A User's Guide to the POT Package (Version 1.1)

Mathieu Ribatet

Department of Hydrological Statistic, INRS,
University of Québec, 490 de la Couronne, G1K 9A9 Canada

Cemagref UR HH, 3bis quai Chauveau
69336 Lyon Cedex 09 France

E-mail: ribatet@hotmail.com
17th October 2006

1 Introduction

1.1 Why the POT package?

The

1.6 Legalese

$z \leq 0$, Coles et al. (1999) Fse the Fniform distribFtion on [0,

3.2 Threshold Selection

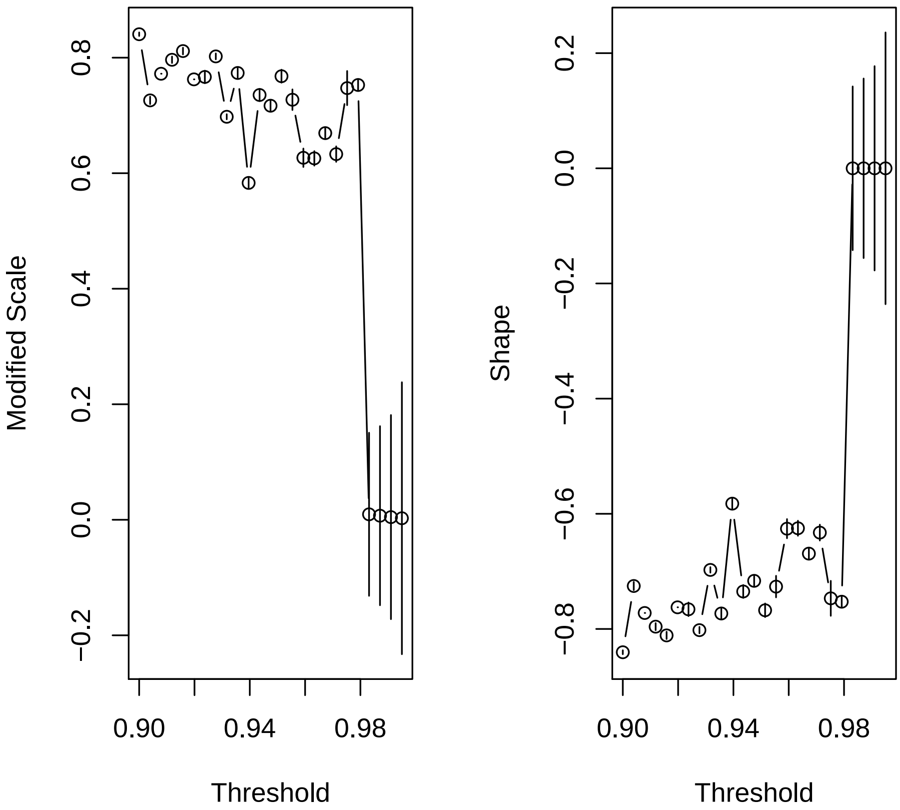

Figure 1: The threshold selection using the tcplot function

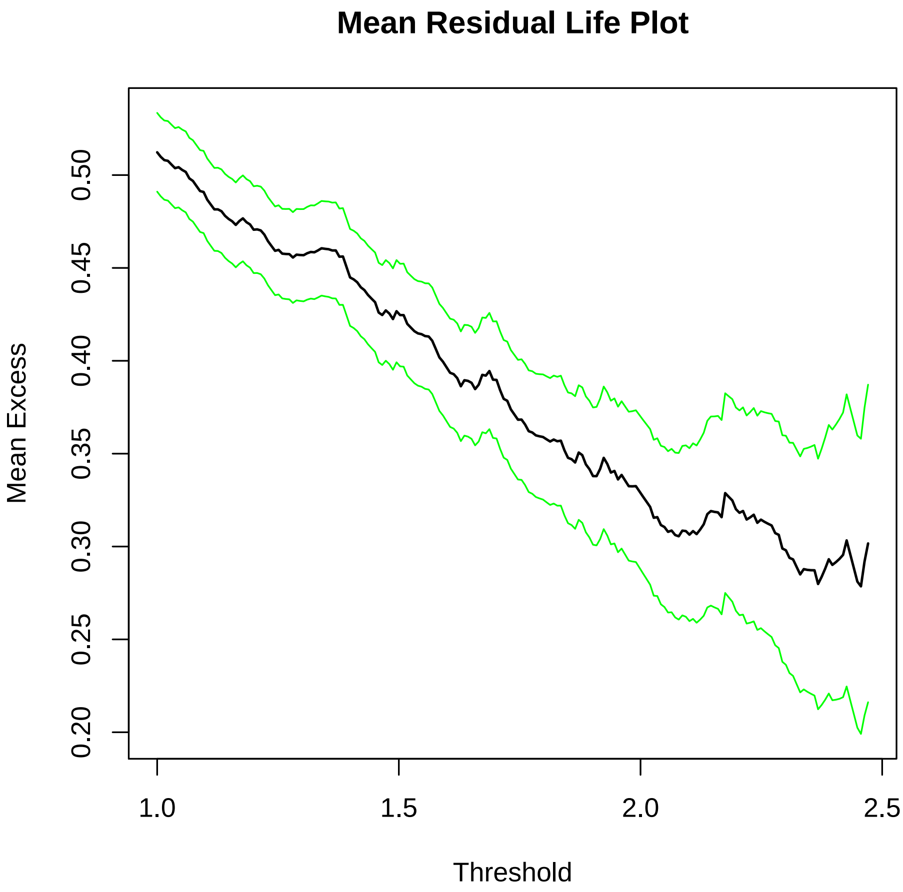

Figure 2: The threshold selection using the mrlplot function

The quantity $E[X - \mu_1 | X > \mu_1]$ is linear in μ_1. Or, $E[X - \mu_1 | X > \mu_1]$ is simply the mean of excesses above the threshold μ_1

3.2.3 L-Moments plot: *lmomplot*

L-moments are summary statistics for probability distributions and data samples. They are analogous to ordinary moments – they provide measures of location, dispersion, skewness, kurtosis, and other aspects of the shape of probability distributions or data samples – but are computed from linear combinations of the ordered data values (hence the prefix L).

For the GPD, the following relation holds:

$$\tau_4 = \tau_3 \frac{1 + 5\tau_3}{5 +}$$

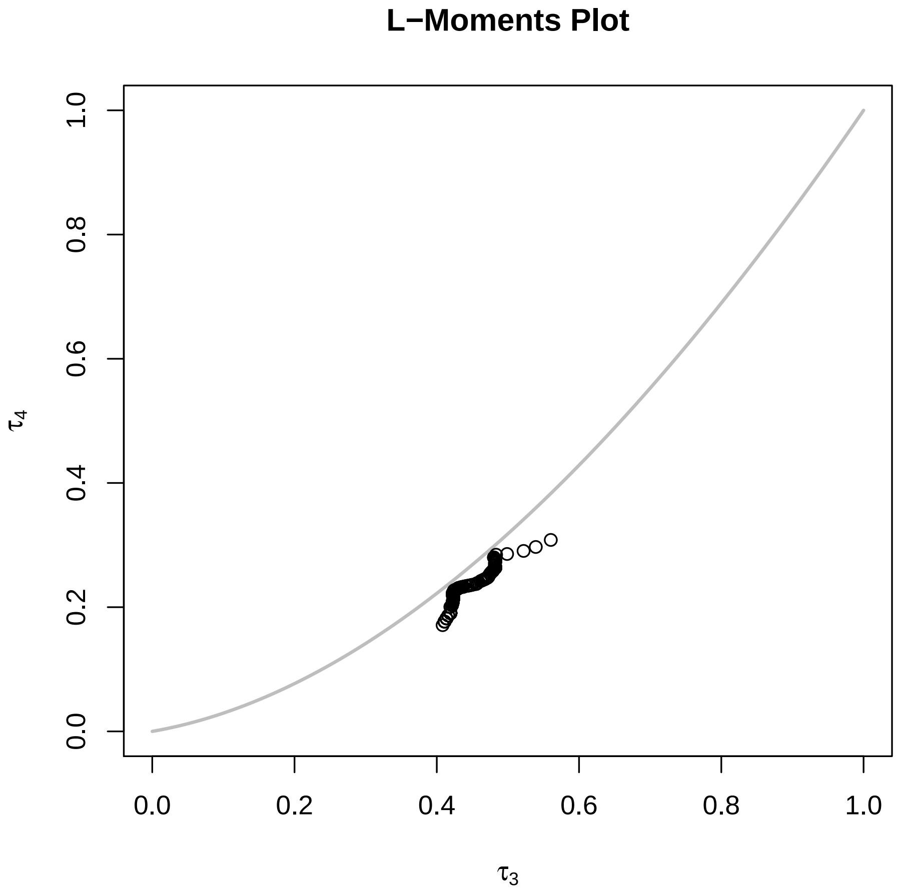

Figure 3: fig: The threshold selection using the lmomplot function

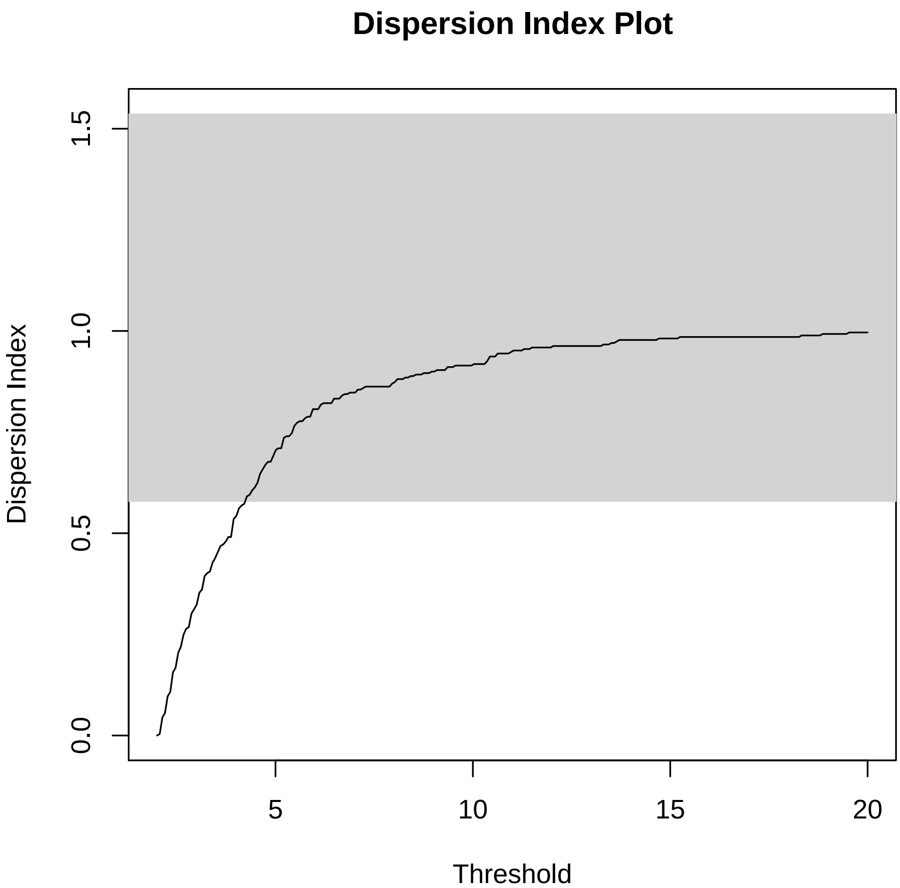

Figure 4: The threshold selection using the diplot function

3.3 Fitting the GPD

3.3.1 The univariate case

The main function to fit the GPD is called **fitgpd**. This is a generic function which can fit the GPD

```
Standard Error Type: observed

Standard Errors
 scale
0.2257

Asymptotic Variance Covariance
       scale
scale  0.05094

Optimization Information
  Convergence: successful
  Function Evaluations: 6
  Gradient Evaluations: 1

> fitgpd(x, thresh = 1, scale = 2, method = "mle")

Estimator: MLE
Deviance: 363.0409
     AIC: 365.0409

Varying Threshold: FALSE
```

```
   scale1     shape1     scale2     shape2      alpha
7.299e-02  4.474e-02  3.098e-02  3.381e-02  2.001e-06

Asymptotic Variance Covariance
        scale1      shape1      scale2      shape2      alpha
scale1   5.327e-03  -2.155e-03   2.007e-06  -1.397e-06   2.200e-11
shape1  -2.155e-03   2.002e-03  -1.774e-07   3.600e-07  -7.491e-11
scale2   2.007e-06  -1.774e-07   9.597e-04  -8.989e-04   1.712e-11
shape2  -1.397e-06   3.600e-07  -8.989e-04   1.143e-03  -1.930e-11
alpha    2.200e-11  -7.491e-11   1.712e-11  -1.930e-11   4.003e-12

Optimization Information
        Convergence: successful
        Function Evaluations: 99
        Gradient Evaluations: 9
```

In the summary, we can see `lim_u Pr[ X_1 > u | X_2 > u] = 0.02` . ThiTf186.07670Td[(.)570Td[(.78(su)1(mmarTf1

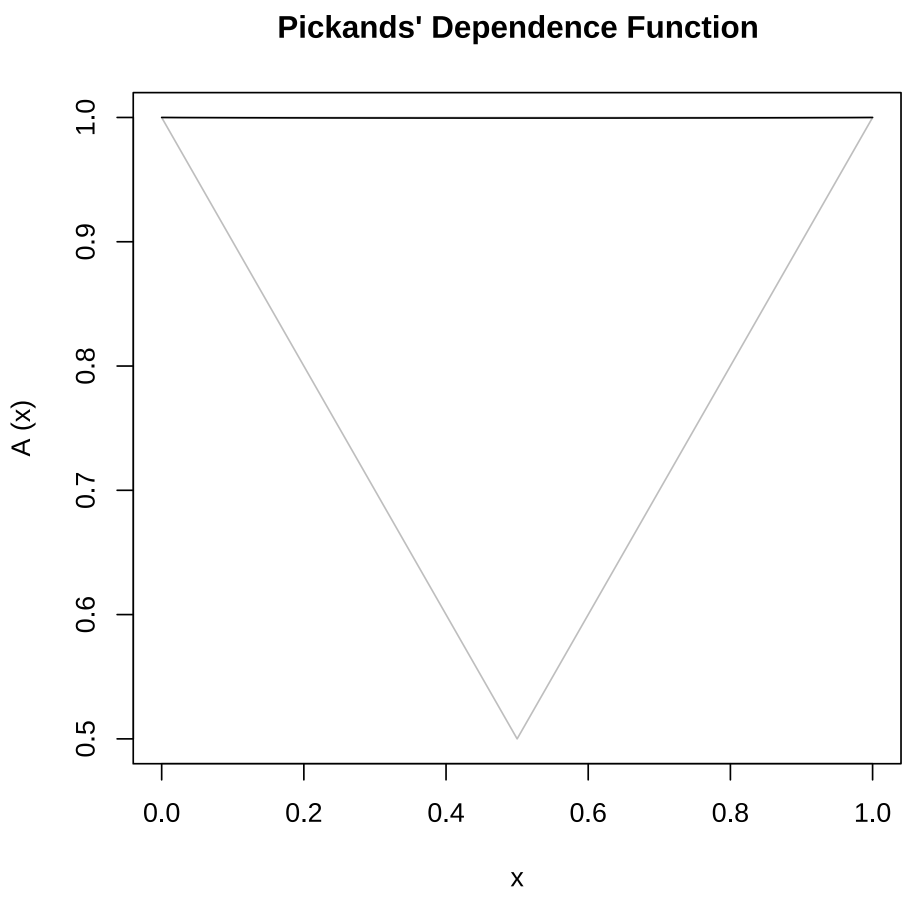

Figure 5: The Pickands' dep8nd8nce function

```
Optimization Information
        Convergence: successful
        Function Evaluations: 53
        Gradient Evaluations: 8
```

Note that as all bivariate ex(i)1(o)reme value dis1(o)ributions are asymptotically dep8nd8nt, the $\bar{\chi}$ s1(o)atis1(o)ic of Coles et al. (1999) is always equal to 1.

Another way to detect the s1(o)r8ngth of dep8nd8nce is to plot the Pickands' dep8nd8nce function – see Figure 5. This is simply done with the `pickdep` function.

```
> pickdep(Mlog)
```

The horizontal line corresponds to indep8nd8nce while the other ones corresponds to p8rfect dep8nd8nce. Please note that by cons1(o)ruction, the *mixed* and *asymetr(i)c mixed* mod8ls can not mod8l p8rfect dep8nd8nce variables.

3.3.3 Markov Chains for Exceedances

structure using a Markov Chains while the joint distribution is obviously a multivariate extreme value distribution. This idea was first introduces by Smith et al. (1997).

In the remainder of this section, we will only focus with first order Markov Chains. Thus, the likelihood for all exceedances is:

$$L(y_1, \ldots, y_n; \theta, \psi) = \frac{\prod_{i=2}^{n} f(y_{i-1}, y_i; \theta, \psi)}{\prod_{i=2}^{n-1} f(y_i;}$$

```
> x <- rgpd(200, 1, 2, 0.25)
> mle <- fitgpd(x, 1, method = "mle")
> mom <- fitgpd(x, 1, method = "moments")
> pwmb <- fitgpd(x, 1, method = "pwmb")
> pwmu <- fitgpd(x, 1, method = "pwmu")
> gpd.fiscale(mle, conf = 0.9)
```

```
If there is some troubles try to put vert.lines = FALSE or change
 the range...
conf.inf conf.sup
1.374242 2.045960

If there is some troubles try to put vert.lines = FALSE or change
 the range...
 conf.inf  conf.sup
0.2454545 0.5424242
```

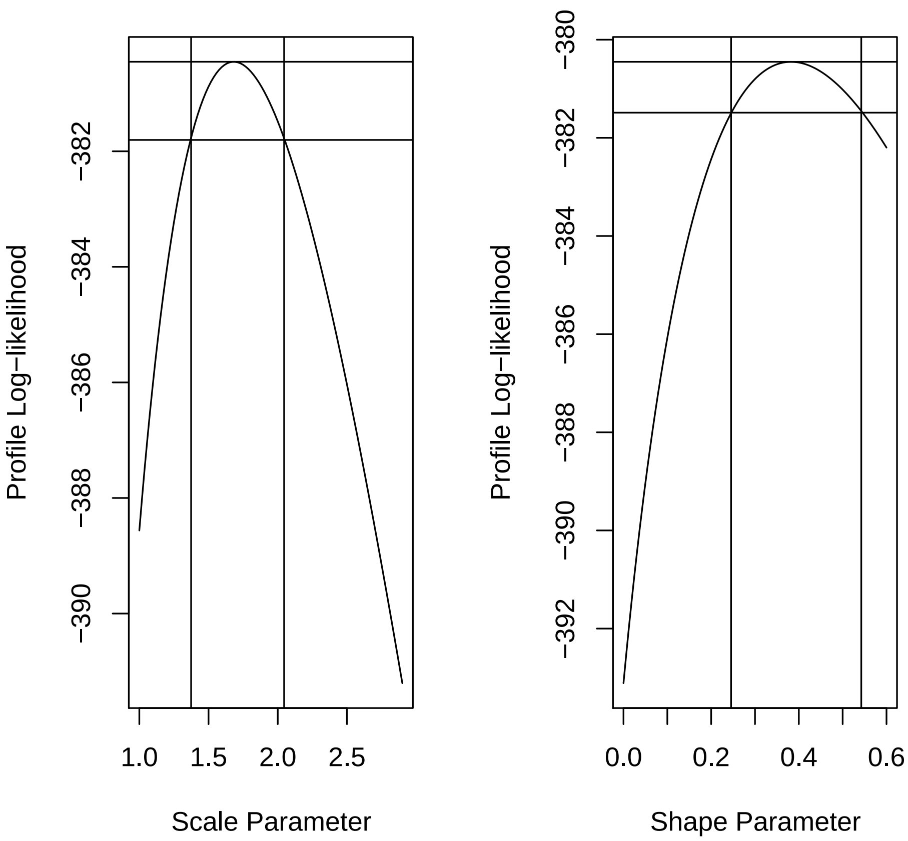

Figure 6: The profile log-likelihood confidence intervals

```
 conf.inf  conf.sup
 7.420834 12.678397
```

```
thn...
conf.inf  conf.sup
```

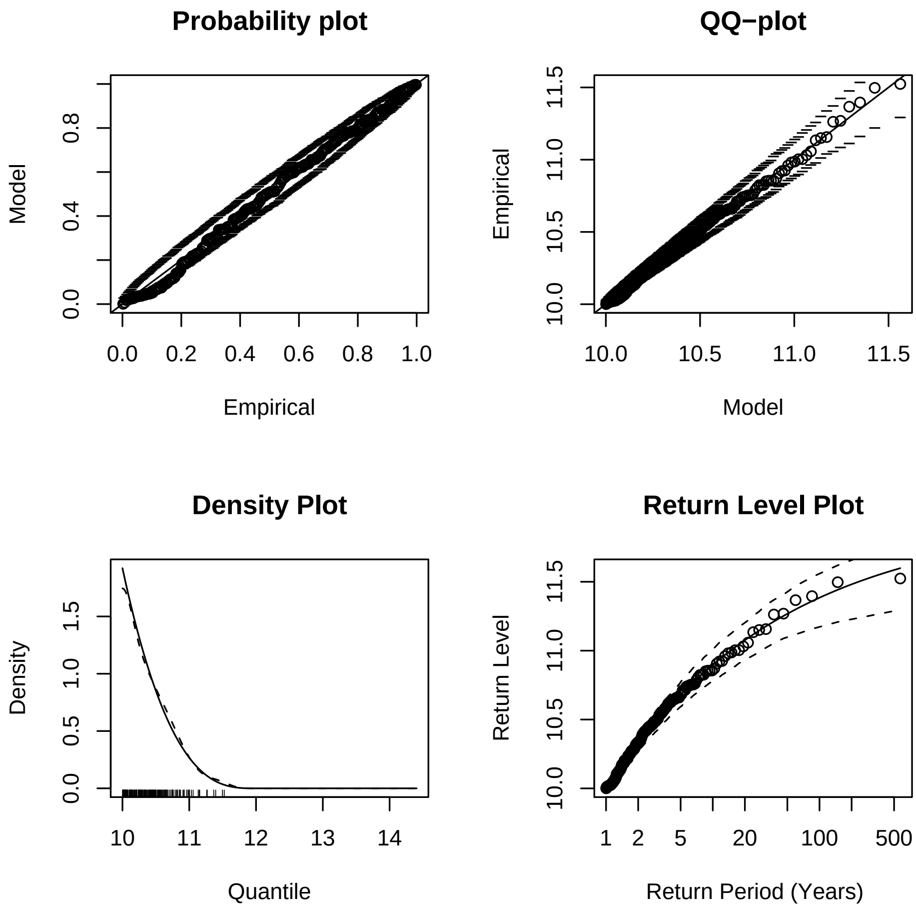

Figure 8: Graphical diagnostic for a fitted POT model (univariate case)

3.5 Model Checking

To check the fitted model, users must call function **plot** which has a method for the `uvpot`, `bvpot` and `mcpot` classes. For example, this is a generic function which calls functions: `pp` (probability/probability plot), `qq` (quantile/quantile plot), `dens`

- which = 4 for a return level plot;

Note that "which" can be a vector like `c(1,3)` or `1:3`.

Thus, the following instruction gives the same graphic.

```
> plot(fitted, which = 1)
> pp(fitted)
```

If a return level plot is asked (4 which), a value for npy is needed. "npy" corresponds to the *mean number of events per year*. This is required o
(i.e. 1) will be chosen.

3.6 Declustering Techniques

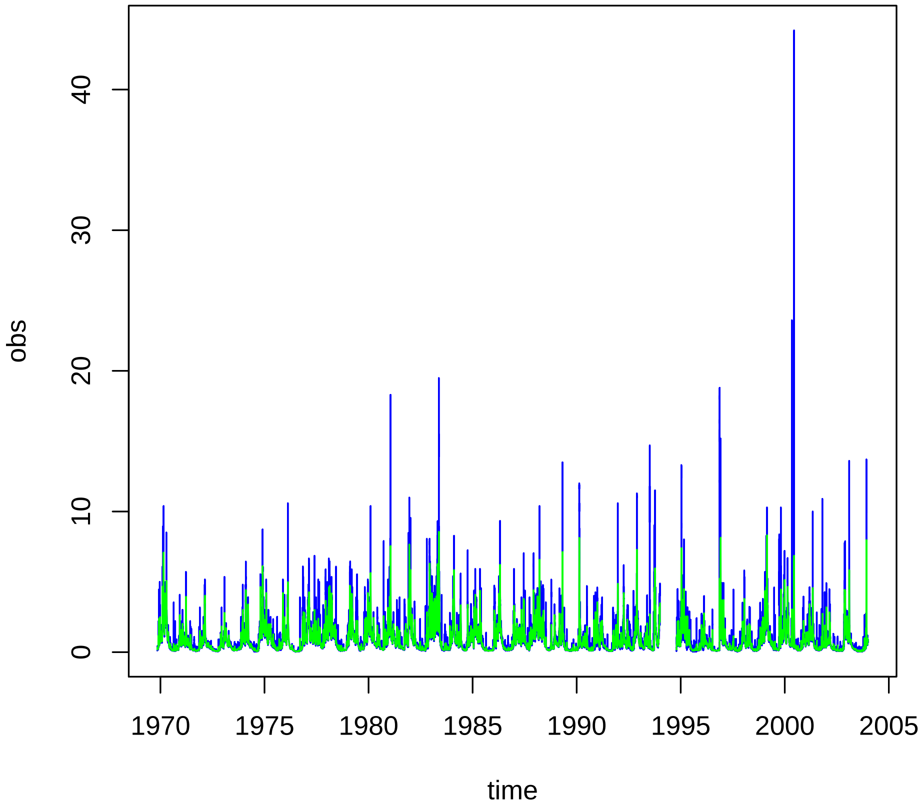

Figure 10: Instantaneous flood discharges and averaged dischaged over duration 3 days. Data ardieres

```
     time            obs
Min.   :1970   Min.   : 0.022
1st Qu.:1981   1st Qu.: 0.236
Median :1991   Median : 0.542
Mean   :1989   Mean   : 1.024
3rd Qu.:1997   3rd Qu.: 1.230
Max.   :2004   Max.   :44.200
               NA's   : 1.000
```

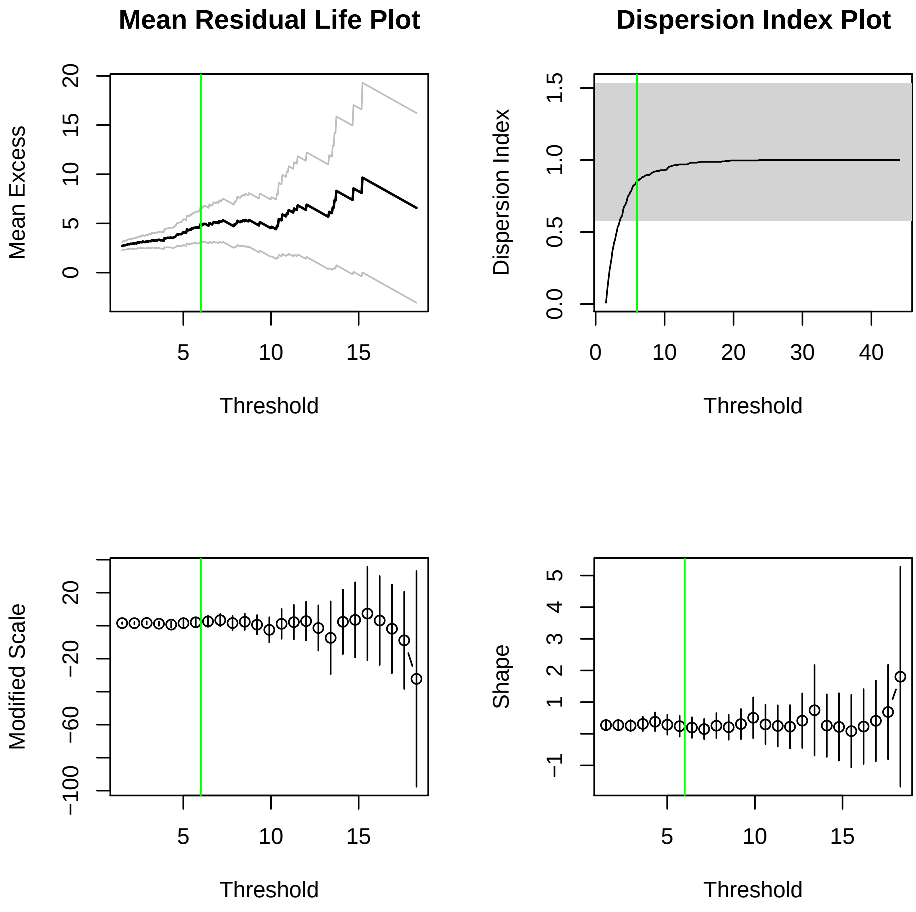

Figure 11: Threshold selection for river Ardières at Beaujeu.

The result of function **fitgpd** gives the name of the estimator, if a varying threshold was used, the threshold value, the number and the proportion of observations above the threshold, parameter estimates, standard error estimates and type, the asymptotic variance-covariance matrix and convergence diagnostic.

Figure 12 shows graphic diagnostics for the fitted model. It can be seen that the fitted model “mle” seems to be appropriate. Suppose we want to know the return level associated to the 100-year return period.

```
> rp2prob(retper = 100, npy = npy)
```

```
       npy retper      prob
1 1.707897    100 0.9941448
```

```
> prob <- rp2prob(retper = 100, npy = npy)[, "prob"]
```

```
If there is some troubles try to put vert.lines = FALSE or change
 the range...
```

A Dependence Models for Bivariate Extreme Value Distributions

A.1 The Logisitic model

The logisitic model is defined by:

$$V(x, y) = \left(x^{-1/\alpha} + y^{-1/\alpha} \right)^{\alpha}, \qquad 0 < \alpha \leq$$

A.5 The Mixed model

Liang Peng and A.H. Welsh. Robust estimation of the generalized pareto distribution. *Extremes*, 4(1): 53–65, 2001.

J. Pickands. Multivariate extreme value distributions. In *Proceedings 43rd Session International Statistical Institute*, 1981.