

Modeling extreme values with a GEV mixture probability distributions

Application to a rain data

Pascal Alain Dkengne Sielenou

2023-10-26

```
# library(xfun)

path <- ".."

xfun::in_dir(dir = path, expr = source("./src/generate_gev_sample.R"))
xfun::in_dir(dir = path, expr = source("./src/calculate_gev_inverse_cdf.R"))
xfun::in_dir(dir = path, expr = source("./src/estimate_gev_mixture_model_parameters.R"))
xfun::in_dir(dir = path, expr = source("./src/plot_gev_mixture_model_pdf.R"))
xfun::in_dir(dir = path, expr = source("./src/plot_gev_mixture_model_cdf.R"))
xfun::in_dir(dir = path, expr = source("./src/estimate_gev_mixture_model_quantile.R"))

library(readr)

pluie <- xfun::in_dir(dir = path, expr = read_csv("./applications/pluie.csv"))

## Rows: 14623 Columns: 1
## -- Column specification -----
## Delimiter: ","
## dbl (1): x
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
x <- pluie$x
x <- x[!is.na(x)]
n <- length(x)

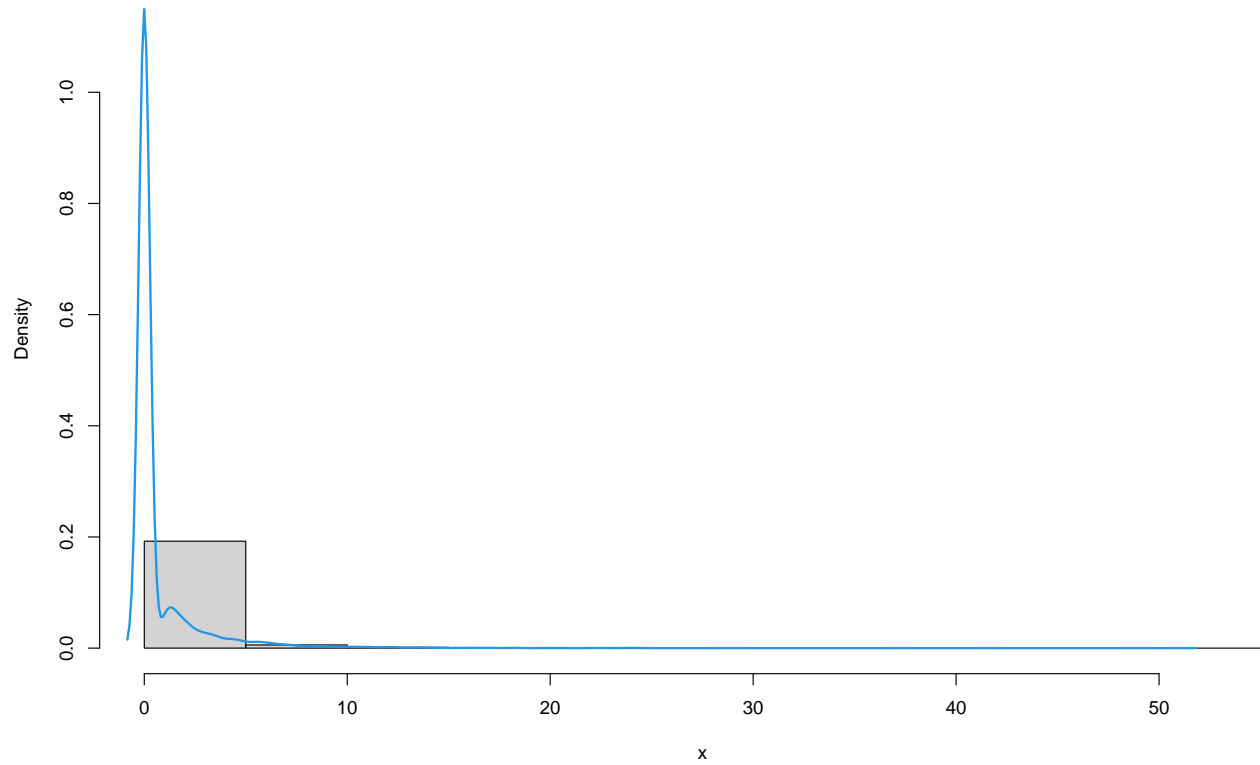
n

## [1] 14623

# Histogram of all data

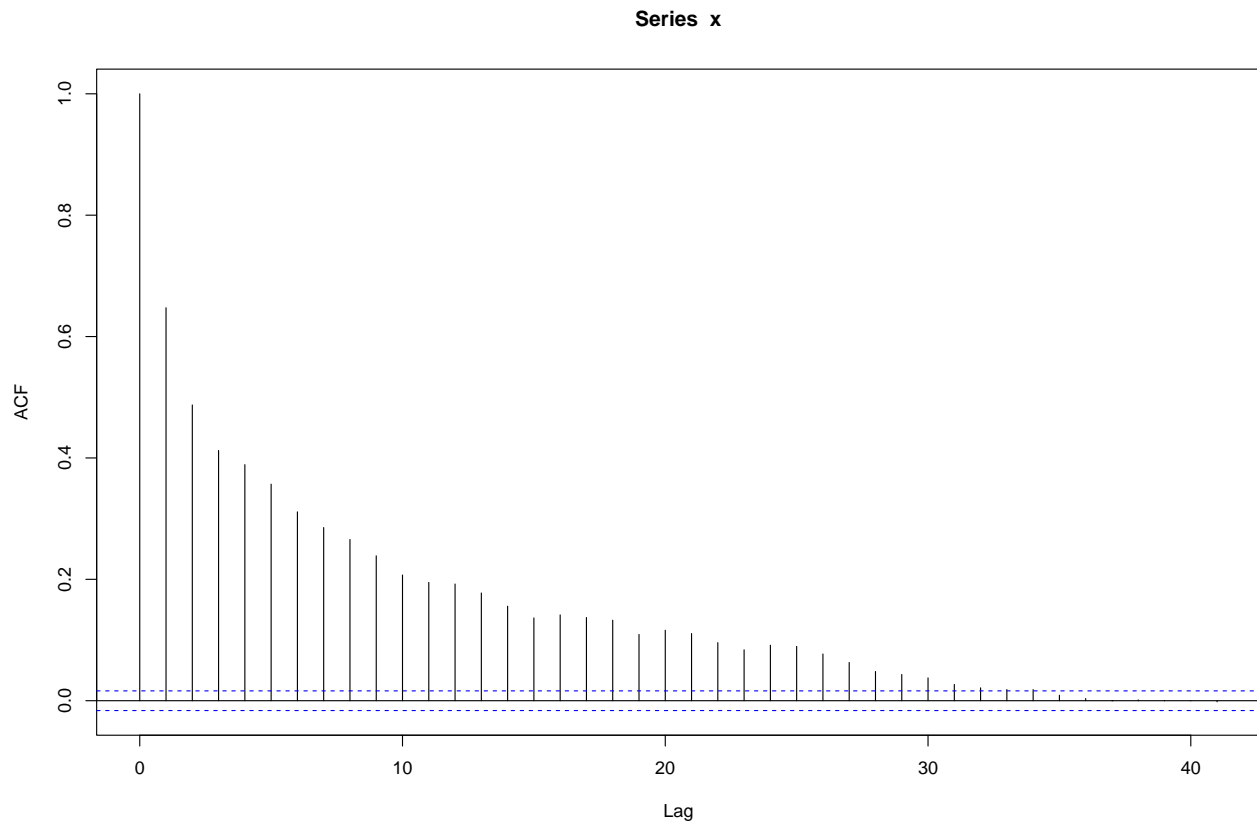
dens_x <- density(x)
hist(x, prob = TRUE, ylim = range(dens_x$y))
lines(dens_x, lwd = 2, col = 4)
```

Histogram of x



```
# Autocorrelation function of all data
```

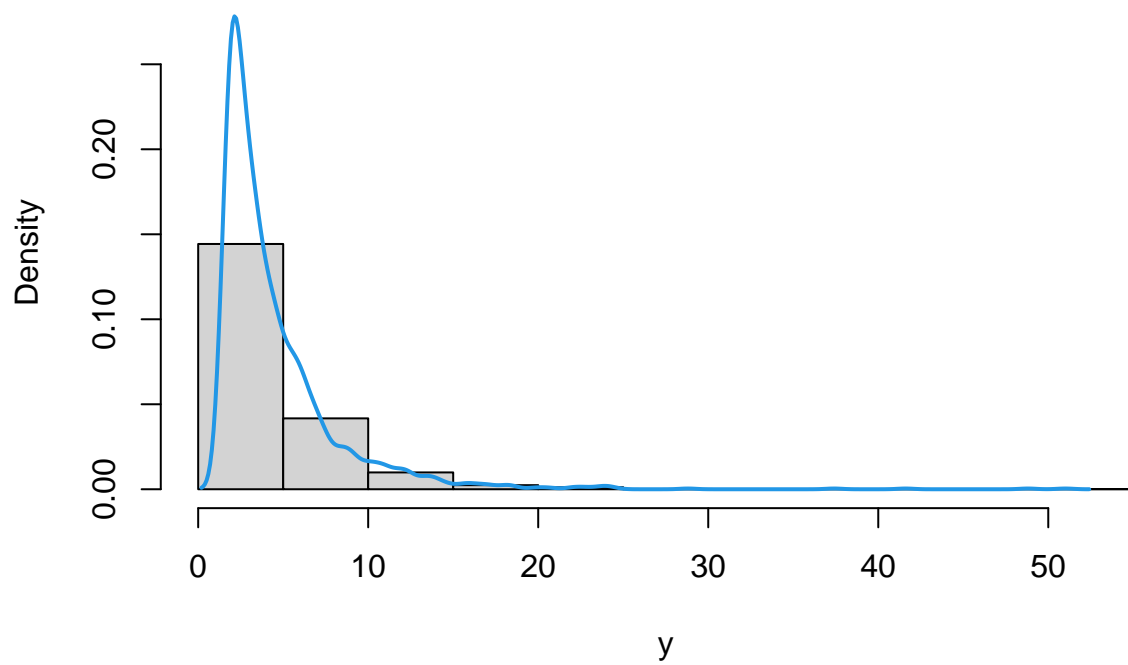
```
acf(x)
```



```
# Histogram of the largest data
```

```
nlargest <- 2000  
y <- extract_nlargest_sample(x, n = nlargest)  
dens_y <- density(y)  
hist(y, prob = TRUE, ylim = range(dens_y$y))  
lines(density(y), lwd = 2, col = 4)
```

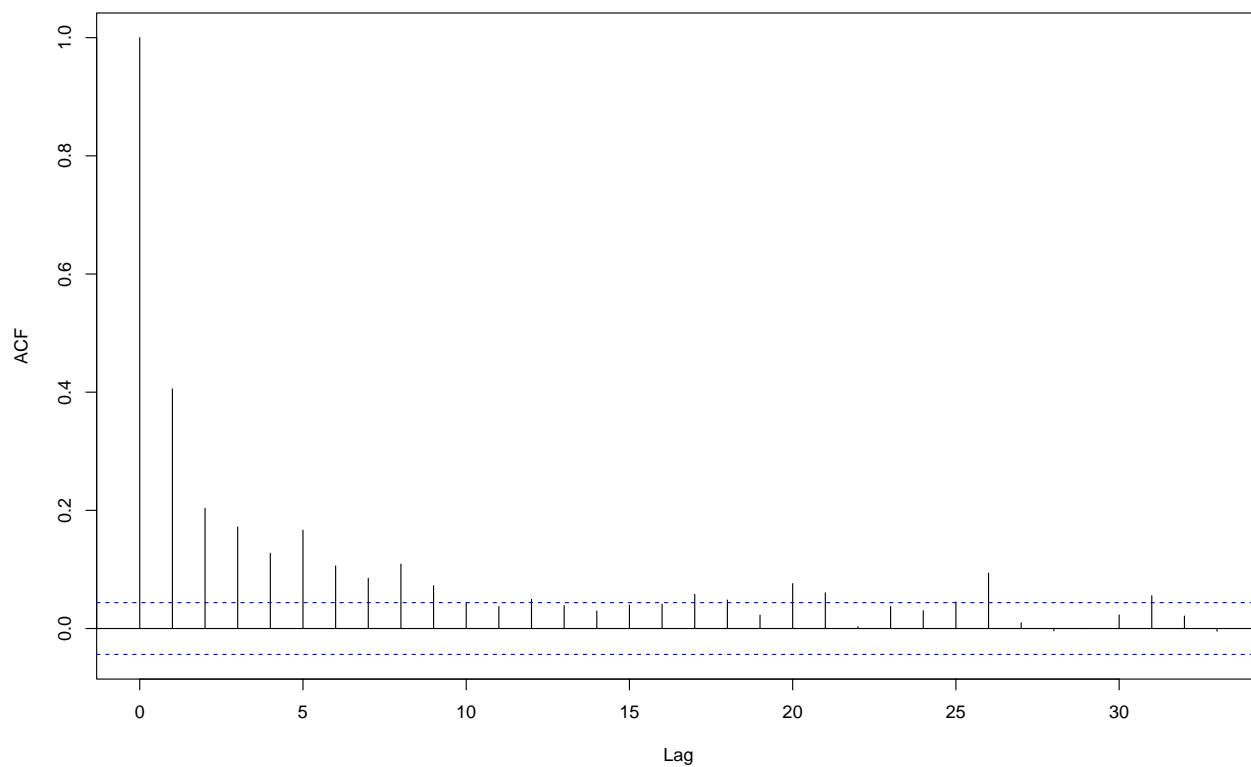
Histogram of y



```
# Autocorrelation function of the largest data
```

```
acf(y)
```

Series y



```
# Estimation of gev mixture models
```

```
gev_mixture_model <- suppressWarnings(estimate_gev_mixture_model_parameters(x = x,  
                                                                           block_sizes = 10:40,  
                                                                           minimum_nblocks = 50,  
                                                                           threshold = NULL,  
                                                                           nlargest = nlargest,  
                                                                           confidence_level = 0.95,  
                                                                           use_extremal_index = TRUE,  
                                                                           use_lower_threshold = FALSE,  
                                                                           maximum_iterations = 1500,  
                                                                           log_mv = TRUE,  
                                                                           log_pw = TRUE,  
                                                                           trace = FALSE,  
                                                                           method = "MLE"))
```

```
## Successful convergence.
```

```
## Successful convergence.
```

```
gev_mixture_model$extremal_indexes
```

```
##          10          11          12          13          14          15  
## 0.2793604651 0.2793604651 0.3413828979 0.2793604651 0.3413828979 0.3069316753  
##          16          17          18          19          20          21  
## 0.3413828979 0.3469704025 0.3413828979 0.3809142450 0.3809142450 0.3809142450  
##          22          23          24          25          26          27  
## 0.3809142450 0.3280070390 0.3469704025 0.3184582985 0.3184582985 0.3413828979  
##          28          29          30          31          32          33  
## 0.2938393366 0.3184582985 0.3184582985 0.3184582985 0.3184582985 0.3184582985  
##          34          35          36          37          38          39  
## 0.3079305589 0.2935114842 0.2938393366 0.3413828979 0.2923565005 0.3079305589  
##          40  
## 0.2923565005
```

```
gev_mixture_model$normalized_gev_parameters_object
```

```
##          loc_star  scale_star  shape_star  
## 10  1.2977551681  1.358834827  0.3717926714  
## 11  1.1759826567  1.415211889  0.3479838639  
## 12  1.4174140625  1.194605799  0.4079956712  
## 13  0.8565862039  1.402068042  0.3664129360  
## 14  1.1569371264  1.364091583  0.3634744334  
## 15  0.6588018056  1.419111332  0.3525851083  
## 16  0.6939011330  1.541530879  0.3180061215  
## 17  0.8164089498  1.484415051  0.3287189325  
## 18  0.5801478438  1.385441221  0.3571152967  
## 19  0.8296480534  1.371627693  0.3538556957  
## 20  0.6496520856  1.454136978  0.3348284867  
## 21  0.7363714978  1.313596445  0.3583401291  
## 22 -0.1188086875  1.820479367  0.2760577105  
## 23  0.4321796555  1.559692390  0.3146385556  
## 24  0.1891199654  1.533880046  0.3195250478  
## 25  0.7094290595  1.333431833  0.3593967718  
## 26  0.8645421099  1.281899269  0.3606150686  
## 27 -0.9548220266  1.823343357  0.2862769747  
## 28  0.5332763059  1.329418645  0.3533977066
```

```
## 29 -0.6555224011 1.803895374 0.2847175776
## 30 0.8369277343 1.224717897 0.3672802508
## 31 -1.1598963312 1.881955037 0.2782003507
## 32 -0.9845403055 1.880675950 0.2624950585
## 33 -1.6226212996 2.052348437 0.2418882121
## 34 0.9388367517 1.298139788 0.3343860366
## 35 0.2914844231 1.365960542 0.3382467378
## 36 -0.1069973355 1.434685957 0.3254873512
## 37 -2.8420314349 2.583474790 0.1855615569
## 38 -0.6166176971 1.774072349 0.2748374215
## 39 -0.2706649938 1.629244995 0.2881871931
## 40 -0.9237856376 1.782973899 0.2744051785
```

```
gev_mixture_model$full_normalized_gev_parameters_object
```

```
##          loc_star  scale_star  shape_star
## 10 -0.08220695918 0.8457750212 0.3717926714
## 11 -0.28153052904 0.9080208185 0.3479838639
## 12 0.37800037152 0.7705295122 0.4079956712
## 13 -0.57179093501 0.8786921810 0.3664129360
## 14 -0.05668031065 0.9229726722 0.3634744334
## 15 -0.71213448837 0.9357396103 0.3525851083
## 16 -0.70941541344 1.0952676270 0.3180061215
## 17 -0.51063121957 1.0481918230 0.3287189325
## 18 -0.65641141070 0.9438469961 0.3571152967
## 19 -0.29182903599 0.9747866376 0.3538556957
## 20 -0.54964515510 1.0525780980 0.3348284867
## 21 -0.33546855648 0.9295131415 0.3583401291
## 22 -1.66129207795 1.3946649338 0.2760577105
## 23 -1.03427168832 1.0982902568 0.3146385556
## 24 -1.18844714355 1.0937128496 0.3195250478
## 25 -0.54155671669 0.8838315834 0.3593967718
## 26 -0.33731673217 0.8484908598 0.3606150686
## 27 -2.64167308736 1.3404367387 0.2862769747
## 28 -0.78832209471 0.8623688015 0.3533977066
## 29 -2.41713994597 1.3023318938 0.2847175776
## 30 -0.30725166316 0.8044834014 0.3672802508
## 31 -3.00424111675 1.3688576709 0.2782003507
## 32 -2.84340729192 1.3927325516 0.2624950585
## 33 -3.67406582654 1.5561281879 0.2418882121
## 34 -0.32502427040 0.8755223101 0.3343860366
## 35 -1.07921106335 0.9023272656 0.3382467378
## 36 -1.55611747490 0.9630156808 0.3254873512
## 37 -5.35927477122 2.1163711972 0.1855615569
## 38 -2.46789240526 1.2652727811 0.2748374215
## 39 -1.89793725206 1.1602859702 0.2881871931
## 40 -2.78481604741 1.2722975176 0.2744051785
```

```
gev_mixture_model$automatic_weights_pw_shape
```

```
## 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35
## 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## 36 37 38 39 40
## 0 0 0 0 0
```

```
gev_mixture_model$automatic_weights_pw_scale
```

```
##           10           11           12           13
## 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00
##           14           15           16           17
## 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00
##           18           19           20           21
## 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00
##           22           23           24           25
## 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00
##           26           27           28           29
## 0.000000000e+00 0.000000000e+00 -1.776356839e-15 0.000000000e+00
##           30           31           32           33
## 0.000000000e+00 0.000000000e+00 0.000000000e+00 2.842170943e-14
##           34           35           36           37
## 0.000000000e+00 3.552713679e-15 0.000000000e+00 1.000000000e+00
##           38           39           40
## 0.000000000e+00 0.000000000e+00 0.000000000e+00
```

```
gev_mixture_model$automatic_weights_pw_loc
```

```
##           10           11           12           13
## 0.000000000e+00 0.000000000e+00 1.000000000e+00 2.220446049e-16
##           14           15           16           17
## 7.216449660e-16 2.220446049e-16 0.000000000e+00 0.000000000e+00
##           18           19           20           21
## 2.220446049e-16 0.000000000e+00 -2.220446049e-16 0.000000000e+00
##           22           23           24           25
## 0.000000000e+00 8.881784197e-16 -8.881784197e-16 0.000000000e+00
##           26           27           28           29
## -2.220446049e-16 8.881784197e-16 8.881784197e-16 8.881784197e-16
##           30           31           32           33
## -1.110223025e-16 0.000000000e+00 3.552713679e-15 3.552713679e-15
##           34           35           36           37
## 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00
##           38           39           40
## 8.881784197e-16 1.776356839e-15 -8.881784197e-16
```

```
gev_mixture_model$weighted_normalized_gev_parameters_object[3, ]
```

```
##           loc_star scale_star  shape_star
## automatic_weights 1.417414062 2.58347479 0.4079956712
```

```
gev_mixture_model$automatic_weights_mw
```

```
##           10           11           12           13           14
## 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00 1.000000000e+00
##           15           16           17           18           19
## 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00 7.771561172e-16
##           20           21           22           23           24
## 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00
##           25           26           27           28           29
## 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00
##           30           31           32           33           34
## 0.000000000e+00 0.000000000e+00 0.000000000e+00 5.551115123e-17 5.551115123e-17
##           35           36           37           38           39
```

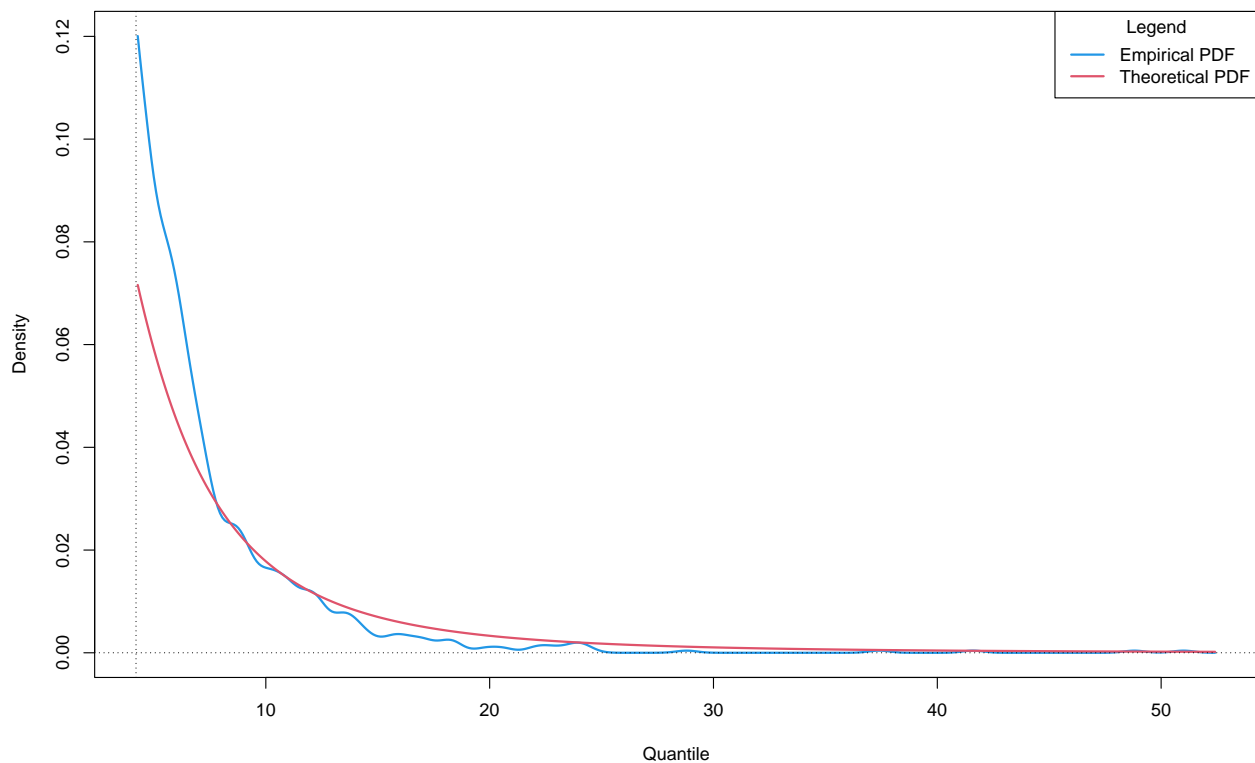
```
## 0.000000000e+00 5.551115123e-17 0.000000000e+00 5.551115123e-17 0.000000000e+00
##          40
## 2.775557562e-17
```

```
# Model diagnostics
```

```
## GEV mixture model with respect to parameters
```

```
plot_gev_mixture_model_pdf(gev_mixture_model,
                             type = "automatic_weights",
                             model_wise = FALSE,
                             zoom = TRUE,
                             xlab = "Quantile",
                             ylab = "Density",
                             main = "Probability Density Function (PDF) Plot")
```

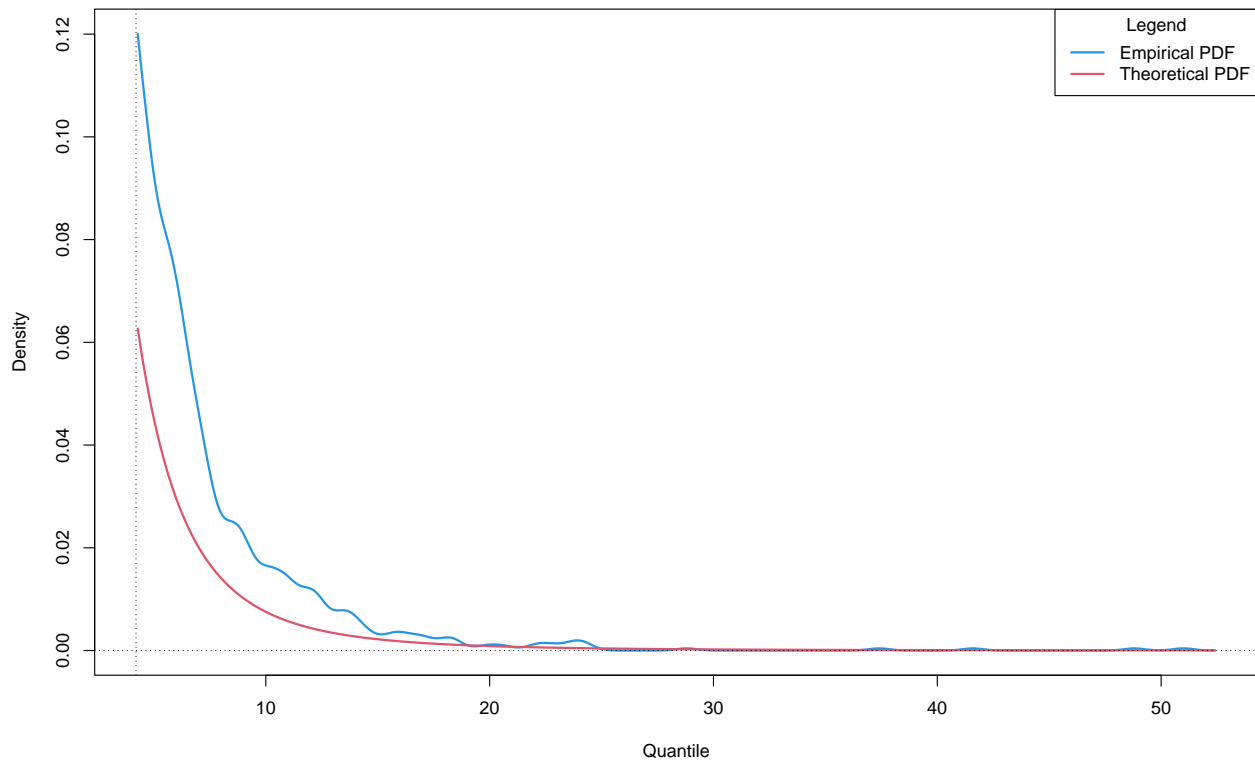
Probability Density Function (PDF) Plot : automatic_weights – model_wise = FALSE : zoom = TRUE



```
## GEV mixture model with respect to distribution functions
```

```
plot_gev_mixture_model_pdf(gev_mixture_model,
                             type = "automatic_weights",
                             model_wise = TRUE,
                             zoom = TRUE,
                             xlab = "Quantile",
                             ylab = "Density",
                             main = "Probability Density Function (PDF) Plot")
```


Probability Density Function (PDF) Plot : automatic_weights – model_wise = TRUE : zoom = TRUE



Estimation of an extreme quantile

```
estimator_types <- c("automatic_weights_mw",
                     "pessimistic_weights_mw",
                     "identic_weights_mw",
                     "automatic_weights_pw",
                     "pessimistic_weights_pw",
                     "identic_weights_pw",
                     "model_wise",
                     "parameter_wise",
                     "empirical")
```

```
alpha <- 10(-6)
```

Quantile from GEV mixture model with respect to parameters

```
rl_pw <- estimate_gev_mixture_model_quantile(gev_mixture_model,
                                             alpha = alpha,
                                             confidence_level = 0.95,
                                             do.ci = TRUE,
                                             estimator_type = estimator_types[4])
```

```
rl_pw
```

```
## [1] 907.2593806
```

Quantile from GEV mixture model with respect to distribution functions

```
rl_mw <- estimate_gev_mixture_model_quantile(gev_mixture_model,
                                             alpha = alpha,
```

```

confidence_level = 0.95,
do.ci = TRUE,
estimator_type = estimator_types[1])

rl_mw

## [1] 210.0716334
## Quantiles from equivalent estimated distributions in GEV mixture model with respect to parameters

est_rl_pw <- suppressWarnings(estimate_gev_mixture_model_quantile(gev_mixture_model,
                                                                    alpha = alpha,
                                                                    confidence_level = 0.95,
                                                                    do.ci = TRUE,
                                                                    estimator_type = estimator_types[8]))

est_rl_pw

##          lower      quantile      upper
## 10 -66.53906710 208.44677401 483.4326151
## 11 -62.02710391 178.06021398 418.1475319
## 12 -177.78991744 270.54289787 718.8757132
## 13 -104.15493132 205.19847922 514.5518898
## 14 -130.02447636 210.06584398 550.1561643
## 15 -117.67636329 191.28707267 500.2505086
## 16 -92.21400015 161.62133835 415.4566769
## 17 -130.45263380 171.16756554 472.7877649
## 18 -171.59559589 201.54076457 574.6771250
## 19 -173.75614232 202.07733538 577.9108131
## 20 -157.19589624 182.01975413 521.2354045
## 21 -194.47754916 201.08278158 596.6431123
## 22 -89.50564054 139.16608209 367.8378047
## 23 -139.78381710 156.73320267 453.2502224
## 24 -154.22164464 163.21719516 480.6560350
## 25 -227.60477817 192.96552425 613.5358267
## 26 -213.34312627 187.47217770 588.2874817
## 27 -139.64231193 145.79294467 431.2282013
## 28 -247.02887746 177.51761461 602.0641067
## 29 -156.68477676 139.78067426 436.2461253
## 30 -247.31667951 189.55345884 626.4235972
## 31 -163.36506822 137.90082040 439.1667090
## 32 -140.63590546 121.71633491 384.0685753
## 33 -115.60452312 112.39747971 340.3994825
## 34 -214.22981974 150.90507027 516.0399603
## 35 -294.56307064 160.52743209 615.6179348
## 36 -248.70042627 151.46409079 551.6286078
## 37 -64.59915337  92.59164309 249.7824396
## 38 -165.09322450 123.89518048 412.8835854
## 39 -165.46395326 128.83451514 423.1329835
## 40 -182.74908250 123.78469112 430.3184647

## Comparison of estimated quantiles

est_rl_pw_range <- range(as.matrix(est_rl_pw))

```

```
## Quantiles from equivalent estimated GEV distributions in GEV mixture model respect to distribution f
```

```
est_rl_mw <- suppressWarnings(estimate_gev_mixture_model_quantile(gev_mixture_model,  
                                                                    alpha = alpha,  
                                                                    confidence_level = 0.95,  
                                                                    do.ci = TRUE,  
                                                                    estimator_type = estimator_types[7]))  
  
est_rl_mw
```

```
##           lower    quantile      upper  
## 14 -130.0244764 210.0658440 550.1561643  
## 19 -173.7561423 202.0773354 577.9108131  
## 33 -115.6045231 112.3974797 340.3994825  
## 34 -214.2298197 150.9050703 516.0399603  
## 36 -248.7004263 151.4640908 551.6286078  
## 38 -165.0932245 123.8951805 412.8835854  
## 40 -182.7490825 123.7846911 430.3184647
```

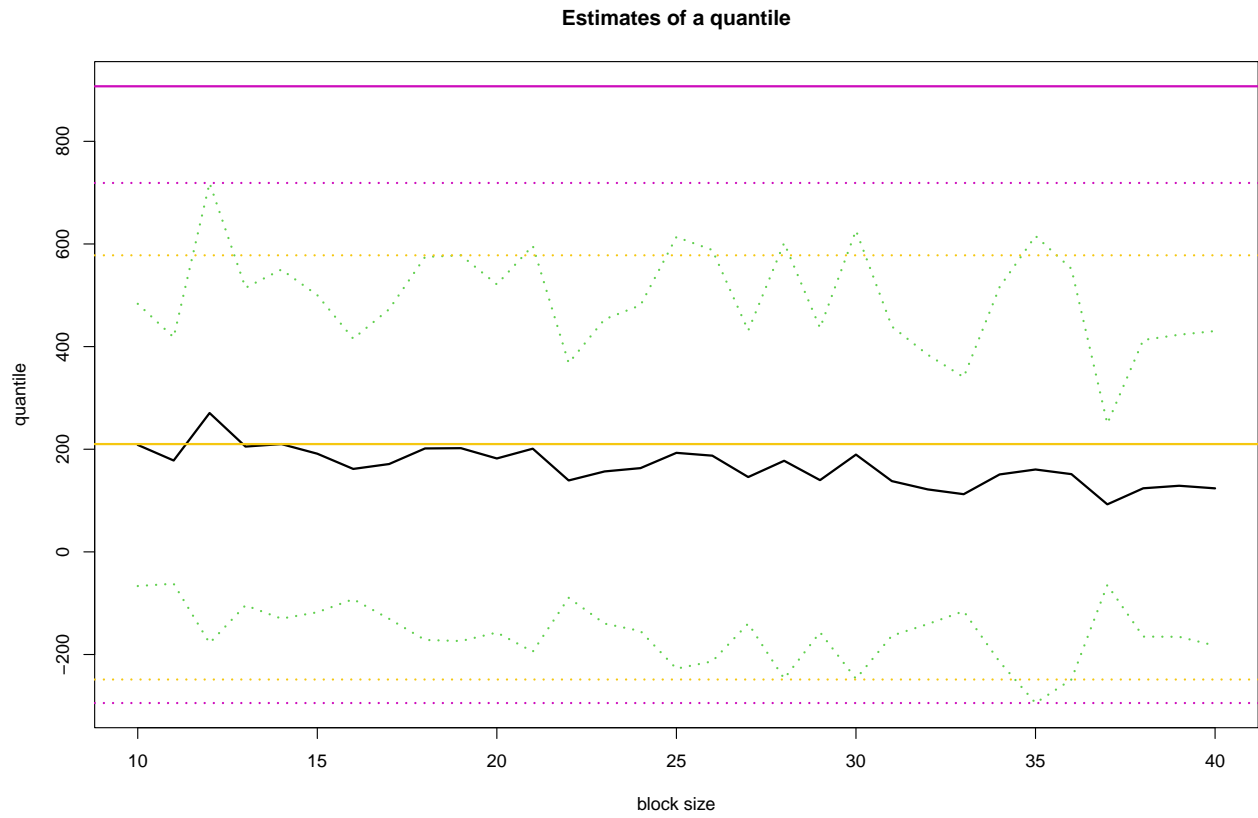
```
est_rl_mw_range <- range(as.matrix(est_rl_mw))
```

```
est_rl_mw_range
```

```
## [1] -248.7004263 577.9108131
```

```
matplot(x = rownames(est_rl_pw),  
        y = est_rl_pw,  
        xlab = "block size",  
        ylab = "quantile",  
        main = "Estimates of a quantile",  
        ylim = range(c(est_rl_pw_range, rl_pw)),  
        cex = 1,  
        cex.lab = 1,  
        cex.axis = 1,  
        type = "l",  
        lty = c("dotted", "solid", "dotted"),  
        lwd = c(2,2,2),  
        col = c(3, 1, 3))
```

```
abline(h = rl_mw, col = 7, lwd = 2)  
abline(h = rl_pw, col = 6, lwd = 2)  
abline(h = est_rl_pw_range, col = 6, lty = "dotted", lwd = 2)  
abline(h = est_rl_mw_range, col = 7, lty = "dotted", lwd = 2)
```



Legend:
yellow: Quantile from GEV mixture model with respect to distribution functions
pink: Quantile from GEV mixture model with respect to parameters