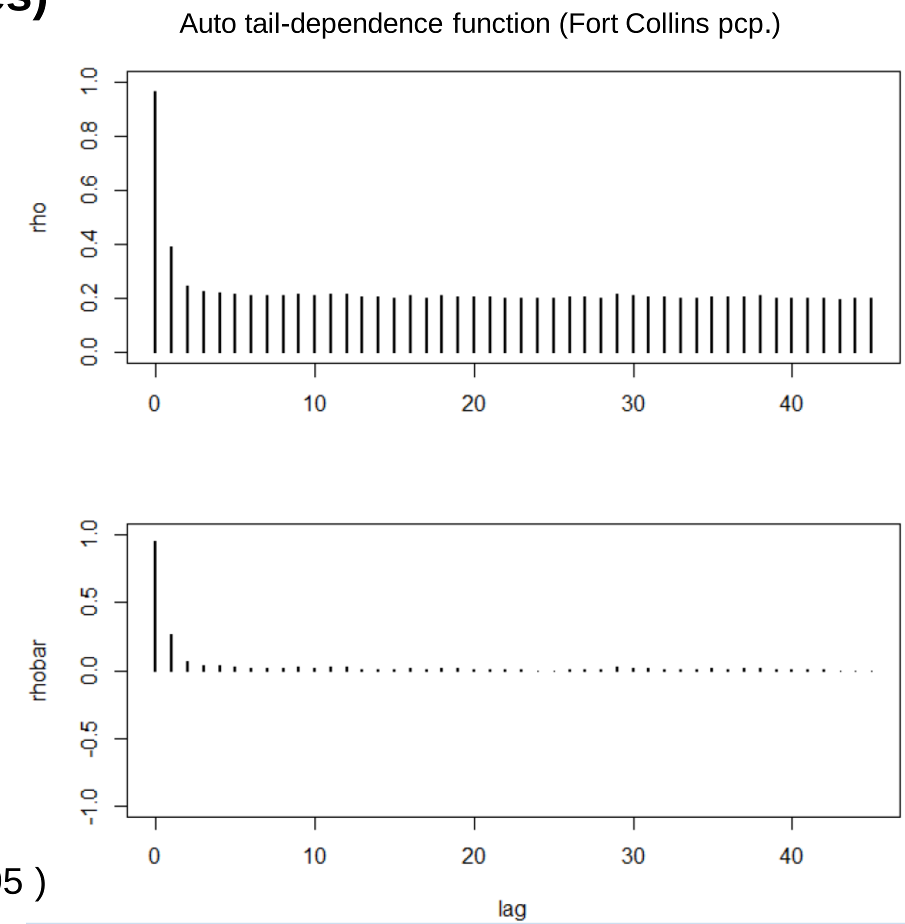




Excesses over a high threshold

Dependence in the data (extremes)

```
atdf( Fort$Prec, u = 0.8 )
```



extRemes: An R software package

extremalindex(Fort$Prec, threshold = 0.395)

Intervals Method Estimator for the Extremal Index
NULL

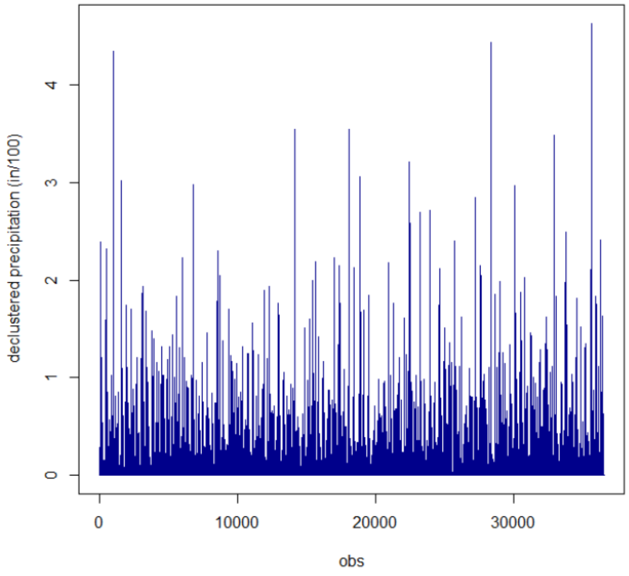
theta.tilde used because there exist inter-exceedance times > 2.

extremal.index	number.of.clusters	run.length
0.6246345	651.0000000	9.0000000

Generalized Pareto Distribution (GPD)

Fort Collins, Colorado precipitation data (in/100)

```
dcFC <- decluster( Fort, which.cols = 6, threshold = 0.395, r = 9 )
Fort <- cbind( Fort, "dcPrec" = c( dcFC ) )
plot( dcPrec ~ obs, data = Fort, ylab = "declustered precipitation (in/100)",
    col = "darkblue", type = "h" )
abline( h = 0.395, lty = 2, lwd = 2 )

extremalindex( Fort$dcPrec, threshold = 0.395 )
```



Now you can do the fitting all over with the declustered data, and obtain more accurate uncertainty information.

Poisson point-process characterization

- If we combine the counting of exceedances with the intensity information (i.e., the excesses), then we have a point-process.
- We can do so orthogonally by fitting the frequency of exceeding the high threshold and the GPD to the data separately, called the orthogonal approach.
- Better, we can fit a two-dimensional Poisson point-process to the data so that we capture the uncertainty in estimating each part all at once.
- Waiting times between excesses are exponentially distributed with unit mean.
- Exponential distribution has a memoryless property (related to the POT-stability property).

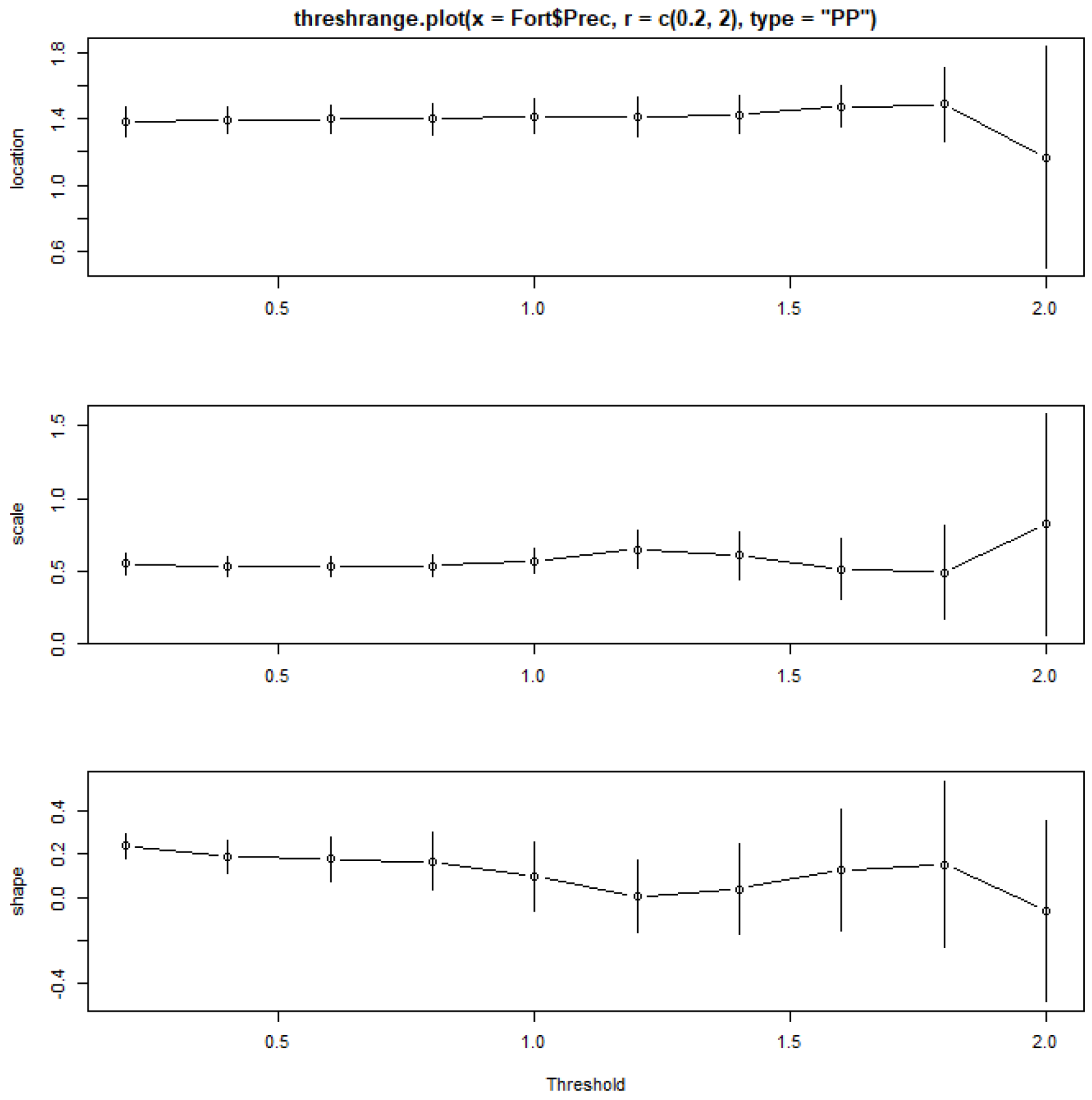
Poisson point-process characterization

Relation of GEV(μ, σ, ξ) to those of a Poisson point process with parameters (λ, σ^*, ξ)

1. ξ is the same
2. $\log \lambda = -\frac{1}{\xi} \log\left(1 + \xi \frac{u-\mu}{\sigma}\right)$
3. $\sigma^* = \sigma + \xi(u - \mu)$

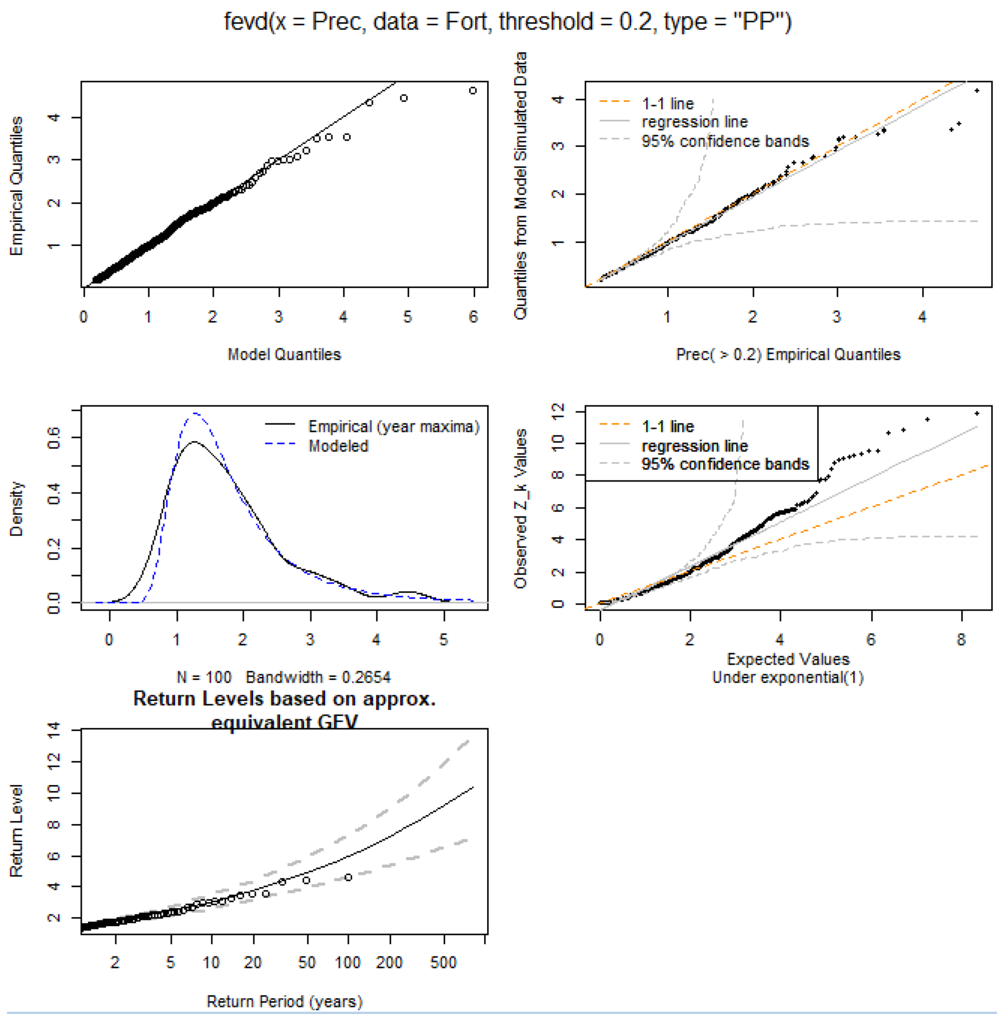
Poisson point process characterization of a GEV

extRemes: An R software package



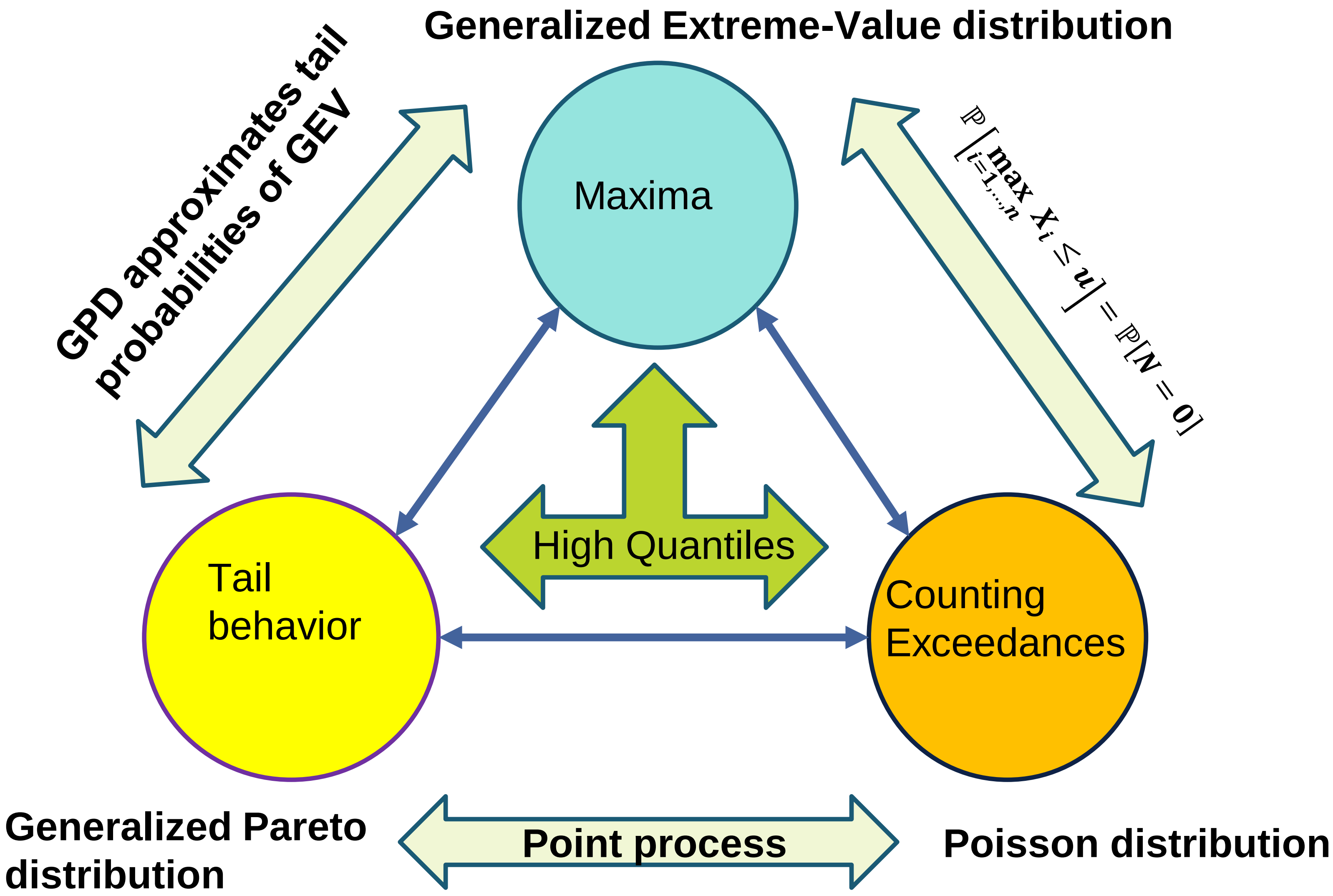
Poisson point process characterization of a GEV

```
fit0 <- fevd( Prec,
    data = Fort,
    threshold = 0.395,
    type = "PP" )
fit0
plot( fit0 )
```



extRemes: An R software package

Extreme-Value Analysis (EVA)



Extreme-Value Analysis (EVA)

The EVD's do not work for all extrema. Some maxima over long blocks, or excesses over a high threshold, do not converge to a non-degenerate distribution at all.

A super-heavy tail distribution, such as the log-Pareto given by $F(x) = 1 - \log^{-1/\alpha}(x)$, is one such type of distribution that has no non-degenerate distribution for its extremes.

Estimation

- **Maximum Likelihood**
  - Must use numerical optimization
  - Regularity assumptions required for the MLE to follow a normal distribution are not met when $\xi \leq -1/2$ (Smith 1985; Büecher and Segers 2017)
  - Many times, the likelihood curve is rather steep and/or has undefined points
  - Easy to incorporate covariates into parameter estimates
- **L-moments**
  - Quick and easy to compute
  - Must use bootstrap methods for uncertainty
  - Not as easy to incorporate covariates into parameter estimates
  - Often used when sample size is small
- **Bayesian**
  - Easy to incorporate uncertainty into parameter estimates
  - Difficult to get good mixing in MCMC
- **Generalized MLE (GMLE, aka penalized MLE; Martins and Stedinger 2000; 2001)**
  - Often useful to avoid problematic areas
- **Various**
  - Hill estimator (not always useful; Resnick 2007, p. 86)
  - Non-parametric (e.g., Huang et al. 2018)
  - weighted composite log-likelihood (Stein 2023)
  - Neural Networks (e.g., Rai et al. 2023)

Uncertainty Estimation

- **Normal Approximation CI's**
  - When using MLE or GMLE
  - Normality assumption may not be valid (e.g., if $\xi \leq -1/2$ or for long return levels)
  - Delta method an be used to obtain CI's for return levels
  - Quick and easy to compute
- **Profile likelihood**
  - Generally the most accurate choice
  - Can be difficult to obtain
  - Difficult to automate
- **Bayesian**
- **Bootstrap**
  - Issues abound for bootstrapping extremes (cf. Bickel and Freedman 1981; see G. 2020 for a recent review)
  - Will never sample a maximum higher than what is observed in the data, and will often obtain samples without the maximum from the data
  - For heavy-tailed distributions, should use an $m < n$ bootstrap
  - Parametric bootstrap is good but can yield intervals that are too narrow (cf. Kyselý 2002; Schendel and Thongwichian 2015; 2017)
  - Test-inversion bootstrap (TIB) generally the best choice (similar to profile likelihood; Schendel and Thongwichian 2015; 2017)
    - Requires use of a root-finding algorithm when covariates are included
    - Often difficult to obtain a solution in general (cf. G. 2020).

References

Bickel, P. J. and D. A. Freedman (1981) Some asymptotic theory for the bootstrap. *The Annals of Statistics*, **9** (6), 1196 – 1217, doi: 10.1214/aos/1176345637.

Büecher, A., and J. Segers, 2017: On the maximum likelihood estimator for the generalized extreme-value distribution. *Extremes*, **20**, 839–872, https://doi.org/10.1007/s10687-017-0292-6.

Gilleland, E., 2020. Bootstrap methods for statistical inference. Part II: Extreme-value analysis. *Journal of Atmospheric and Oceanic Technology*, **37** (11), 2135 - 2144, doi: 10.1175/JTECH-D-20-0070.1.

Huang, W. K., Nychka, D. W., and Zhang, H. (2019). Estimating precipitation extremes using the log-histospline. *Environmetrics*, 30(4):e2543

Kyselý, J., 2002: Comparison of extremes in GCM-simulated, downscaled and observed central-European temperature series. *Climate Res.*, **20**, 211–222, https://doi.org/10.3354/cr020211.

Martins ES, Stedinger JR (2000). "Generalized Maximum-Likelihood Generalized Extreme-Value Quantile Estimators for Hydrologic Data." *Water Resources Research*, **36**(3), 737–744. doi:10.1029/1999wr900330.

Martins ES, Stedinger JR (2001). "Generalized Maximum-Likelihood Pareto-Poisson Estimators for Partial Duration Series." *Water Resources Research*, **37**(10), 2551–2557. doi:10.1029/2001wr000367.

References

Rai, S., A. Hoffman, S. Lahiri, D. W. Nychka, S. R. Sain, S. Banyopadhyay, 2023. Fast parameter estimation of Generalized Extreme Value distribution using Neural Networks. doi: 10.48550/arXiv.2305.04341.

Resnick, S. I., 2007: *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling.* Springer Series in Operations Research and Financial Engineering, Springer, 404 pp.

Schendel, T., and R. Thongwichian, 2015: Flood frequency analysis: Confidence interval estimation by test inversion bootstrapping. *Adv. Water Resour.*, **83**, 1–9, https://doi.org/10.1016/j.advwatres.2015.05.004.

Schendel, T., and R. Thongwichian, 2017: Confidence intervals for return levels for the peaks-over-threshold approach. *Adv. Water Resour.*, **99**, 53–59, https://doi.org/10.1016/j.advwatres.2016.11.011.

Smith, R. L., 1985: Maximum likelihood estimation in a class of nonregular cases. *Biometrika*, **72**, 67–90, https://doi.org/10.1093/biomet/72.1.67.

Stein, Michael L., 2023. A weighted composite log-likelihood approach to parametric estimation of the extreme quantiles of a distribution, Extremes, 10.1007/s10687-023-00466-w.