Problem 1 Maxima of blocks of m independent background observations with distribution function F are modelled by a fitted GEV G, and a T-block return level is sought.

(a) Explain why it would be reasonable to solve the equation $G(x_p) = 1 - 1/T$, and show that this gives $x_p \doteq \eta + \tau(T^\xi - 1)/\xi$ for large T.

(b) Another possibility is to solve the equation $F(x_p) = 1 - 1/(mT)$. Show that this yields $x_p = \eta + \tau([-m\log\{1 - 1/(mT)\}]^{-\xi} - 1)/\xi$, and deduce that this gives the same approximation as in (a).

(c) Suppose that a 20-year return level is to be estimated based on the weekly maxima of hourly background data. Give the appropriate values of p, m and T, and, supposing that $\eta = 0$, $\tau = 1$ and $\xi = 0.1$, compute the exact and approximate values of x_p from (a) and (b). Comment.

Problem 2 Suppose that $X_1, X_2, \ldots \stackrel{\text{iid}}{\sim} \text{GPD}(\sigma, \xi)$ and that N has a Poisson distribution with mean λ.

(a) Find the distribution of $M = \max(X_1, \ldots, X_N)$. How does it differ from a GEV distribution?

(b) Rainfall maxima are increasing, and the question arises whether this is because there are more large values, or because the values themselves are larger. Suggest how the result in (a) could be used to investigate this, based only on block maxima (i.e., the background data are unavailable).

Problem 3 A random sample $X_1, \ldots, X_n$ is available from a distribution F that satisfies the extremal types theorem with sequences $\{a_n\} > 0$ and $\{b_n\}$. Let $u_n = b_n + a_n u$ and $p_u = \text{P}(X_j > u_n)$, and suppose that the generalized Pareto distribution can be used to approximate the distribution of $X - u_n$ conditional on $X > u_n$. Show that log likelihood based on the observations $x_1, \ldots, x_{n_u}$ that exceed u_n can be written as

$$L = p_u^{n_u} \times (1 - p_u)^{n - n_u} \times \prod_{j=1}^{n_u} h(x_j - u_n) = L_1 \times L_2 \times L_3,$$

say, where $h(x - u) = \{-\dot{\Lambda}(x)\}/\Lambda(u)$, with $\Lambda(z) = \{1 + \xi(z - \eta)/\tau\}_+^{-1/\xi}$, where $x > 0$ and $u, z \in \mathbb{R}$.

(a) Show that $np_u \to \Lambda(u)$ and deduce that $n^{n_u} L_1 \to \Lambda(u)^{n_u}$ and $L_2 \to \exp\{-\Lambda(u)\}$ as $n \to \infty$.

(b) Show that

$$n^{n_u} L \to \exp\{-\Lambda(u)\} \prod_{j=1}^{n_u} \{-\dot{\Lambda}(x_j)\},$$

and deduce that the likelihoods based on threshold exceedances and on the point process approximation should give similar inferences for large n.

Problem 4 A moving maximum process (Example 24) with standard Fréchet marginal distribution can be simulated using the code

```
n <- 10000; a <- 1; i <- c(1:n)
z <- 1/rexp(n+1)   # independent Frechet variables
x <- pmax(a*z[i],z[i+1])/(a+1)  # moving maximum series
par(mfrow=c(1,2))  # two adjacent panels for figures
plot(i,x,log="y",pch=20)  # should see clustering of high values, but need log axes
qqplot(z,x,log="xy") # compare the marginal distributions of x and z
```

Here `z` contains independent data for which $\theta = 1$, since there should be no clustering, whereas `x` has clustered data. To see the effect of the clustering, we estimate the extremal index θ over a range of quantiles for each of these vectors:

```
tl <- quantile(z, probs = c(0.1,0.95))
exiplot(z,tl)  # plots estimated theta between the limits given by tl
abline(h=1,col="red")
tl <- quantile(x, probs = c(0.1,0.95))
exiplot(x,tl)
abline(h=max(a,1)/(a+1),col="red")
```

The red lines show the true values of θ for very large n; recall that $\theta = \max(a, 1)/(a + 1)$ for the moving maximum process.

Discuss the difference between the two plots. Try with other values of a. Do you see what you expect?