# Extreme value analysis of the annual maxima of daily and monthly total precipitation in Scotland

Thomas Uriot, s1216723@ed.ac.uk

June 28, 2021

**Abstract**

An extreme value analysis based on the block maxima approach is conducted on rain data in the South, East and North of Scotland. A GEV density is fitted to the sequence of maxima and the estimated return levels with their associated confidence intervals are computed, for various return periods. It is shown that the East region has low return levels for both daily and monthly maxima. The South has the highest return level for daily precipitation while the North has the highest return levels for monthly precipitations. It is also argued that extrapolation made on the monthly and daily data in South Scotland may not be appropriate since the quantile plots depart from linearity.

**Keywords:** Gumbel distribution; generalized extreme value; extreme value statistics; block maxima; rain; Scotland

# 1 Introduction

## Background

The field of extreme value theory or extreme value analysis was pioneered by the British statisticians Leonard Tippett (1902–1985) and R.A Fisher (1890-1962). Tippett worked on making cotton thread stronger when he realized that the strength of a thread was controlled by the strength of its weakest fibres. Tippet and Fisher then obtained three asymptotic limits describing the distributions of extremes known as the Fisher–Tippett–Gnedenko theorem. In contrast to classical statistics which are more focused on the mean or median of a probability distribution, extreme value analysis deals with extreme deviations from those values. While the most common applications of extreme value analysis are in finance (insurance in particular) and natural hazards, there are new applications cropping up in genetics (mutations are considered as rare events) or athletics to model the probability of a new world record occurring. A typical example of the use of extreme value analysis concerns flooding. When building a dam, it is of interest to know by how much the water level could rise up to and how likely it is to happen.

Two approaches exist for practical extreme value analysis. The first method consists of blocking the data into sequences of observations of euqal length. For each sequence (block), the maximum (minimum) is extracted, generating a series of maxima (minima). A distribution called the *generalized extreme value* (GEV) distribution is then fitted to these extremes values, as an approximation. The second method relies on extracting, from a continuous record, the peak values reached for any period during which values exceed (falls short) a certain threshold. This method is called the peak over threshold method and gives rise to the generalized Pareto (GP) as an approximation for the excess over a high threshold. The main differences between these methods is that in the later case, there may be several or no values being extracted in any given time period, depending on how extreme the threshold is. The block maxima approach always leads to a single value being extracted for each time period.

### Literature review

The paper by Richard W. Katz (2010) is a good starting point as it offers a short historical account of extreme value statistics as well as future challenges. It also provides a simple comparison between the block maxima and the peak over threshold methods. Rivest. A (1989) provides a characterization of the Gumbel's family of extreme value distributions proposed by Gumbel (1960). The Gumbel distribution is a particular case of the GEV distribution, which arises as a limit of the GEV. For applications of the Gumbel distribution, R.I. Harris (1996) analyses annual maximum wind speeds using transformed data. An application of the peak over threshold method is given in J. A. Ferreira and C. Guedes Soares (1998), where the authors investigates wave height data. A major current topic of interest is the analysis of temporal trends. Climate change is one of the main area of the application where temporal trends are present. This is because 'greenhouse' effect may result is gradually more frequent natural hazards such as storms and floods. The paper by P. Hall and N. Tajvidi (2000) propose fitting nonparametric models to extreme value data with temporal trends. Finally, an analysis on rainfalls is given in the paper by Sergio M. (2005). The author makes use of the peaks over threshold approach, and fits a Poisson–generalized Pareto model. This can provide an alternative method to the one based on the block maxima presented here.

### Aims and plan

As mentioned, this paper is based on the block maxima method to analysis the annual maxima of daily and monthly total precipitation in 3 different geographical locations in Scotland (South, North, East). A GEV distribution is fitted to the maximas and the return levels are computed for different return periods. Being able to obtain accurate estimates for the return levels is very important in practice since extreme precipitation events are in the origin of severe erosion, landslide triggering, or flash floods, which can have regionally devastating power and pose a severe hazard to lives and property. Previous extreme value analysis have been applied to rain falls in the UK but none of them make use of up-to-date data.
The aim of this paper is to perform an extreme value analysis on the 3 regions within Scotland, compute the return levels and their associated confidence intervals for each region, and compare them. Firstly, the data used and the block maxima method will be exposed. Then, the extreme value analysis will be carried out and the results presented.

# 2 Data and theory

## 2.1 Data

The data used is taken from the Met Office Hadley Centre observations datasets (HadUKP). The monthly and daily precipitations series for Scotland (and sub-regions) begin in 1931. There are two precipitations series (monthly and daily) per region (East, North, South), giving 6 series in total, recorded in millimeters. The monthly series all start in 1931 and finish in 2015. The daily series start in 1931 and are still ongoing in 2016. However, since there are not enough days recorded in 2016, the data for this year is left out of the analysis.

## 2.2 Block maxima approach and GEV fitting

As was briefly discussed in the introduction, if data are blocked into $n$ sequences of iid observations $X_1, \ldots, X_N$ of equal length $m$, generating a series of block maxima $M_1, \ldots, M_n$, then the $M_i$ are assumed to be distributed according to a GEV distribution. If the $X_i$ are independent then, the $M_i$ are also independent. However, independence of the $M_i$ is a reasonable approximation even though the $X_i$ form a dependent series. The choice of the length of blocks is important in order to have identically distributed $X_i$. In the case of precipitations (or any other weather related variables), blocks of one year means that the assumption that individual block maxima have a common distribution is plausible, as it gets ride of the seasonal effects.Figure 1 shows the time series plots of the block (yearly) maxima and of all the months in the data for the monthly precipitations series in East Scotland. This is shown as an illustration of independence, one could choose either monthly or daily and any regions.

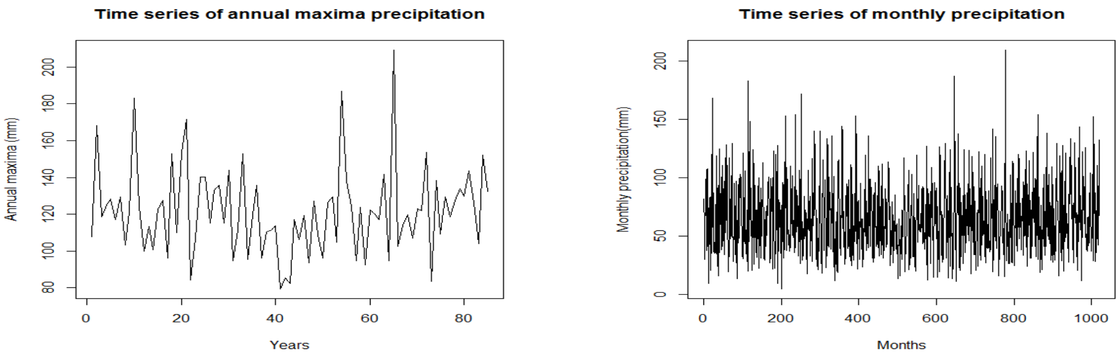


Figure 1: Time series plot for the East Scotland monthly precipitation data.

One can see that the average value of precipitation over some time boxes of same length is not the same, which implies that there is independence between the $X_i$ (months). The GEV distribution fitted to the block maxima $M_i$ is given by

$$G(x; \mu, \sigma, \xi) = \exp\left[-\left(1 + \xi \frac{x - \mu}{\sigma}\right)_+^{-\frac{1}{\xi}}\right], \quad -\infty < \mu, \;\; \xi < \infty, \;\; \sigma > 0, \tag{1}$$

where the range of x is such that $1 + \xi(x - \mu)/\sigma > 0$. The parameter $\xi$ controls the shape of the density which, has heavy right tail and finite lower support point if $\xi > 0$, and a finite upper support point if $\xi < 0$. The location parameter $\mu$ specifies the center of the distribution and the scale parameter $\sigma$ determines the size of deviations about the location parameter. The Gumbel distribution

$$G(x; \mu, \sigma, 0) = \exp\left[-\exp\left(-\frac{x - \mu}{\sigma}\right)\right], \quad -\infty < x < \infty, \tag{2}$$

arises as a particular case of (1) as $\xi \to 0$. Figure 2 below gives the fitted GEV densities plotted over the histograms for both daily and monthly annual maxima in the 3 geographical regions. Table 1 summarizes the values of maximum likelihood estimates of the parameters. One can see how $\xi$ affects the shape of the GEV density.

| Parameters | East Daily | East Monthly | North Daily | North Monthly | South Daily | South Monthly |
|---|---|---|---|---|---|---|
| $\hat{\mu}$ | 22.9 | 111.40 | 31.1 | 240 | 31 | 206.83 |
| $\hat{\sigma}$ | 6.3 | 19.69 | 5.26 | 40.23 | 6.89 | 37.74 |
| $\hat{\xi}$ | -0.025 | -0.053 | -0.063 | -0.212 | 0.022 | -0.226 |

Table 1: Maximum likelihood estimators.

The value of $\hat{\xi}$ is almost 0 for all of them except for the monthly annual maxima in the South and North of Scotland where the density is not as peaked around $\mu$. For the low estimates of $\xi$, the densities are close to Gumbel densities.

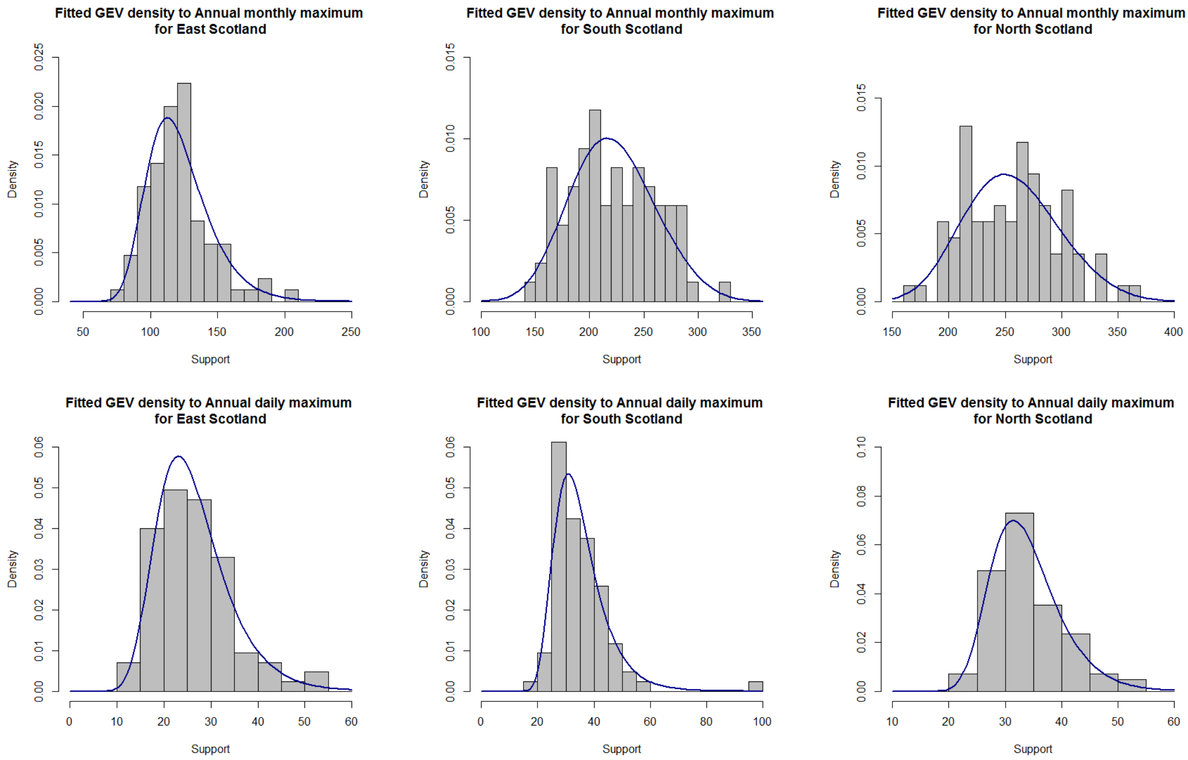


Figure 2: Time series plot for the East Scotland monthly precipitation data.

## 2.3 Return levels

The level that is expected to be exceeded on average once in $T = 1/p$ years is called the return level associated with return period T and is given by the level $x_p$ that satisfies

$$G(x_p) = 1 - p,$$

and so

$$x_p = \begin{cases} \mu - \frac{\sigma}{\xi}\left[1 - (-log(1-p))^{-\xi}\right] & \xi \neq 0 \\ \mu - \sigma \log[-\log(1-p)] & \xi = 0. \end{cases} \tag{3}$$

One can the maximum likelihood estimate $\hat{x}_p$ of $x_p$ by substituting the maximum likelihood estimates of the parameters $\mu$, $\sigma$ and $\xi$ of the GEV density. The maximization of the likelihood function is done using the software R and the code is provided in Appendix A. The estimated return levels will be computed for each location and for both daily and monthly annual maxima at different return periods T. Associated 95% confidence intervals will also be computed. The confidence interval for the estimated return level $\hat{x}_p$ is computed using the normalized profile likelihood defined as

$$D(x_p) = 2[l(\hat{\mu}, \hat{\sigma}, \hat{\xi}) - l(x_p, \hat{\sigma}, \hat{\xi})], \tag{4}$$

where $D(x_p) \sim \chi_1^2$. In (4), $l(\hat{\theta})$ denotes the log likelihood evaluated at the parameters. the parameter $x_p$ appears after solving for $\mu$ in terms of $\sigma$, $\xi$ and $x_p$ in (3). Using that the normalized profile likelihood is distributed as a $\chi_1^2$, one can obtain a confidence interval $C_\alpha$ such that

$$C_\alpha = [x_p : D(x_p) \leq C_\alpha]$$

.

Figure 3 shows a couple of plots of normalized profile likelihood for the daily and monthly annual maxima in East Scotland, evaluated at $p = 0.01$. The vertical line is the 95th quantile of$\chi_1^2$ which is 3.84. The values of $x_{0.01}$ for which the profile likelihood is smaller than 3.84 form the 95% confidence interval for $x_{0.01}$.

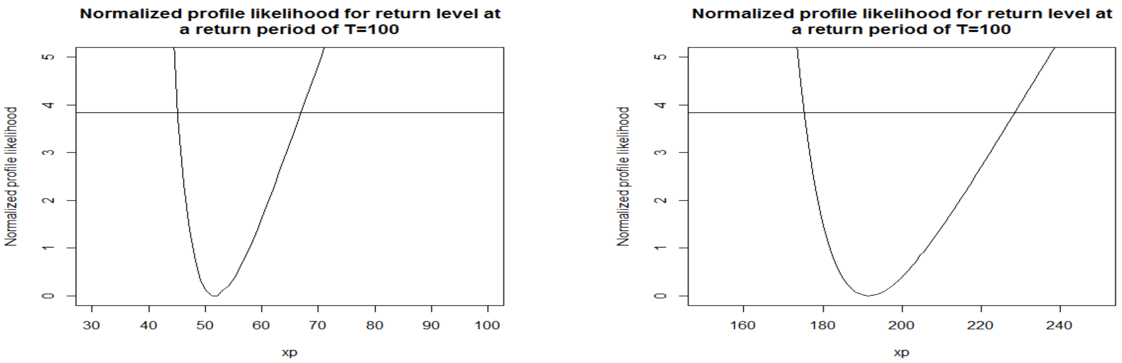


Figure 3: Normalized profile likelihood plots for return period of T=10 years. Daily annual maxima for East Scotland is on the left, monthly annual maxima on the right.

# 3 Analysis and results

## 3.1 Results

Figure 3 shows the return level plots for the 6 different precipitations series. The plots separate daily and monthly. The return level is plotted against $-\log[-\log(1-p)]$. When $\xi = 0$, the plot is linear, if $\xi < 0$, the plot is concave with asymptotic limit $\mu - \sigma/\xi$ as $T \to \infty$ (i.e $p \to 0$). If $\xi > 0$, the plot is convex and has not finite bound.

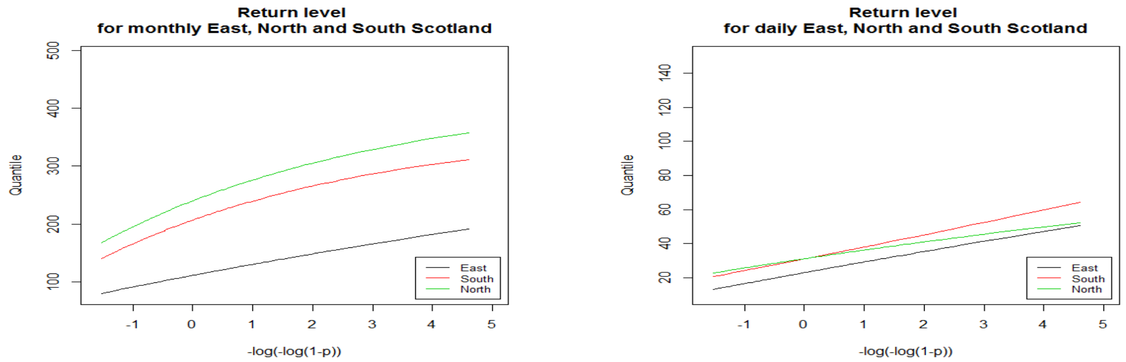


Figure 4: Return level plots of the GEV distribution for monthly and daily aannual maxima.

Looking at Table 1, one can see that $\hat{\xi}$ is clearly smaller than 0 for North and South monthly annual maxima and the plots of the return level are concave. The estimate of the asymptotic limit for the South Scotland is $206.83 - (37.74/(-0.226)) = 373$. The interpretation of that result is that a level of 373 mm of rain in a month is never expected to be exceeded in the South. Similarly, for North Scotland, one has $240 - (40.23/(-0.212)) = 429$. Next comes the most important part of the paper. The return level for several return periods (10, 100, 1000 and 10000 years) and their associated 95% confidence intervals are computed. My findings for the estimates of the return levels and their confidence intervals for some return periods $T$ are computed as explained in section 2.3 and summarized in Table 2 and Table 3 below. Table 2 is for daily annual maxima and Table 3 for monthly annual maxima.

| T (years) | East | North | South |
|---|---|---|---|
| 10 | 37(35 41) | 42(40 45) | 47(43 52) |
| 100 | 51(45 66) | 52(48 62) | 64(58 89) |
| 1000 | 63(53 103) | 61(53 84) | 82(70 138) |
| 10000 | 75(58 122) | 68(56 114) | 101(81 154) |

Table 2: Estimated return levels for daily annual maxima and associated 95% confidence intervals (given in brackets).

| T (years) | East | North | South |
|---|---|---|---|
| 10 | 153(146 164) | 312(301 327) | 273(263 288) |
| 100 | 191(175 227) | 358(341 403) | 315(341 403) |
| 1000 | 225(197 285) | 385(360 446) | 339(315 399) |
| 10000 | 254(214 314) | 402(368 463) | 353(322 413) |

Table 3: Estimated return levels for monthly annual maxima and associated 95% confidence intervals (given in brackets).

As explained in section 2.3, it is expected that the level of monthly annual precipitations in East Scotland exceeds 254 mm once every 1000 years.

## 3.2 Model testing

The assessment of the GEV model can be made with reference to observed data. Quantile plots are used here and are displayed in Figure 5. Departures from linearity in the quantile plot indicate model failure. One can see that the quantile plots for monthly South ans North seem to have a cubic trend, especially pronounced for South Scotland. In addition, the quantile plot for South Scotland seems to be more logarithmic that linear and contains a clear outlier on the top right of the plot. These quantiles plots indicate that extrapolation made on the monthly and daily data in South Scotland may not be appropriate. The remaining plots do not depart significantly from linearity.

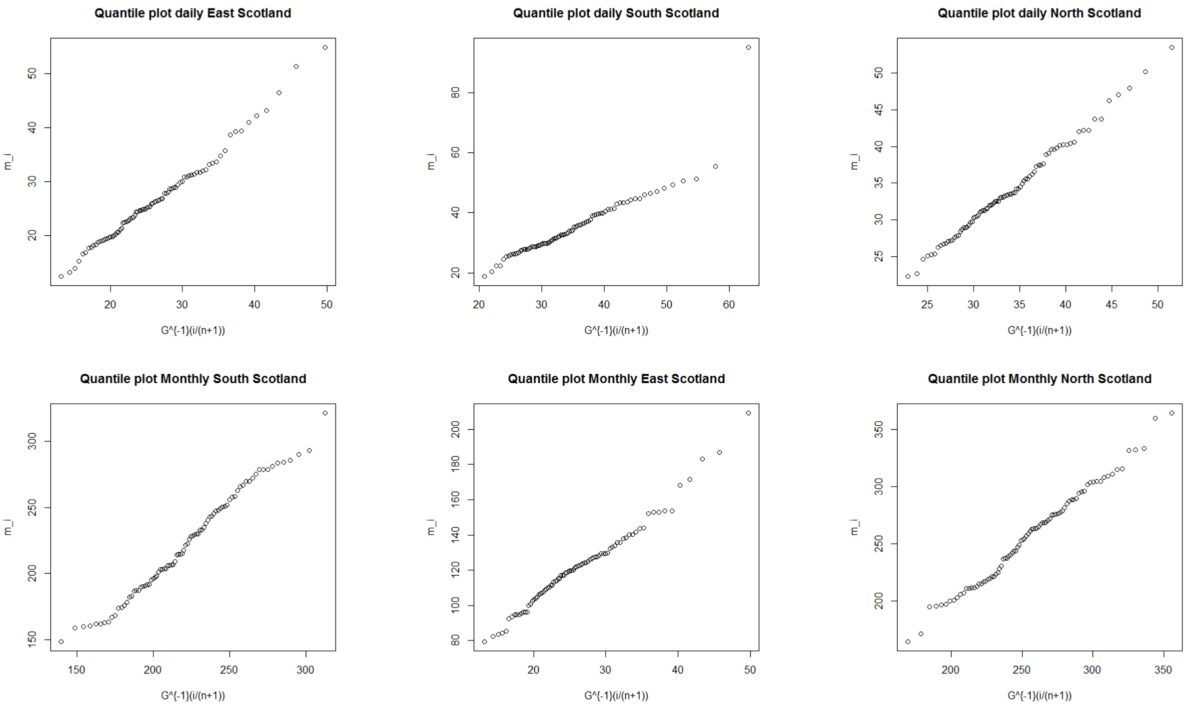


Figure 5: Time series plot for the East Scotland monthly precipitation data.

# 4 Discussion

According to my results shown in section 3.1, there are obvious differences between the regions in Scotland. I found that the daily annual maximum for the South of Scotland is much than the East and North which have similar return levels. For instance, it is expected that the daily annual maxima in South Scotland exceeds 82 mm once every 1000 years, while it is expected to exceed 63 and 61 mm for East and North respectively, over the same return period. However, when looking at Table 3, the most extreme monthly maxima seem to happen in the Northen region while East has the lowest return levels. This could mean that precipitations levels over short periods of time in South Scotland can be very extreme but that there are not a lot of rainy days in a month. On the other hand, in the North of Scotland, it seems like the daily precipitations are not as intense but that it rains more consistently over the whole month. The Eastern region has relatively low return levels for both daily and monthly annual maxima. In practice, my findings would suggest that areas in North and South Scotland should prepare to experience extreme monthly return levels in the North and extreme daily return levels in the South the coming 100 years. In further research on the subject, it would be of interest to study whether there are any temporal trends in the occurrence of extreme precipitations level in Scotland. In other words, to investigate if extreme levels are more frequent now than 50 years ago, for example. This may be more easily answered using the peak over threshold methods since the block maxima methods only tells us the value of the maxima but does not tell us anything about the frequency of extreme events (since it only takes one maxima per block). On the other hand, the peak over threshold method allows us to take into account all the values exceeding the threshold, so that one could take into account the frequency and temporal trends.

# 5 Appendix

Only Daily and Monthly East Scotland is shown. The rest is similar and it just needs to input the new data and change the range of $x_p$ in the sequences and plotting.

## Daily East Scotland

```

datmes <- read.table('C:/Users/user/Downloads/datades.txt', header=TRUE)
########################################
attach(datmes)
fix(datmes)
library(ismev)
library(evd)

############################################# obtain the vector of maxima

datmes$X1931 <- datmes$X1931 * (-1) #Multiply by -1 so that it doesnt return as maximum
years <-NULL
ammes <-NULL
for (i in 1:85)
{
years[i] <- (-1930-i)
ammes[i] <- max(subset(datmes, X1931==years[i]))
}
ammes

plot(ammes,type="l",col=1)
#####################################################
#Time serie splot
vecmes <- as.vector(t(datmes[,-c(1)]))
plot(vecmes,type="l",col=1,ylim=c(2,60))

################################## Likelihood function

llik <- function(theta,X)
{
mu <- theta[1]
sigma <- exp(theta[2])
xi <- theta[3]
n <- length(X)
l <- NULL
if(abs(xi)>5e-2)
{
if(sum(1+xi*(X-mu)/sigma <= 0)==0)  #If there are no zero, do this
{
l <- -n*log(sigma)-(1+(1/xi))*
sum( log( (1+xi*(X-mu)/sigma) ) )-
sum(( ( (1+xi*(X-mu)/sigma)^(-1/xi) ) ))
}
if(sum((1+xi*(X-mu)/sigma) < 0)>0 || xi>1 || xi<(-1))
{
l <- -1e40 # when likelihood is 0, log lik is -inf
}
```

```
}
if(abs(xi)<=5e-2)
{
l <- -n*log(sigma) - sum((X-mu)/sigma)-
sum(exp(-(X-mu)/sigma))
}
#l <- -l
return(l)
}

##################################### maxllik function
ammes
mean(ammes)                    #Compute the values to use as initial values in optim
log(sqrt(var(ammes)))
min(ammes)
max(ammes)

maxllik <- function(theta,X)
    {
        res <- optim(par=theta,fn=llik,X=X,control=list(fnscale=-1))
        return(res)
    }

mleval<-maxllik(c(mean(ammes),log(sqrt(var(ammes))),0.3),ammes)$val
mle<-maxllik(c(mean(ammes),log(sqrt(var(ammes))),0.3),ammes)$par
mle<-c(mle[1],exp(mle[2]),mle[3])
mle
mleval
######################################## GEV density and histogram plot

x <- seq(0,60,len=100)
hist(ammes,seq(0,60,5),col="grey",xlab="Support",freq=F,ylim=c(0,0.06),main="Fitted GEV density to Annual daily maximum
for East Scotland")
lines(x,dgev(x,mle[1],mle[2],mle[3]), type="l",col="darkblue", lwd=2)

########################################### Profile likelihood function
llik_xp <- function(par,x,X,p)
{
sigma <- exp(par[1])
xi <- par[2]
n <- length(X)
l <- NULL
if(abs(xi)>5e-3)
{
mu <- x + (sigma/xi)*(1-(-log(1-p))^(-xi))
if(sum(1+xi*(X-mu)/sigma <= 0) ==0 )
{
l <- -n*log(sigma)-(1+(1/xi))*
sum( log( (1+xi*(X-mu)/sigma) ) )-
sum(( ( (1+xi*(X-mu)/sigma)^(-1/xi) ) ))
}
if(sum((1+xi*(X-mu)/sigma) < 0)>0 || xi>1 || xi<(-1))
{
l <- -1e40
}
}
if(abs(xi)<= 5e-3 )
{
mu <- x +(sigma)*log(-log(1-p))
l <- -n*log(sigma) - sum((X-mu)/sigma)-
sum(exp(-(X-mu)/sigma))
}
return(l)
}

profile_llik_xp <- function(par,x,X,p)
{
    res <- NULL
    for(i in 1:length(x))
    {
    res[i] <- optim(par=par,fn=llik_xp,X=X,x=x[i],p=p,
    control=list(fnscale=-1,
    reltol=1e-14,maxit=10000))$value
```

```r
    }
    return(res)
}

# x is a sequence of return level for a chose p (here p=0.01). The compute the profile
    likelihood
# for the values of xp=x.
x <- NULL
x <- seq(xp(c(mle[1],log(mle[2]),mle[3]),0.01)-40,xp(c(mle[1],log(mle[2]),mle[3]),0.01)
    +40)
mle3<- profile_llik_xp(c(log(mle[2]),mle[3]),x,ammes,0.01)
mle3


# return level function xp
xp <- function(par,p)
  {
    mu     <- par[1]
    sigma  <- exp(par[2])
    xi     <- par[3]
    yp     <- -log(1-p)
    if(abs(xi)>5e-2)
      {
        ret <- mu -(sigma/xi)*(1-yp^(-xi))
      }
    else
      {
        ret <- mu - sigma*(log(yp))
      }
    return(ret)
  }

########################## Normalised profile likelihood and plot

# Compute the normalized profile likelihood
Normalisedprof <-NULL
for(i in 1:length(x))
{
Normalisedprof[i] <- (2*(mleval-mle3[i]))
}
Normalisedprof

# Plot the normalized profile likehood and compute the 95% conf interval for thre return
    level
# for the chosen p from before (0.01)
calphaUp <- qchisq(0.95,1)
plot(x,Normalisedprof,type="l",xlim=c(30,100),ylim=c(0,5),xlab='xp',ylab='Normalized
    profile likelihood')
title('Normalized profile likelihood for return level at
a return period of T=100')
abline(calphaUp,0)
mle3
mleval

conf <- x[Normalisedprof<3.84]
max(conf)
min(conf)
######################################### return level plot

xp1 <- function(par,p)
  {
            ret <- NULL
            mu     <- par[1]
    sigma <- exp(par[2])
    xi     <- par[3]
            for(i in 1:length(p))
    if(abs(xi)>5e-3)
      {
        ret[i] <- mu -(sigma/xi)*(1-(-log(1-p[i]))^(-xi))
      }
    else
      {
        ret[i] <- mu - sigma*(log(-log(1-p[i])))
      }
```

```
    return(ret)
  }


# Plot the return levels for a sequence of p. Simply change the p to a fixed value
# in the xp1() in order to obtain point estimates of return level for given return period.
p <- seq(0,0.99,len=100)
p
returnlevel <- xp1(par=c(mle[1],log(mle[2]), mle[3]),p)
returnlevel
yp <- -log(-log(1-p))
yp

# Plot all the return levels on same plot. The mle values for South and North are taken from
# the exact same code as here but simply changing the dataset and some range for the xp when
# plotting. The values are then copy pasted into the code below.

plot(yp,returnlevel,type="l",xlim=c(-1.6,5),ylim=c(10,150),xlab="-log(-log(1-p))",ylab="Quantile",col=1) # East
lines(yp,xp1(par=c(31,1.93, 0.022),p), type="l",lwd=1,col=2) #South
lines(yp,xp1(par=c(31.1,1.66,-0.063),p), type="l",lwd=1,col=3) #North
title('Return level
for daily East, North and South Scotland')
legend( 'bottomright', inset=0.02,
        legend=c("East","South","North"),
        col=c(1,2,3), lwd=1, lty=c(1,1,1),
         merge=FALSE,cex=0.8)
###################################################################### Quantile plot

orderedammes <- NULL
ordered <- order(ammes)
for (i in 1:85)
{
orderedammes[i] <- ammes[ordered[i]]
}


GevInversed <- NULL
for (i in 1:85)
{
GevInversed[i] <- (mle[1]-(mle[2]/mle[3])*(1-(-log(i/86))^(-mle[3])))
}

plot(GevInversed,orderedammes,xlab="G^{-1}(i/(n+1))",ylab="m_i",col=1)
title('Quantile plot daily East Scotland')
```

# Monthly East Scotland

```

datmes <- read.table('C:/Users/user/Downloads/datames.txt', header=TRUE)
########################################
attach(datmes)
datmes
fix(datmes)
library(ismev)
library(evd)


#############################################
matplot(t(datmes[-c(1,14)]),type="l",col=1)
ammes <- apply(datmes[,-c(1,14)],1,max)

#Annual maxima
ammes
plot(ammes,type="l",col=1,ylab='Annual maxima (mm)',xlab='Years')
title('Time series of annual maxima precipitation')
```

```
#Time serie splot
vecmes <- as.vector(t(datmes[,-c(1,14)]))
plot(vecmes,type="l",col=1,ylab='Monthly precipitation(mm)',xlab='Months')
title('Time series of monthly precipitation')

################################## Likelihood function

llik <- function(theta,X)
{
mu <- theta[1]
sigma <- exp(theta[2])
xi <- theta[3]
n <- length(X)
l <- NULL
if(abs(xi)>5e-2)
{
if(sum(1+xi*(X-mu)/sigma <= 0)==0)  #If there are no zero, do this
{
l <- -n*log(sigma)-(1+(1/xi))*
sum( log( (1+xi*(X-mu)/sigma) ) )-
sum(( ( (1+xi*(X-mu)/sigma)^(-1/xi) ) ))
}
if(sum((1+xi*(X-mu)/sigma) < 0)>0 || xi>1 || xi<(-1))
{
l <- -1e40 # when likelihood is 0, log lik is -inf
}
}
if(abs(xi)<=5e-2)
{
l <- -n*log(sigma) - sum((X-mu)/sigma)-
sum(exp(-(X-mu)/sigma))
}
#l <- -l
return(l)
}

##################################### maxllik function
ammes
mean(ammes)
log(sqrt(var(ammes)))
min(ammes)
max(ammes)

maxllik <- function(theta,X)
    {
        res <- optim(par=theta,fn=llik,X=X,control=list(fnscale=-1))
        return(res)
    }

mleval<-maxllik(c(mean(ammes),log(sqrt(var(ammes))),0.3),ammes)$val
mle<-maxllik(c(mean(ammes),log(sqrt(var(ammes))),0.3),ammes)$par
mle<-c(mle[1],exp(mle[2]),mle[3])
mle
mleval
######################################## GEV density and histogram plot

x <- seq(40,250,len=100)
hist(ammes,seq(40,250,10),col="grey",xlab="Support",freq=F,ylim=c(0,0.025),main="Fitted GEV density to Annual monthly maximum for East Scotland")
lines(x,dgev(x,mle[1],mle[2],mle[3]), type="l",col="darkblue", lwd=2)

########################################### NEW LAB

llik_xp <- function(par,x,X,p)
{
sigma <- exp(par[1])
xi <- par[2]
n <- length(X)
l <- NULL
if(abs(xi)>5e-3)
{
mu <- x + (sigma/xi)*(1-(-log(1-p))^(-xi))
```

```
if(sum(1+xi*(X-mu)/sigma <= 0) ==0 )
{
l <- -n*log(sigma)-(1+(1/xi))*
sum( log( (1+xi*(X-mu)/sigma) ) )-
sum(( ( (1+xi*(X-mu)/sigma)^(-1/xi) ) ))
}
if(sum((1+xi*(X-mu)/sigma) < 0)>0 || xi>1 || xi<(-1))
{
l <- -1e40
}
}
if(abs(xi)<= 5e-3 )
{
mu <- x +(sigma)*log(-log(1-p))
l <- -n*log(sigma) - sum((X-mu)/sigma)-
sum(exp(-(X-mu)/sigma))
}
return(l)
}


profile_llik_xp <- function(par,x,X,p)
{
    res <- NULL
    for(i in 1:length(x))
    {
    res[i] <- optim(par=par,fn=llik_xp,X=X,x=x[i],p=p,
    control=list(fnscale=-1,
    reltol=1e-14,maxit=10000))$value
     }
     return(res)
}


x <- NULL
x <- seq(xp(c(mle[1],log(mle[2]),mle[3]),0.01)-40,xp(c(mle[1],log(mle[2]),mle[3]),0.01)
    +60)
mle3<- profile_llik_xp(c(log(mle[2]),mle[3]),x,ammes,0.01)
mle3


# return level function xp
xp <- function(par,p)
  {
    mu     <- par[1]
    sigma  <- exp(par[2])
    xi     <- par[3]
    yp     <- -log(1-p)
    if(abs(xi)>5e-2)
      {
        ret <- mu -(sigma/xi)*(1-yp^(-xi))
      }
    else
      {
        ret <- mu - sigma*(log(yp))
      }
    return(ret)
  }

########################## Normalised prof lik

Normalisedprof <-NULL
for(i in 1:length(x))
{
Normalisedprof[i] <- (2*(mleval-mle3[i]))
}
Normalisedprof
calphaUp <- qchisq(0.95,1)
plot(x,Normalisedprof,type="l",xlim=c(150,250),ylim=c(0,5),xlab='xp',ylab='Normalized
    profile likelihood')
title('Normalized profile likelihood for return level at
a return period of T=100')
abline(calphaUp,0)
mle3
```

```r
mleval

conf <- x[Normalisedprof <3.84]
max(conf)
min(conf)
######################################### return level plot

xp1 <- function(par,p)
  {
            ret <- NULL
            mu    <- par[1]
    sigma <- exp(par[2])
    xi    <- par[3]
            for(i in 1:length(p))
    if(abs(xi)>5e-3)
      {
        ret[i] <- mu -(sigma/xi)*(1-(-log(1-p[i]))^(-xi))
      }
    else
      {
        ret[i] <- mu - sigma*(log(-log(1-p[i])))
      }
    return(ret)
  }

p <- seq(0,0.99,len=100)
p
returnlevel <- xp1(par=c(mle[1],log(mle[2]), mle[3]),0.0001)
returnlevel
yp <- -log(-log(1-p))
plot(yp,returnlevel,type="l",xlim=c(-1.6,5),ylim=c(78,490),xlab="-log(-log(1-p))",ylab="
    Quantile",col=1) # East
lines(yp,xp1(par=c(206.83,3.6, -0.226),p), type="l",lwd=1,col=2) #South
lines(yp,xp1(par=c(240,3.69,-0.212),p), type="l",lwd=1,col=3) #North
title('Return level
for monthly East, North and South Scotland')
legend( 'bottomright', inset=0.02,
        legend=c("East","South","North"),
        col=c(1,2,3), lwd=1, lty=c(1,1,1),
         merge=FALSE,cex=0.8)

########################################################### Model checking

orderedammes <- NULL
ordered <- order(ammes)
for (i in 1:85)
{
orderedammes[i] <- ammes[ordered[i]]
}


GevInversed <- NULL
for (i in 1:85)
{
GevInversed[i] <- (mle[1]-(mle[2]/mle[3])*(1-(-log(i/86))^(-mle[3])))
}

plot(GevInversed,orderedammes,xlab="G^{-1}(i/(n+1))",ylab="m_i",col=1)
title('Quantile plot Monthly East Scotland')
```

# 6 References

Richard W. Katz (2010) Statistics of extremes in climate change. *Climate change* (2010) **100**, 71-76.

P. Hall and N. Tajvidi (2000) Nonparametric estimation of the dependence function for a multivariate extreme value distribution. *Journal of Multivariate Analysis* **99**, Issue 4, April 2008, 577–588.