Solution 1

(a) Write $(1+\xi x/\sigma)_+^{-1/\xi} = \exp\{-\xi^{-1}\log(1+\xi x/\sigma)_+\}$ and note that as $\log(1+a) \sim a + o(a)$ when $a \to 0$, $\lim_{\xi\to 0} \xi^{-1}\log(1+\xi x/\sigma)_+ = x/\sigma$ for any $x/\sigma > 0$. As the exponential function is continuous,

$$\lim_{\xi\to 0} \mathrm{P}(X > x) = \lim_{\xi\to 0} \exp\left\{-\xi^{-1}\log(1+\xi x/\sigma)_+\right\} = \exp\left\{-\lim_{\xi\to 0}\xi^{-1}\log(1+\xi x/\sigma)_+\right\} = \exp(-x/\sigma),$$

for $x/\sigma > 0$, as required.

(b) The support is $S_\xi = \{x : f_X(x) > 0\}$, where $f_X(x) = \sigma^{-1}(1+\xi x/\sigma)_+^{-1/\xi-1}$, with $\sigma > 0$ and $\xi \neq 0$, and $f_X(x) = \sigma^{-1}\exp(-x/\sigma)$ when $\xi = 0$. Hence $S_\xi = \{x : (1+\xi x/\sigma)_+ > 0\}$ when $\xi \neq 0$ and $S_0 = \mathbb{R}_+$. If $\xi > 0$, then $(1+\xi x/\sigma)_+ > 0 \iff (1+\xi x/\sigma) > 0 \iff x > 0$, so $S_\xi = \mathbb{R}_+$. When $\xi < 0$, $(1+\xi x/\sigma)_+ > 0 \iff (1+\xi x/\sigma) > 0 \iff -\sigma/\xi > x > 0$, so in this case $S_\xi = (0, -\sigma/\xi)$. Hence

$$S_\xi = \begin{cases} [0,\infty), & \xi \geq 0, \\ [0, -\sigma/\xi), & \xi < 0. \end{cases}$$

(c) Take $\xi > 0$. Then provided $1 - 1/\xi > 0$, i.e., provided $\xi < 1$, the hint gives

$$\mathrm{E}(X) = \int_0^\infty (1+\xi x/\sigma)_+^{-1/\xi}\,\mathrm{d}x = \left[\frac{\sigma}{\xi}\frac{1}{1-1/\xi}(1+\xi x/\sigma)_+^{1-1/\xi}\right]_0^\infty = \frac{\sigma}{1-\xi}.$$

It is easy to check that when $\xi \geq 1$ the integral is infinite, so $\mathrm{E}(X)$ is undefined.

When $x < 0$ the calculation is the same as above, except that the integral is on the interval $(0, -\sigma/\xi)$, and when $\xi = 0$, $X \sim \exp(1/\sigma)$, so $\mathrm{E}(X) = \sigma$.

If you are curious about the hint, note that

$$\mathrm{E}(X) = \int_0^\infty x f_X(x)\,\mathrm{d}x = \int_0^\infty \int_0^x 1\,\mathrm{d}y f_X(x)\,\mathrm{d}x = \int_0^\infty \int_y^\infty f_X(x)\,\mathrm{d}x\,\mathrm{d}y = \int_0^\infty \mathrm{P}(X > y)\,\mathrm{d}y.$$

(d) We have

$$\mathrm{P}(X > u + x \mid X > u) = \frac{\{1+\xi(x+u)/\sigma\}_+^{-1/\xi}}{(1+\xi u/\sigma)_+^{-1/\xi}} = \left(\frac{1+\xi u/\sigma + \xi x/\sigma}{1+\xi u/\sigma}\right)_+^{-1/\xi} = (1+\xi x/\sigma_u)_+^{-1/\xi},$$

where $\sigma_u = \sigma + \xi u$. This implies that $X - u \mid X > u \sim \mathrm{GPD}(\xi, \sigma_u)$ and therefore that $\mathrm{E}(X - u \mid X > u) = \sigma_u/(1-\xi) = (\sigma + \xi u)/(1-\xi)$, which is linear in u with intercept $\sigma/(1-\xi)$ and slope $\xi/(1-\xi)$.

Solution 2 To run the code we load the packages:

```
load(evd, mev, ismev, scales, lubridate, gridExtra, ggplot2, dplyr, tidyr, ggdist,
 ggpubr, xts)
```

The following code gives Figure 1.

```
load("eskrain.RData")

time.seq <- seq(from=min(date(eskrain)), to=max(date(eskrain)), length=149016)
precip_numeric <- as.numeric(eskrain)
esk.rain <- data.frame(date=as.Date(time.seq), precip=precip_numeric)

u <- 5
plot_esk <- plot(esk.rain, type="h", ylab="Hourly rainfall (mm)", xlab="Time")
points(esk.rain[esk.rain$precip > u,], col="red", cex=.25, pch=20)
abline(u, 0, col="red")
```

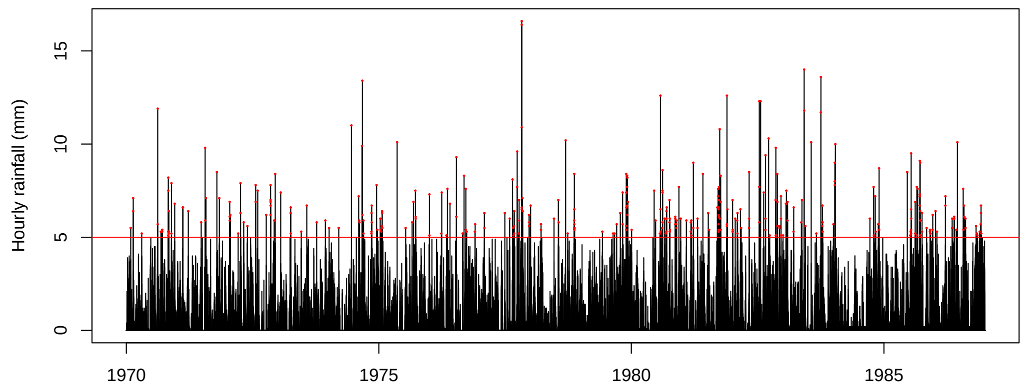

Figure 1: Precipitation from 1970–1986 in Eskdalemuir.

(a) We run the following to find the maxima for the different periods:

```
### start by taking daily maxima
daily.max <- apply(matrix(esk.rain$precip, ncol=24, byrow=T), max)

### then use daily maxima to compute weekly maxima
weekly.max <- apply(matrix(daily.max, ncol=7, byrow=T), max)

### finally, take monthly maxima. We use here months with fixed lengths of 30 days
monthly.max <- apply(matrix(daily.max, ncol=30, byrow=T), max)
```

There are many dry days with zero precipitation, giving a positive probability mass at zero, and this will impact the fit of the GEV to daily data. For instance, analysis of the daily maxima results in a warning on the convergence of the optimisation routine, and gives the following fit

```
(fit.daily <-  fgev(daily.max))

Call: fgev(x = daily.max)
Deviance: 10475.06

Estimates
    loc    scale    shape
0.02312  0.05478  2.19920

Standard Errors
      loc      scale      shape
2.025e-06  1.082e-03  7.285e-02

Optimization Information
  Convergence: iteration limit reached
  Function Evaluations: 1184
  Gradient Evaluations: 100
```

The diagnostic plots in Figure 2 show that the fitted model does not describe the data well.

```
par(mfrow=c(1, 4))
plot(fit.daily)
```

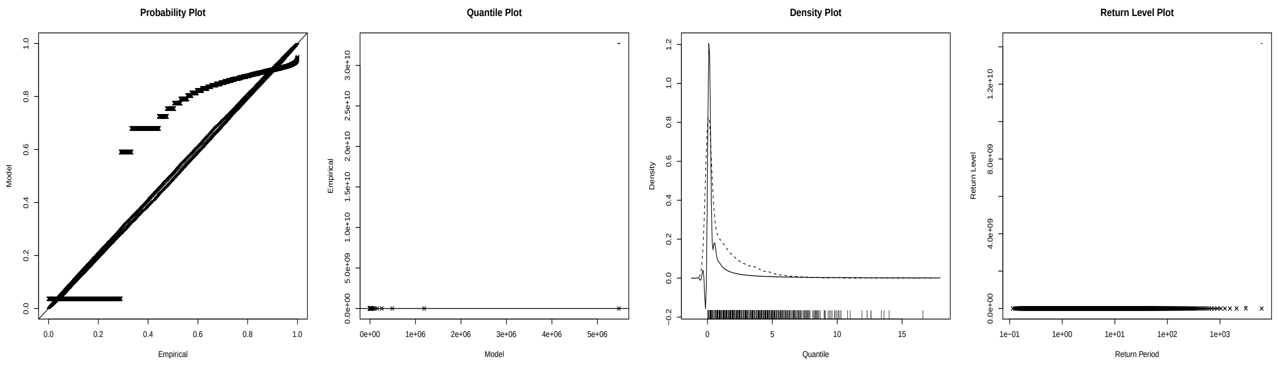

Figure 2: Diagnostic plots of the GEV fit for daily maxima.

Fitting the GEV model to the weekly maxima gives a much better description of the data. The MLEs are $\widehat{\nu} = 2.25$, $\widehat{\sigma} = 1.83$ and $\widehat{\xi} = 0.05$. This is confirmed by the plots in Figure 3. Nonetheless, the probability plots reveal that the fit to the bulk of the distribution is not ideal.

```
(fit.weekly <-  fgev(weekly.max))

Call: fgev(x = weekly.max)
Deviance: 3934.709

Estimates
    loc    scale    shape
2.25220  1.83398  0.05482

Standard Errors
    loc    scale    shape
0.07141  0.05403  0.02997

par(mfrow=c(1, 4))
plot(fit.weekly)
```

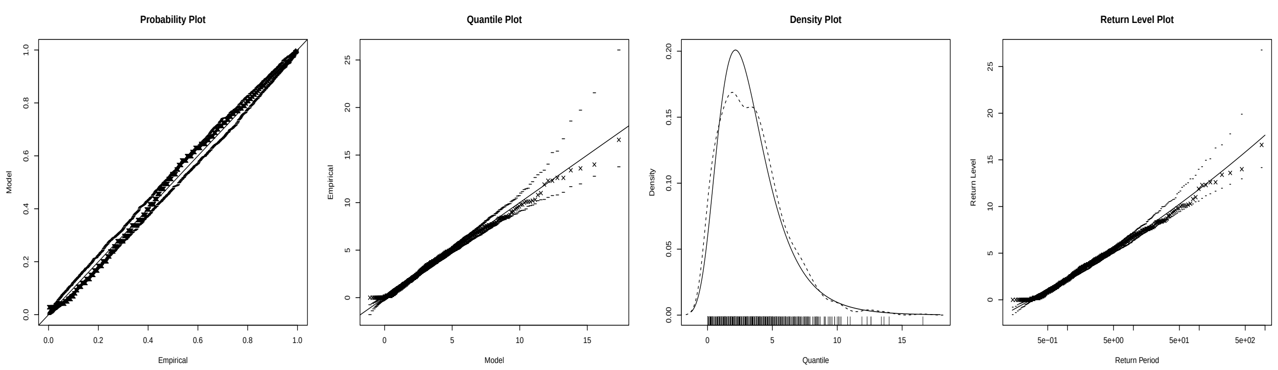

Figure 3: Diagnostic plots of the GEV fit for weekly maxima.

Finally, we analyse the monthly maxima. The resulting MLEs are $\widehat{\nu} = 4.78, \widehat{\sigma} = 2.05$ and $\widehat{\xi} = -0.02$. The diagnostic plots in Figure 4 show a good fit.

```
(fit.monthly <-  fgev(monthly.max))

Call: fgev(x = monthly.max)
```

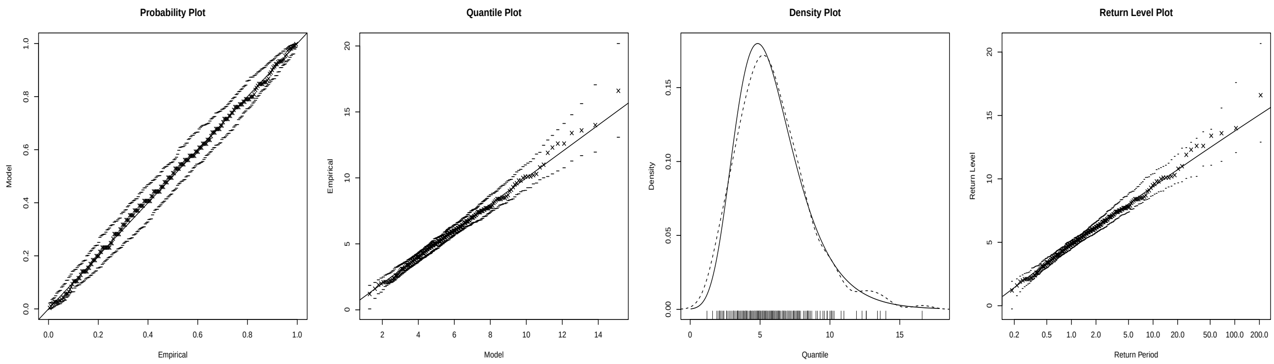

Figure 4: Diagnostic plots of the GEV fit for monthly maxima.

	$\widehat{\mu}$	$\widehat{\sigma}$	$\widehat{\xi}$
Weekly	2.25 (0.07)	1.83 (0.05)	0.05 (0.03)
Monthly	4.78 (0.16)	2.05 (0.11)	-0.02 (0.05)

Table 1: Parameter estimates for the fitted GEV models.

```
Deviance: 944.4168

Estimates
     loc     scale     shape
 4.78404   2.04649  -0.02086

Standard Errors
    loc    scale    shape
0.15925  0.11445  0.04841

par(mfrow=c(1, 4))
plot(fit.monthly)
```

(b) We only look at the models for the weekly and monthly maxima. Table 1 gives the MLEs and their standard errors (in brackets). Suppose that we want to compute 0.95% confidence intervals (CIs). We use standard likelihood theory (e.g., slide 24) to compute the CIs with confidence level $1 - 2\alpha$. Here $\alpha = 0.025$, so $z_{1-\alpha} = 1.96$ and then obtain the CIs in Table 2, using for instance

```
fit.monthly$estimate+qnorm(0.975)*c(-1,1)*fit.monthly$std.err
```

	Weekly	Monthly
$\widehat{\mu}$	(2.11, 2.39)	(4.47, 5.10))
$\widehat{\sigma}$	(1.73, 1.94)	(1.82, 2.27)
$\widehat{\xi}$	(-0.004 , 0.114)	(-0.12, 0.07)

Table 2: 95% confidence intervals for the GEV parameters.

These CIs are based on the normal approximation to the distribution of the estimates, but we should check whether this is reasonable using the profile log likelihood plots to see whether there are asymmetries:

```
par(mfrow=c(1,3))
plot(profile(fit.monthly))
```

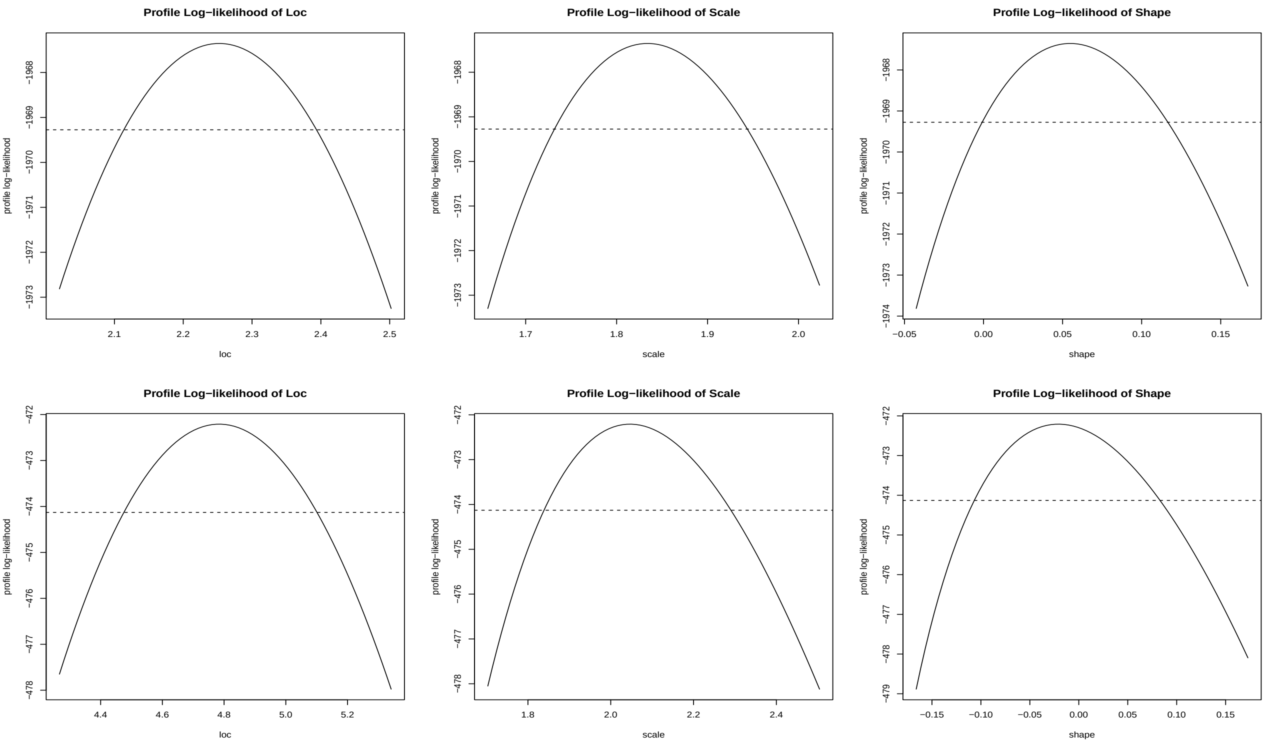

Figure 5: Profile log likelihoods for the three GEV parameters. The top panel corresponds to weekly maxima and the bottom panel to monthly maxima.

Figure 5 shows fairly symmetric profiles, so the confidence intervals above should be reasonable, except maybe for the shape parameter.

(c) The confidence interval for ξ in Table 2 does not exclude the possibility that $\xi = 0$. For a test based on the likelihood ratio statistic between a GEV model with three parameters and the Gumbel model, i.e., the GEV model with only location and scale parameters, we can use the difference of deviances for the two models:

```
fit.monthly0 <- fgev(monthly.max, shape=0)
ratio <- fit.monthly0$dev - fit.monthly$dev)
qchisq(0.95,1)
3.841459
1-pchisq(ratio,df=1)
0.672707
```

As the p-value equals 0.67 and `ratio` does not exceed the theoretical 95% quantile of a χ^2_1 random variable, we cannot reject the Gumbel model under which $\xi = 0$.

Solution 3 We now study the excess precipitation above a threshold u for the Eskdalemuir rain data.

(a) We start by looking at the mean residual life plot

```
mrlplot(esk.rain$precip)
```

Figure 6 shows that the mean excesses exhibit a rather stable 'linear' behaviour for $2 \le u \le 5$, see also slide 92. There is a slight upward trend in the figure, suggesting that $\xi > 0$. These impressions are confirmed by the parameter stability plots in Figure 7, which show stable behaviour of the estimates for such values of u (though they are less suggestive that $\xi > 0$):

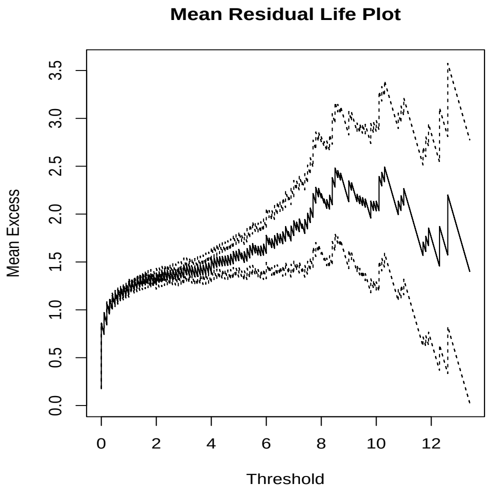

Figure 6: Mean residual life plot

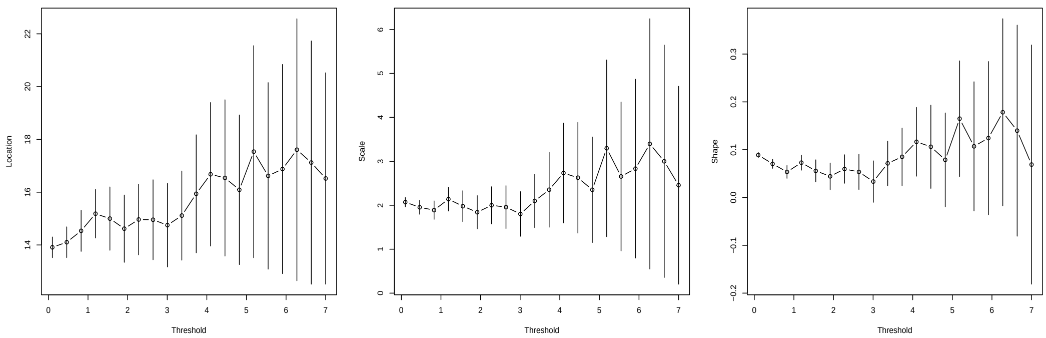

Figure 7: Parameter stability plots

```
tcplot(esk.rain$precip, tlim=c(0.1,7), model="pp", nt = 20)
```

(b) We run the code

```
# here we take a fixed threshold u=5, but you can choose u based on part (a)
thresh.fit <- fpot(esk.rain$precip, threshold=5, model="pp", start=list(loc=10,
 scale=1.2, shape=0.1), npp=365.25*24)

Call: fpot(x = esk.rain$precip, threshold = 5, model = "pp", start = list(loc = 10,      sca
Deviance: -394.7777

Threshold: 5
Number Above: 356
Proportion Above: 0.0024

Estimates
     loc     scale     shape
10.13628   1.86637   0.06696

Standard Errors
    loc    scale    shape
0.35380  0.23673  0.05379
```

```
# next, we look at the resulting diagnostic plots
par(mfrow=c(1,4))
plot(thresh.fit)
```

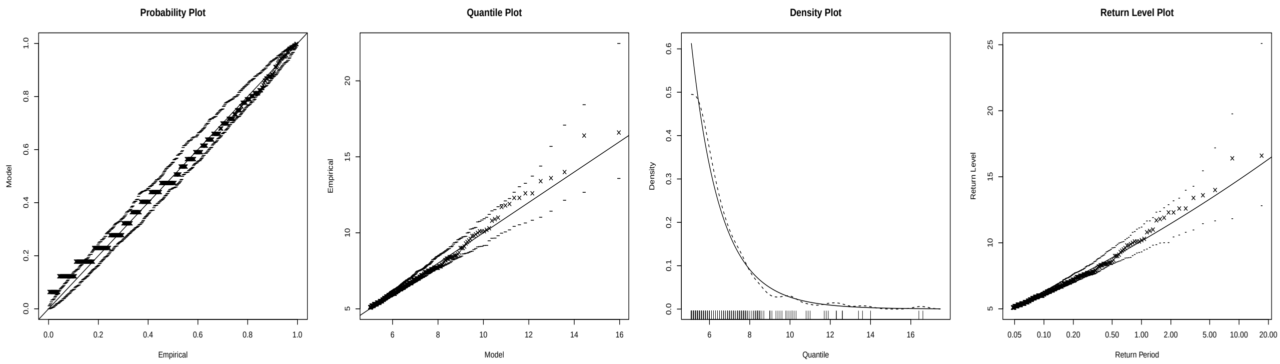

Figure 8: Diagnostic plots for the Poisson process fit

The diagnostic plots in Figure 8 show that the Poisson process fits the data quite well.

(c) To compare the fit of the GPD model with the Poisson process in (b) we take $u = 5$:

```
(thresh.fit_gpd <- fpot(esk.rain$precip, threshold=5, start=list(scale=1.2,
 shape=0.1)))

 Call: fpot(x = esk.rain$precip, threshold = 5, start = list(scale = 1.2,
  shape = 0.1))
Deviance: 1058.954

Threshold: 5
Number Above: 356
Proportion Above: 0.0024

Estimates
  scale    shape
1.52216  0.06725

Standard Errors
  scale    shape
0.11488  0.05387

# next, we look at the resulting diagnostic plots
par(mfrow=c(1,4))
plot(thresh.fit_gpd)
```

The shape parameter estimates are similar, but but the scale parameter estimates are a it different, because the Poisson process fit estimates τ, but the GPD fit estimates $\sigma_u = \tau + \xi(\eta - u)$, and here ξ is not estimated to be zero. A comparison between Figures 8 and 9 shows good fits in both cases.

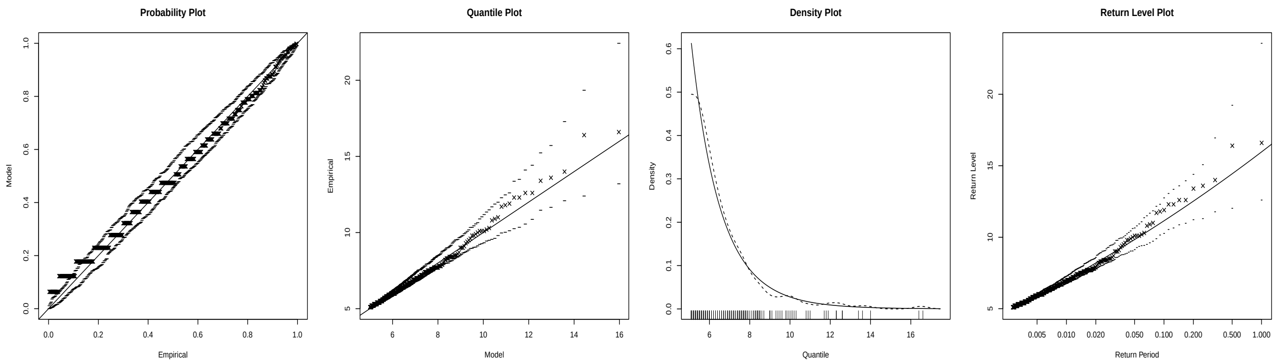

Figure 9: Diagnostic plots for the GPD fit