The Extremes Toolkit (extRemes)

Weather and Climate Applications of Extreme Value Statistics

Eric Gilleland

National Center for Atmospheric Research (NCAR), Research Applications Laboratory (RAL)

Richard W. Katz[1, 2]

NCAR, Institute for the Study of Society and Environment (ISSE)

The Extremes Toolkit (extRemes)

Web Page

http://www.assessment.ucar.edu/toolkit

Acknowledgements

1. The Extremes Toolkit is funded by the National Science Foundation (NSF) through the NCAR Weather and Climate Impact Assessment Science Initiative (now Program) with additional support from the NCAR Geophysical Statistics Project (GSP). Initial work on the toolkit was performed by **Greg Young**.
2. We thank **Stuart Coles** for his permission to use his S functions and **Alec Stephenson** for supplying his R port of these functions to the R-CRAN website.

Outline

- Motivation and Goals for The Extremes Toolkit
- The R programming language
- Overview of Functionality of `extRemes`
- Example
- On-going work

Motivation and Goals for The Extremes Toolkit

Motivation

The toolkit was motivated by the continued use, particularly by non-statisticians, of traditional statistical distributions (e.g., normal, lognormal, gamma, etc.) in situations where extreme value theory is applicable.

Goals

- To provide a GUI prototype to interact with a high-level language capable of advanced statistical applications.
- Available to a wide audience (not just statisticians).
  - Weather and Climate Impacts (and related) scientists.
  - Free!
  - Consistent across major platforms (e.g., Windows, Linux, Mac OS).
  - Small Learning Curve.

The R Project for Statistical Computing

`extRemes` is *written in* and requires R, but does not require familiarity with R.

About R

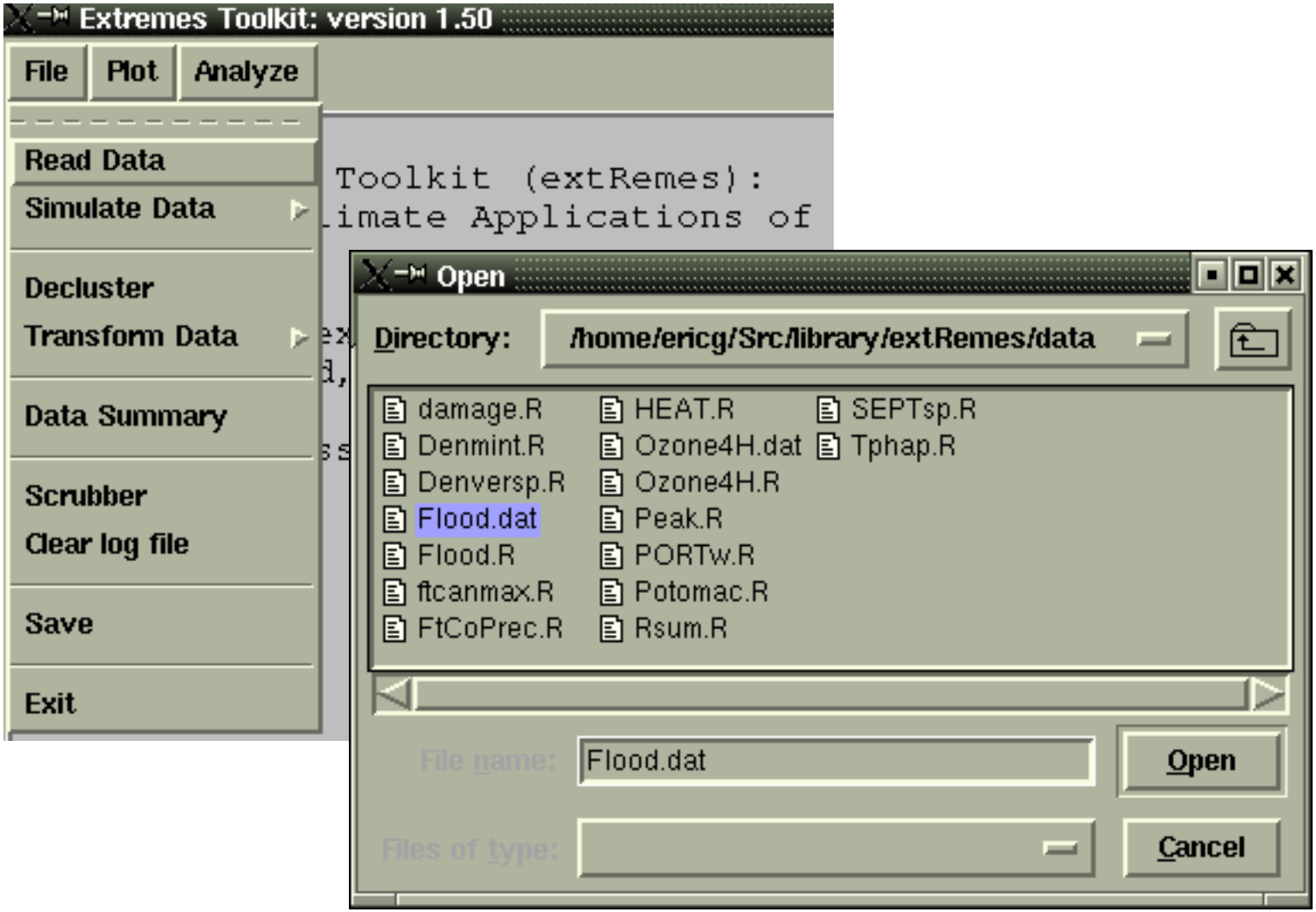
- R is a language and environment for statistical computing and graphics.
- It's Free!
- Curiously similar to S-Plus.
- Mostly identical across platforms.
- C, C++ and Fortran code can be linked and called at run time.
- Many researchers contribute packages to R (over 500 packages on CRAN).

`http://www.r-project.org`

Functionality of `extRemes`

- Data management
  - Opening data files
  - Simulating Data (GEV, GPD)
  - Data Transformations
- Log file (`extRemes.log`) showing executed code
- Graphical Displays
- Fitting data
- Mostly, `extRemes` is a GUI interface to `ismev`, but has some additional functionality. For example,
  - Runs Declustering
  - Extremal Index
- Other Diagnostics

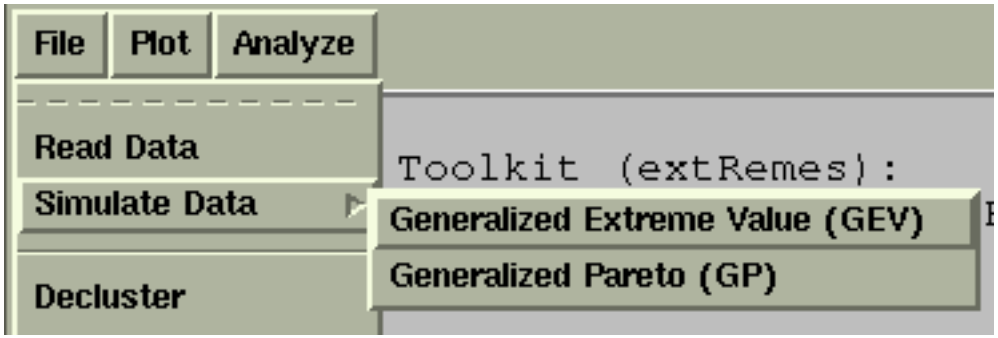
Data management: Opening data files

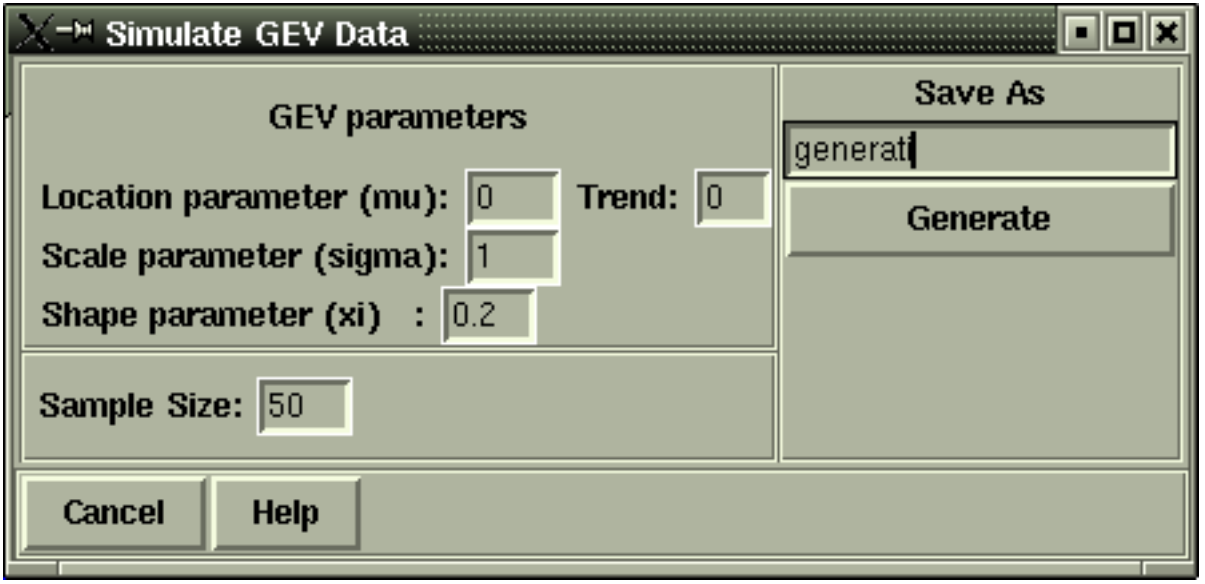


Data management: Opening data files

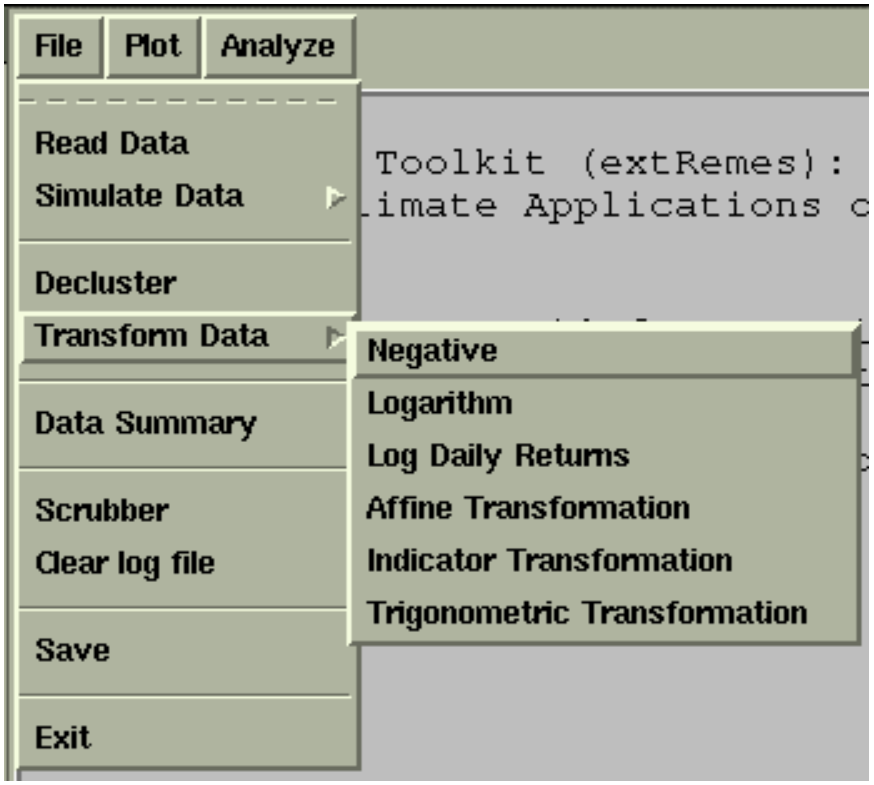

```
> [1] "Successfully opened file: Flood.dat"
                    OBS      HYEAR     USDMG    DMGPC    LOSSPW
N              66.00000   66.00000 66.000000 66.00000   66.0000
mean           33.50000 1964.50000  2.629076 12.92844  270.5659
Std.Dev.       19.19635   19.19635  3.168426 13.88106  293.8702
min             1.00000 1932.00000  0.116800  0.92420   14.5300
Q1             17.25000 1948.25000  0.690225  3.78565   92.1600
median         33.50000 1964.50000  1.395600  7.53605  163.9700
Q3             49.75000 1980.75000  3.381075 16.55457  333.7825
max            66.00000 1997.00000 17.167800 68.32760 1453.1300
missing values  0.00000    0.00000  0.000000  0.00000    0.0000

 Saving workspace (may take a few moments) ...
```

Data management: Simulating Data





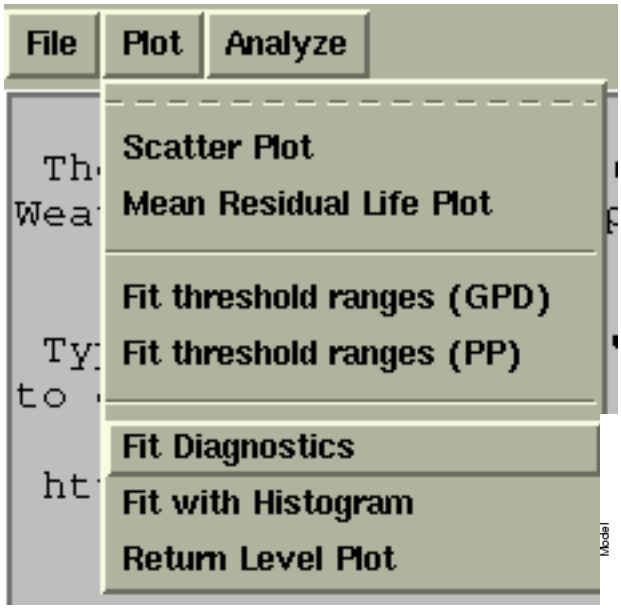
Data management: Data Transformations

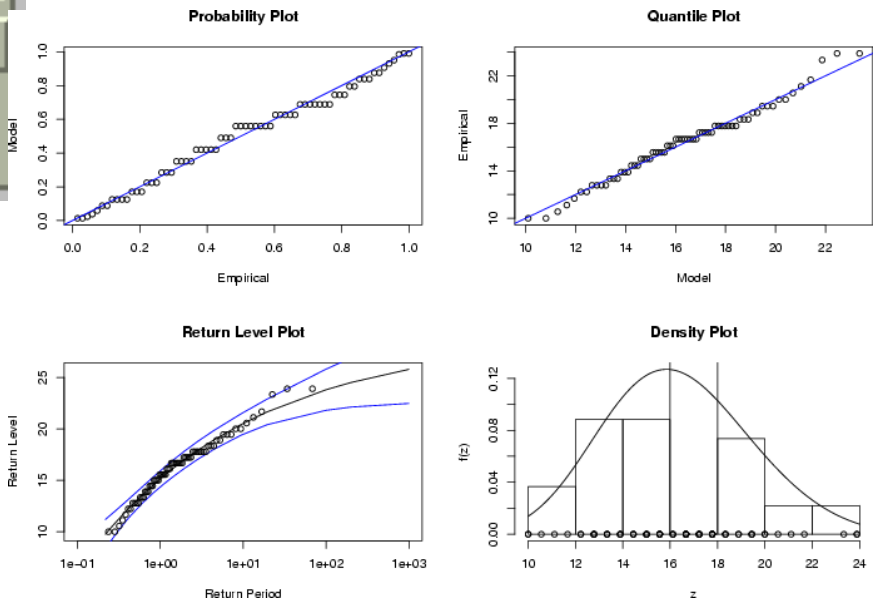


Log file

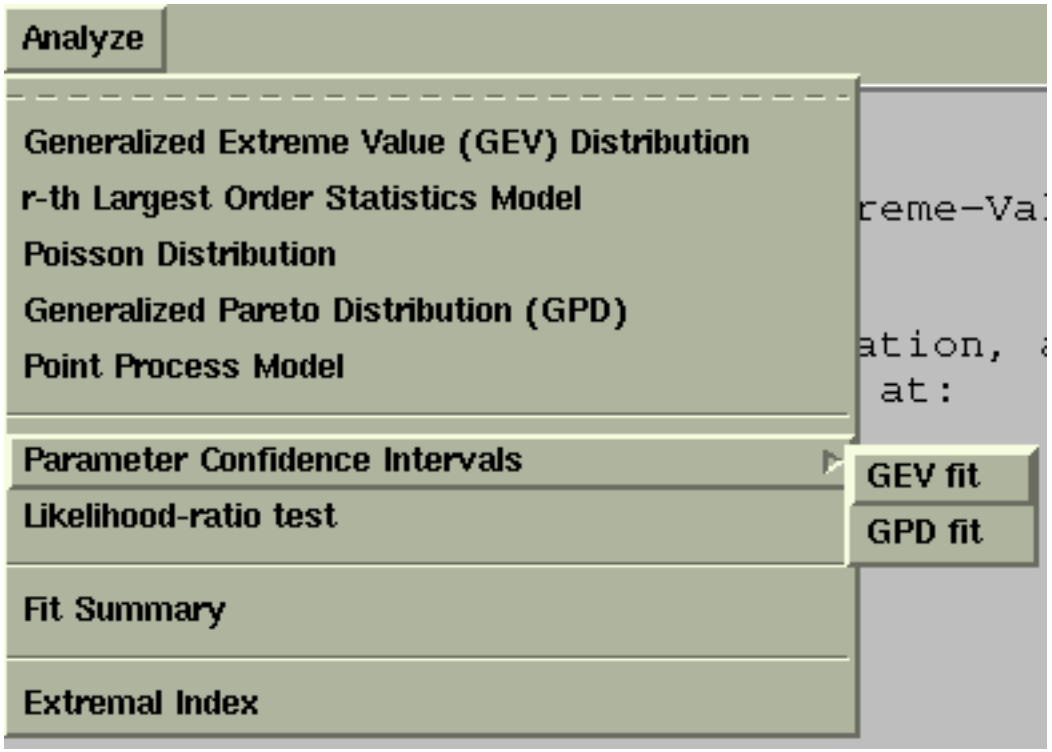
```
p <- c( 0, 1, 0.2)
gev.sim <- gen.gev(p=p, n=50, trend=0)
plot( gev.sim)
gev.sim <- cbind(1:50, gev.sim)
colnames( gev.sim) <- c("obs","gev.sim")
gev.sim <- as.extRemesDataObject( gev.sim)
gev.sim[["name"]] <- "GEV Simulated"
gev.sim[["params"]] <- c(0, 1, 0.2, 0)
gev.sim[["generated"]] <- TRUE
assign( "generati", gev.sim, pos=".GlobalEnv")
save.image()
```

Graphical Displays



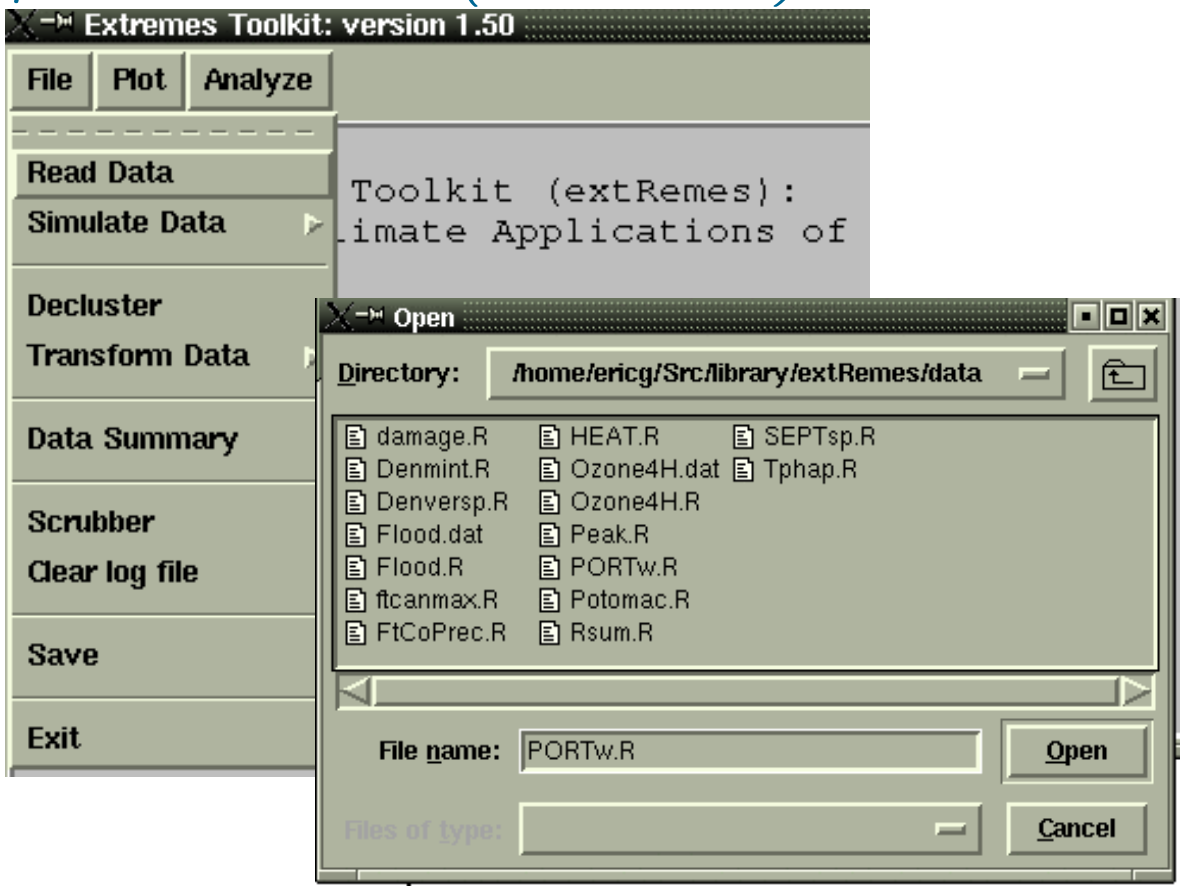


Fitting data



Example: Port Jervis Temperature Data

Open the dataset (PORTw.R)



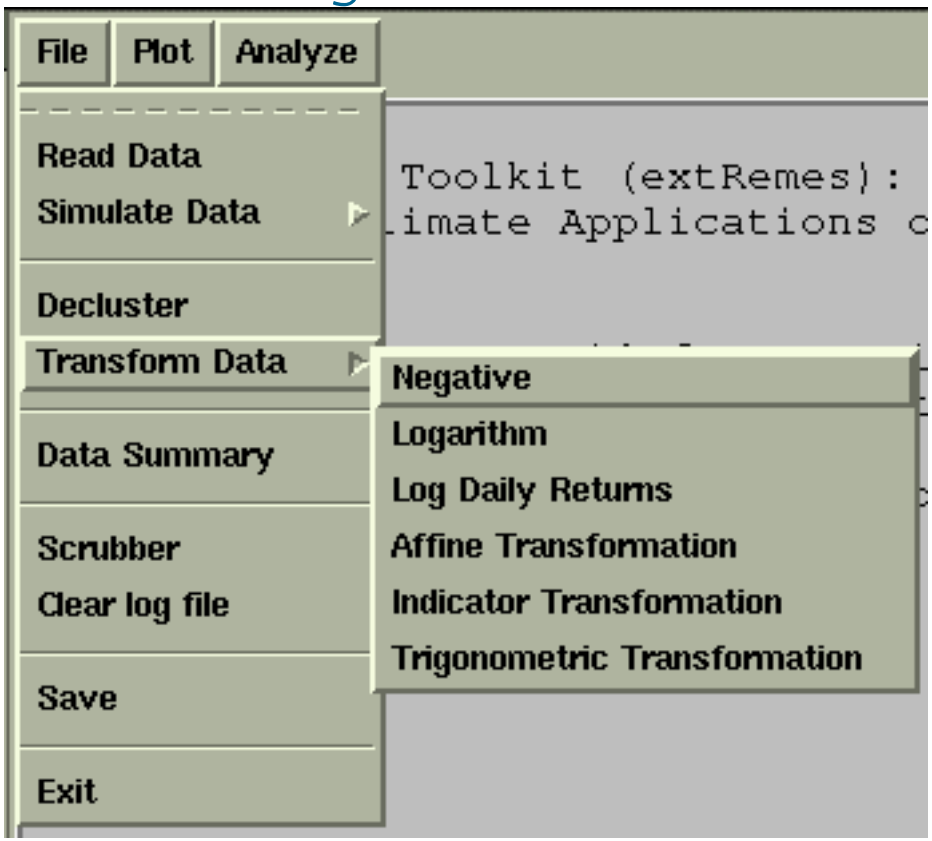
Example: Port Jervis Temperature Data

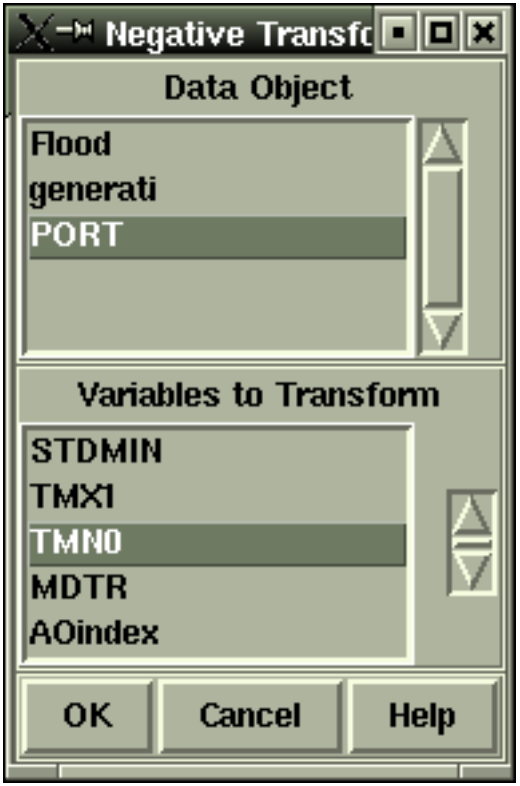

```
> [1] "Successfully opened file: PORTw.R"
                    Year     MTMAX      MTMIN    STDTMAX    STDMIN      TMX1
N                 68.00000 68.000000  68.000000 68.0000000 68.0000000 68.000000
mean            1961.01471  2.872543  -7.151029  5.0702144  5.9775104 16.315363
Std.Dev.          20.21119  1.445223   1.659042  0.8015449  0.7967736  3.104209
min             1927.00000 -0.360200 -10.667900  3.6730700  4.2801700 10.000000
Q1              1943.75000  1.739197  -8.152800  4.6058525  5.4579200 14.305525
median          1961.50000  2.811730  -6.836400  5.0246850  5.8272150 16.666700
Q3              1978.25000  3.816355  -6.202150  5.3763175  6.4117725 17.777800
max             1995.00000  6.208790  -3.211200  7.3393700  8.1993800 23.888900
missing values     0.00000  0.000000   0.000000  0.0000000  0.0000000  0.000000
                     TMN0     MDTR      AOindex
N                68.000000 68.00000 68.00000000
mean            -22.001631 10.02631 -0.04672824
Std.Dev.          3.450322  0.77566  1.07917299
min             -28.888900  8.57050 -2.61315000
Q1              -24.027775  9.47535 -0.69499500
median          -21.944450  9.95680 -0.02762000
Q3              -19.444400 10.41050  0.70904750
max             -13.888900 12.50940  2.90473000
missing values    0.000000  0.00000  0.00000000

 Saving workspace (may take a few moments for large workspaces) ...

 Workspace saved.
```

Daily *block* maxima and minima Winter temperature (^oC) at Port Jervis, New Jersey
with a covariate for the North Atlantic Oscillation index from 1927 through 1995.

Example: Port Jervis Temperature Data

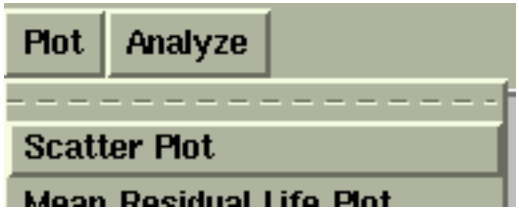
Take the negative of minimum temperature data

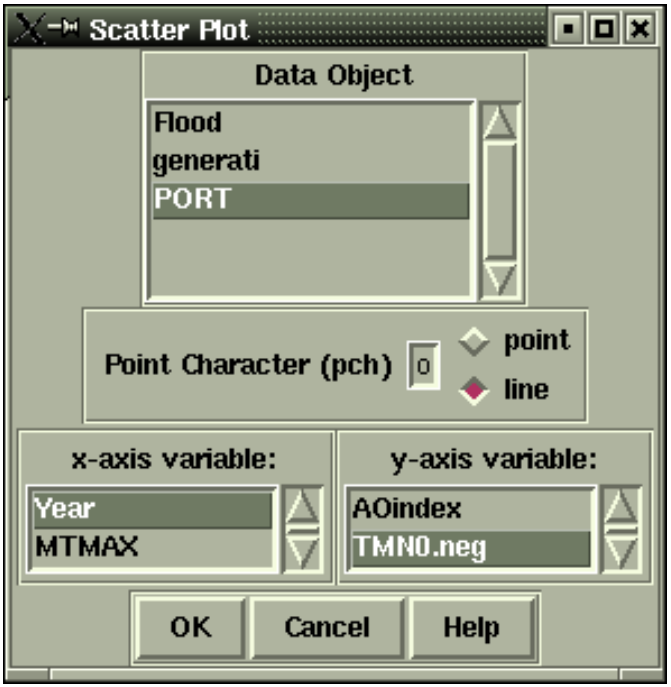


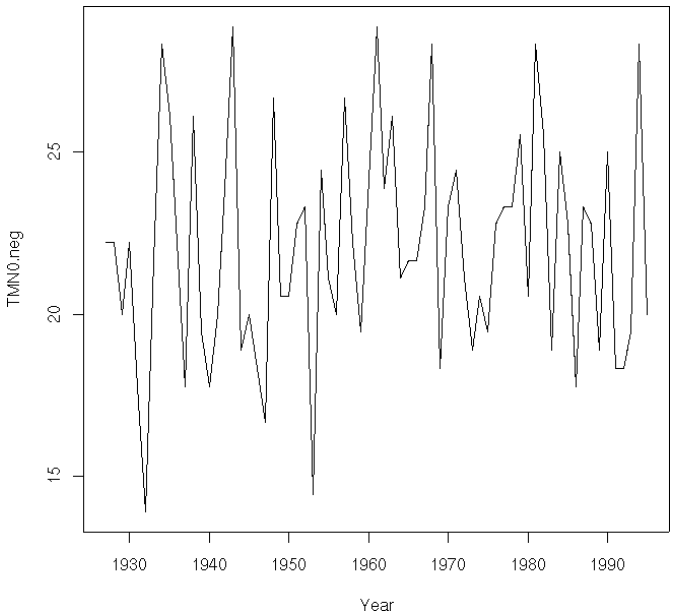


Example: Port Jervis Temperature Data

Plot minimum temperature against time

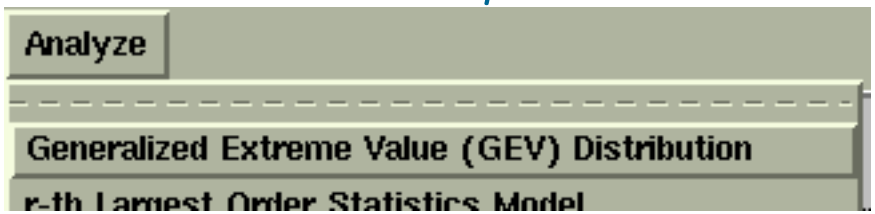


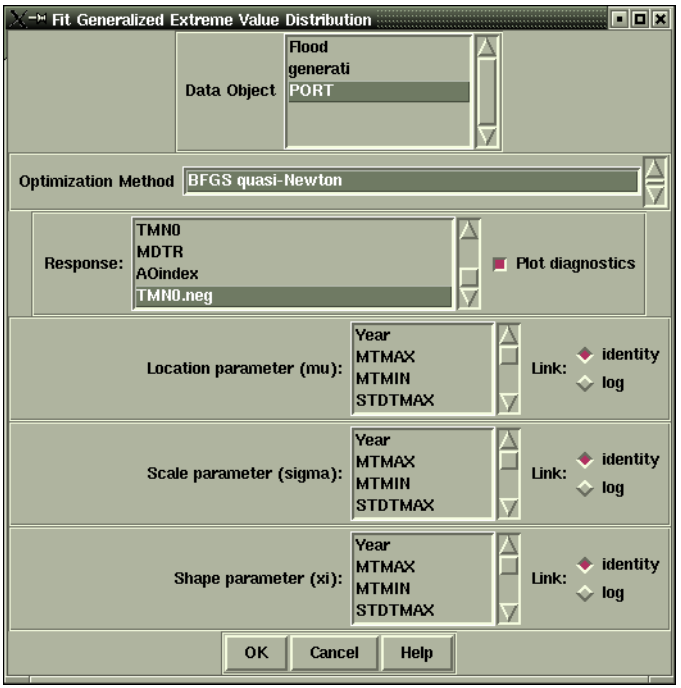


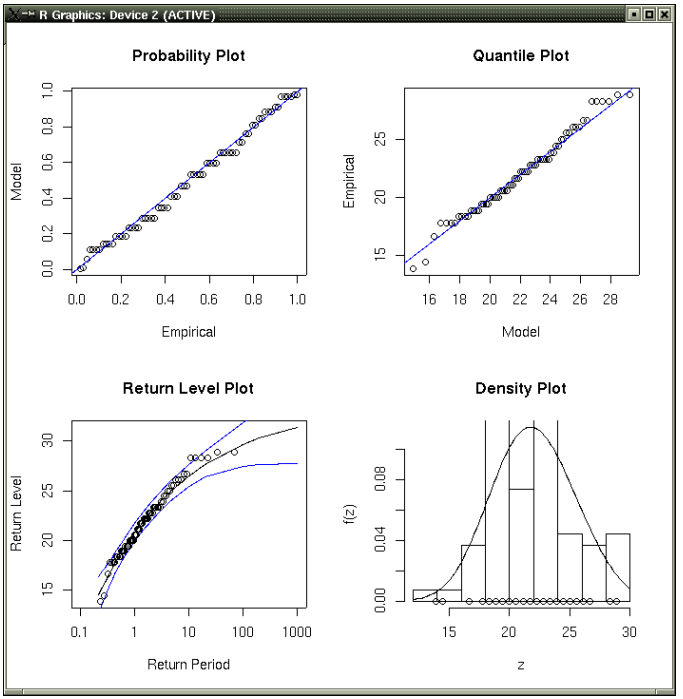


Example: Port Jervis Temperature Data

Fit minimum temperature data to a GEV distribution





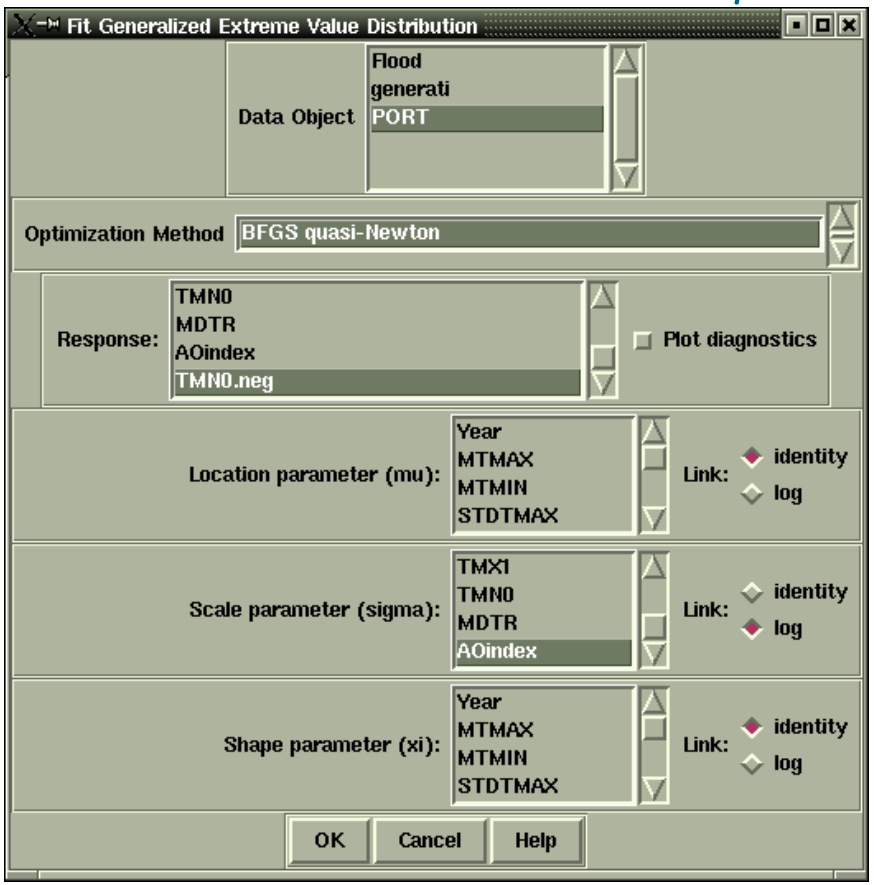


	MLE	Stand. Err.
MU: (identity)	20.76976	0.45498
SIGMA: (identity)	3.34555	0.32424
Xi: (identity)	-0.26401	0.09217

[1] "Negative log-likelihood: 179.55663

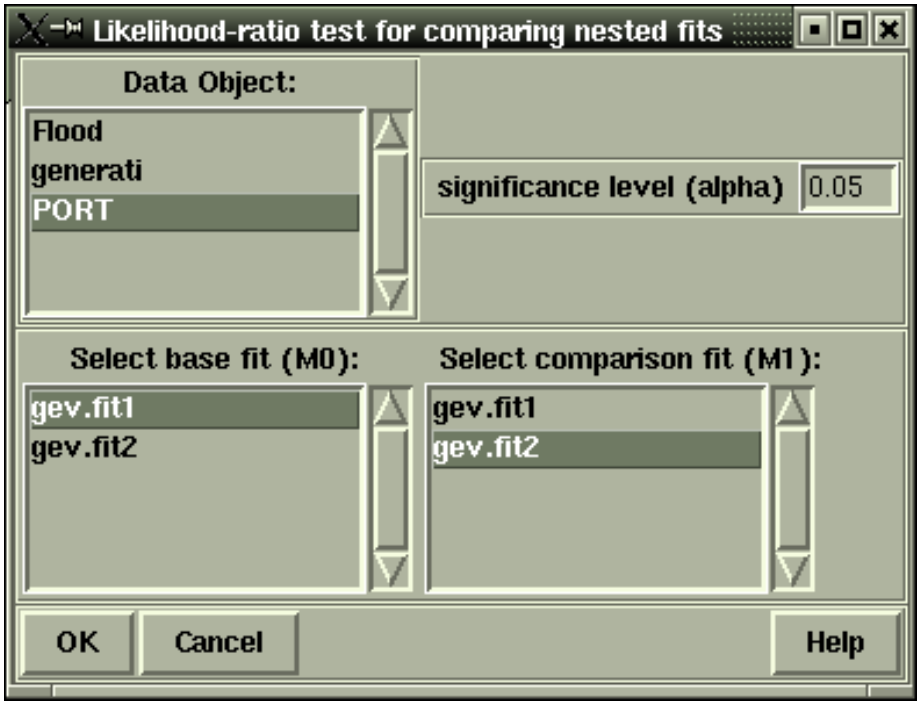
Example: Port Jervis Temperature Data

Fit minimum temperature data to a GEV with AO index as a covariate in the scale parameter



Example: Port Jervis Temperature Data

Likelihood ratio test between the two fits




```
***************
 Likelihood-ratio test statistic for models: M0 = gev.fit1 and M1 = gev.fit2 is:
 0.6072 <= 3.8415 = 1 - 0.05 quantile of a Chi-square with 1 degrees of freedom.
***************
  p-value = 0.435834
***************
```

Tutorial

One of the major components of `extRemes` is the tutorial, which can be found at

http://www.assessment.ucar.edu/toolkit

Table of Contents

Ongoing work

(and work I'd like to have done)

- Spatial `extRemes`
- More functionality (e.g., goodness of fit tests, Hill estimator, L-moments?, ...)
- More user-friendliness (e.g., "stickies", ...)
- Find out who is using it, and how many. Important for future funding.