Australian National University
Centre for Mathematics and its Applications (CMA)
Geoscience Australia
Canberra

STATISTICAL MODELING OF EXTREME VALUES, 2008
COMPUTER ASSIGNMENT 3

Modeling extremes of non-stationary sequences

You should read Chapters 6 and 7 in both `Coles` book and the `extRemes` tutorial before you start this computer session.

Maximum sea-level Data

The `R` source file `fremantle.R` contains annual maximum sea-level data in Fremantle, south Australia. Note that the library `extRemes` does not yet have all the functionality needed to carry out this analysis so you should start by directly attaching the package `ismev` to `R`. Load the dataset to `R` by running this command `data(fremantle)` from `R`.
Analyze the data by incorporating different trend models in the parameters of GEV according to the five models discussed in the lectures; see also the discussion in Section 6.3.1 of the Coles book. Use likelihood ratio test to compare different models and choose a model which seems to fit to the data best. The `R` function `gev.fit` in `ismev` package can be used to include covariates in the block maxima model so before start you should read pages 7-10 in the accompaniment to the Coles library and the help page for`gev.fit` to become familiar with the syntax of the function.

1. Analyze different models by likelihood ratio test. Which model does seem to be adequate for this dataset?
   Answer:

Wooster Temperature Series

Daily minimum winter temperature (degrees below 0 F.) are given in the R source file `wooster.R`. If you saved your workspace after the computer assignment 2, the `wooster` dataset will be available in your working directory otherwise you have to read it into `extRemes` as usual.
In Computer Assignment 2 in order to obtain an approximately stationary series we considered only the temperatures in winter. As discussed in Chapter 7, it is possible to analyze the whole dataset by modeling exceedances over a time-varying threshold. In Section 7.7, six different models have been briefly discussed. Repeat the analysis for all the models and compare them by using likelihood ratio test; see also Table 7.1 in the Coles book. Include the estimates of the parameters for each model in your report and enclose a number of plots which are useful to understand and compare the models. The R function `pp.fit` will be used in this analysis. The full syntax of `pp.fit` has been discussed in the accompaniment to Coles book, pages 18-21 which is also available in the online help system in R. To construct the model matrix `ydat` in R, the following commands might be used. For convenience, these commands are also stored in the file `assignment3.txt` in the following location:

- `http://www.maths.lth.se/matstat/kurser/fms155mas231/datasetsR.html`

You can copy and paste from this file to R directly.

```
x = 1:length(wooster$data[,2])
usin = function(x, a, b, d)
{
        a + b * sin(((x - d) * 2 * pi)/365.25)
}
wu = usin(x, -30, 25, -75)
winter = c(rep(c(rep(1, 61), rep(0, 273), rep(1, 31)), 5),
        1)
spring = c(rep(c(rep(0, 61), rep(1, 91), rep(0, 365 - 91 -
        61)), 5), 0)
summer = c(rep(c(rep(0, 61 + 91), rep(1, 91), rep(0, 365 -
        91 - 61 - 91)), 5), 0)
fall = c(rep(c(rep(0, 61 + 91 + 91), rep(1, 91), rep(0, 365 -
        91 - 61 - 91 - 91)), 5), 0)
rescale.covariate  = function(x)
{
```

```
r.x = range(x)
x.01 = (x-r.x[1])/diff(r.x)
2*x.01-1
}
ydat = cbind(sin((x * 2 * pi)/365.25), cos((x * 2 * pi)/365.25
        ), rescale.covariate(x), winter, spring, summer, fall)
```

In the matrix `ydat` the first and second columns will be used for modeling periodic trend in parameters and column 3 contains the rescaled index of time which can be used for linear trend. For example to obtain the estimates in models 3 and 6 the following commands can be used, respectively.

```
fit.model3 = pp.fit( - wooster$data[,2], threshold = wu, ydat = ydat,
mul = 1:2, sigl = 1:2, siglink = exp, method=''BFGS'')
```

```
fit.model6 = pp.fit( - wooster$data[,2], threshold = wu, ydat = ydat,
 mul = c(1, 2), sigl = c(1, 2), shl = 4:7, siglink = exp, method=''BFGS'')
```

Here the argument `method=''BFGS''` makes the program to use quasi-Newton method as optimization algorithm.

1. Analyze your results. Which model would be adequate for this dataset?
   Answer: